Innovation and Growth: Theoretical Models and Analytical Simulations of Spatial, Clustering and Competition Effects

A thesis submitted

in fulfilment of the

requirements for the degree

of

Doctor of Philosophy (Ph.D.) in Economics

at

The University of Waikato

by

Steven Bond-Smith

KO TE TANGATA

THE UNIVERSITY OF

WAIKATO

Te Whare Wānanga o Waikato

2014

Innovation and Growth: Theoretical Models and Analytical Simulations of Spatial, Clustering and Competition Effects

Steven Bond-Smith

Abstract

This thesis expands upon endogenous growth theory to incorporate a more detailed understanding of innovation. The similarities in modelling techniques between endogenous growth theory and the New Economic Geography (NEG) in the existing literature creates an opportunity to incorporate features of innovation within a model of both geography and growth. The thesis explores three key areas. Firstly, innovation is based on knowledge spillovers. But knowledge spillovers are also subject to spatial characteristics in its transfer between firms, locations, regions and nations. Similarly, knowledge is used by innovators in many sectors, not just the sector where the knowledge was originally developed. As a result, innovating firms have an incentive to agglomerate alongside other innovators to benefit from each other's knowledge. The geographic constraints of knowledge spillovers are incorporated into a regional model of growth with creative destruction. However, spillovers between industries are not equal. Industries or firms producing products which are closely related have a strong incentive to locate in close proximity to benefit from related-technology spillovers. Some of these related innovation clusters could potentially be isolated, but sustainable due to knowledge transfer between related firms within the cluster. It is this innovation clustering which could provide the opportunity for distant or isolated economies to sustain high levels of productivity. Lastly, the thesis considers a way in which firms enter a market such that entry is not necessarily available to all potential entrants. This limit on the market supplied by innovating firms is characterised by discrete instead of free entry. Entry by discrete firms allows entering firms to invest in innovation in response to actual (instead of potential) competitors. They are able to escape competition by sustaining incremental innovations. Greater competition in larger markets has greater investment in innovation by all participants. This thesis provides new insights for endogenous growth theory, emphasising the need for policies regarding innovation and economic growth that are industry and location specific.

Acknowledgements

Thank you to the Royal Society of New Zealand Marsden Fund for financial support in conducting this study. Most of all, thank you to Prof Philip McCann and Prof Les Oxley for offering me this opportunity to pursue my passion for economics with an excellent topic that opened up a range of research questions as part of the wider Marsden Fund project "Is understanding innovation the key to understanding economic growth?". Also thank you to the University of Waikato, for both the excellent undergraduate and masters study that has fuelled my interest in economics and for the support of a Waikato Doctoral Scholarship. The journey I have taken in my tertiary study has been sustained by continued support from the university and many members of the Department of Economics at various times.

I am particularly grateful to Prof Jacques Poot for taking over as chief supervisor on a project perhaps outside Jacques' main fields of research. Nonetheless, Jacques' past research on growth and his knowledge of spatial economics was fundamental to my development as an applied theorist. On that note, I also thank my other supervisors, Prof Les Oxley for his committed supervision over the entire journey of my thesis and Prof Philip McCann for his supervision despite the geographical barriers.

Thank you to my employer during the majority of my Ph.D. study. As a consultant at Covec, I have had the benefit of a wide variety of projects to expand my knowledge of economics beyond the specialities of my Ph.D. Thank you to Dr Tim Denne and my other colleagues who have been a great support. The flexible employment arrangements alongside my research at the University of Waikato allowed me to pursue my Ph.D. from Auckland, to be a teaching assistant for microeconomics for a semester in Hamilton and to provide economic and public policy advice on interesting issues. This experience as an economic consultant is immensely valuable for my future career.

My Ph.D. study also provided me the opportunity to travel and share my research with an international community of scholars. I was able to present at international conferences in Seoul, Miami, Leuven, Melbourne and Canberra. I gave seminars at the University of Groningen, VU University Amsterdam, the London School of Economics and the University of Cambridge. Thank you to all those who attended these seminars and provided helpful comments to improve my research. I am also grateful for the opportunity to network with scholars in similar fields of research and for the discussions which helped to develop the ideas behind the models in this thesis. The greatest opportunity to develop my research was being invited for research visits at the University of Groningen, VU University of Amsterdam, the London School of Economics and the Australian National University. A particular thank you to Prof Philip McCann and Prof Jacques Poot for their support in finding these opportunities in the Netherlands. I would also like to thank my hosts. Prof Philip McCann, thank you for hosting me at the University of Groningen, making me feel a part of the university life while in Groningen and the opportunity to share my research with other Ph.D. students studying similarly specialised topics in economics. These opportunities are rare or unavailable in New Zealand for our field and it was this opportunity which lead me down a path to further pursue research questions in economic geography and visit other leading universities that specialise in our field. Thank you to Prof Henri de Groot at VU University Amsterdam. It was a pleasure sharing an office while I was in Amsterdam and your mentoring while I was there has been very valuable. Thank you to Prof Andrés Rodríguez-Pose for hosting a visit to the London School of Economics. The chance to conduct research at such a prestigious institution and the opportunity to attend seminars with leading academics will forever be in my memory and is an asset to my education. Thank you to Prof Martin Richardson for hosting a visit at the Australian National University, making me feel welcome in Australia and for several discussions regarding my future career in economics.

Lastly, thank you to my partner Dr. Daniela Strube. Your intelligence and the ability to discuss our mutual passion for understanding economic phenomena and its policy implications has been an inspiration in conducting my research and writing this thesis.

Contents

List of figures

List of abbreviations

BRIC	Brasil, Russia, India and China
CES	Constant elasticity of substitution
GDP	Gross domestic product
GPT	General-purpose technology
ICT	Information and communication technology
NEG	New Economic Geography
OECD	Organisation for Economic Cooperation and Development
TFP	Total factor productivity

Chapter 1

Introduction

1.1 Innovation and Growth

Economic growth is what determines the material well-being of billions of people. Endogenous growth theory has become the essential explanation for sustained economic growth. It has provided a plausible and comprehensive account of how technology is the key determinant of economic growth. Through invention and innovation, the world has found alternative resources, renewable resources and increased productivity. With innovation as the engine of endogenous growth many countries have developed policy approaches that focus on stimulating innovation. These policies typically focus on the institutions which govern economic systems, the nature of firm interaction and the protection of intellectual property.

Standard economic growth theory including endogenous growth modelling is typically based on an aggregate view of the economy, implicitly assuming that there would be a single "best practice" policy solution that is universally valid. This is in stark contrast to empirical observations of cross-country differences in terms of innovation and growth outcomes. For example, New Zealand's economic growth has lagged behind the rest of the countries in the Organisation for Economic Co-operation and Development (OECD) despite having what is often considered "best practice" economic policy (McCann, 2009). While Australia has now achieved at least 22 years of growth without recession, despite the global financial crisis, its innovation performance during this period has been much lower than comparable peers (Hollanders and Es-Sadki, 2014; Dutta and Lanvin, 2013; Schwab and Sala-i-Martín, 2013; Andrew et al., 2009; EIU, 2009). Even within the United States, the vast majority of patents are developed in only a handful of places such as Silicon Valley or Boston (Crescenzi and Rodríguez-Pose, 2013). Other locations are unable to achieve the same levels of innovation based economic growth. This suggests that there are certain characteristics about these locations and their economies which make innovation less likely, discourage research and development and lead to a reliance on factor endowments such as minerals or climate.

While the research on innovation has developed numerous insights into the nature of innovation, the implications of these findings for theoretical models of growth have not been consistently incorporated into the endogenous growth literature. As a result, growth policy is developed based on a best practice approach rather than one which incorporates the unique, localised and contextual understanding of the nature of innovation in different locations, regions and nations. This spatial perspective of innovation helps explain how growth can vary between locations, despite having similar institutions and factor endowments.

Incorporating spatial features to understand economic phenomena is the ambition of economic geography. Profit-maximising firms choose certain locations over others, because the characteristics of these locations make them a more advantageous place to produce. The New Economic Geography (NEG) (Krugman, 1991b) takes a theoretical modelling approach based on the same microeconomic foundations as endogenous growth models. The characteristics that make a location attractive for firms emerge endogenously, because firms find it advantageous to locate close to consumers, labour, factors of production and suppliers. The characteristics have circular causality, reinforcing each other such that small differences can lead to diverse economic outcomes. Conveniently, NEG and endogenous growth theory have a lot in common, offering avenues to develop theoretical models of growth that add a spatial dimension to innovation.

There are potentially also many other factors that help explain why some locations and industries are poor performers in terms of innovation. For example, if there are barriers to entry, market participants can sustain a strong market position with less investment in innovation. If there is less threat from competitors, potential or actual, there are less incentives to innovate. There are many obvious examples of this disincentive to innovate in natural monopolies such as telecommunications or electricity networks, but these effects might also be present in other markets where the barriers to entry are not so obvious. In particular, small markets may be too small to sustain anything more than a few firms. These firms experience a privileged position that could be detrimental to economic growth. Examples could include industries such as airlines or supermarket chains where network effects, branding or large fixed costs might make firm entry expensive, yet still a prerequisite to innovation. But since theoretical models of growth aggregate the wider economy with a continuum of firms, most endogenous growth models are not equipped to take account of the unique characteristics of individual industries or the nature of competition.

This thesis addresses the failure of endogenous growth theory to thoroughly understand innovation by considering several characteristics of knowledge and innovation within the theory of growth. Three insightful theoretical models of endogenous growth are obtained which help explain variation in economic growth outcomes and offer policy implications that are based on local characteristics. The thesis explores the outcomes for growth with the use of simulations and discusses the implications for the location of economic activity and for policy approaches to support growth. This introductory chapter explains the motivation behind conducting this research and introduces the reader to endogenous growth theory, the economics of innovation and the NEG and concludes with the objectives of the thesis.

1.2 Motivation

New Zealand's recent economic performance does not live up to its "best practice" reputation in terms of economic policy or the ambitions of its policy-makers. New Zealand's universities continue to fall in international rankings and 24% of university graduates move overseas (Hendy and Callaghan, 2013). It is a paradox why growth in New Zealand has lagged behind the rest of the OECD despite New Zealand often ranking highly on international policy comparisons. McCann (2009) suggests that economic geography could offer an explanation. By this account, New Zealand is simply not an advantageous location for innovation activity, with the exception of a few industries that remain in New Zealand because of factor endowments. Closer economic integration has seen a tide of people shifting to Australia, many New Zealand firms become subsidiaries of Australian parent companies or completely shifted offshore.

On a similar point, Australia has considerably lower rates of innovation than other developed countries.[1] It is perhaps largely the strength of investment in the mining of resources that has seen Australia now achieve at least 22 years of sustained economic growth (Garnaut, 2013). Conditions for innovative activity are slightly better in global cities such as Sydney and Melbourne, but Grimes et al. (2011) find that all Australasian cities are more peripheral than the major cities of peripheral countries in Europe. Australia and New Zealand are not densely populated and far away from the rest of the world. Cities are even far away from each other and even more so if the lack of regional infrastructure connecting these cities is taken into account.

McCann (2009) offers some specific policy tools that focus on making New Zealand, and Auckland in particular, more important to the international economy. New Zealand's location offers pathways between three of the four BRICs countries. Adding Indonesia, this location offers new opportunities in the near future as centres of economic activity shift to Asia and South America. It is possible that Auckland could become a more globally connected city, offering professional services that are useful to the wider South

[1] Schwab and Sala-i-Martín (2013) rank Australia 21st for competitiveness. Dutta and Lanvin (2013) rank Australia 19th for innovation. EIU (2009) rank Australia 20th for innovation performance between 2004 and 2008. Andrew et al. (2009) rank Australia 22nd for innovation. Hollanders and Es-Sadki (2014) note that Australia is performing worse than the EU on seven innovation indicators and only performing better on three indicators of innovation related to the public sector.

East Asia region to facilitate international economic activity. It would require local and national governments to take the opportunity to improve Auckland's connectivity. This policy approach holds the potential of bringing Auckland and New Zealand closer to the rest of the world in terms of spatial economic distance. It is these characteristics and the advantage of agglomeration that have maintained Sydney and Melbourne as centres of knowledge industries that can facilitate business in the wider Australasia and South East Asia regions. Sydney has the potential to remain a financial powerhouse in the so-called "Asian century" if it can remain important for international business over alternatives such as Hong Kong, Singapore or Shanghai.

These characteristics of the New Zealand and Australian economies appear unique. But similar characteristics also exist in other places and on different scales. The southern peripheral countries of Europe have been unable to achieve the economic outcomes of the so-called "Blue Banana" industrial region. Stretching from Birmingham, via London, through the Randstad, Ruhr, Zurich, as far as Milan, Turin and Florence, this area has retained the industrial heart of Europe compared to the southern periphery that includes the economically poorer regions stretching from Portugal, through Sardinia and southern Italy, to Greece. In the United States, the economic fortunes of the "rust belt" have declined, while New York and San Francisco have flourished. While the "rust belt" was once a core industrial manufacturing region it has been unable to retain economic activity in the switch to a services-based economy. What is it about some locations that makes them poor locations for innovation? More importantly, if we properly understand innovation, is there anything that can be done about it?

While the recent economic history of New Zealand is a large motivation for pursuing this thesis, the conclusions of the research here can be applied more generally. Ultimately, this thesis seeks explore if understanding innovation can provide a key to sustaining innovation-based growth in countries, regions or locations that suffer due to being located in the periphery.

1.3 Introduction to the literature

Endogenous growth theory is the starting point of the models developed in this thesis. While neoclassical growth theories recognised the role of technological change to economic growth (Solow, 1957; Swan, 1956), these models assumed that technological change was exogenous and theoretical models could only explain growth from increases in the stock of capital and labour. The core idea of endogenous growth theory is that the incentives for investing in technological change come from economic factors. Innovation in the form of new products, processes, brands, services, and markets embodies technological progress. These innovations are responses to economic incentives, as firms attempt to stay ahead of competition by producing better, cheaper or entirely new products. For example, many innovations are the result of research and development undertaken by profit-seeking firms. Increasing returns to scale provide the profit incentive that can be secured by the successful development of an innovation. As a result, economic policies relating to labour markets, trade, competition, education, taxes and intellectual property can influence both investment in research and development and the rate of innovation by changing the economic incentives of entrepreneurs.

There are two main mechanisms for innovation in endogenous growth models: increasing product variety and increasing quality. Product variety growth models, such as Romer (1990), Grossman and Helpman (1991a, Ch. 3) and Young (1993), offer models where firms earn a monopolistic profit forever, because an innovation leads to a new variety such that there is a continually expanding variety of goods. Innovations build upon existing knowledge and create new knowledge as an input to future innovations. The other type of growth models assume that innovations produce new versions that replace existing varieties by being of higher quality or offering higher productivity. Grossman and Helpman (1991b) and Young (1998) consider models in which each variety is repeatedly replaced by a higher quality version. Aghion and Howitt's (1992) model of cost-reducing innovations suggests a monopolist producer of the intermediate good is repeatedly replaced by firms producing varieties that offer higher productivity for producers of final goods. This group of endogenous growth models embodies Schumpeter's (1934) idea of creative destruction where new versions destroy the value in existing varieties.

While Helpman (1992) finds increasing variety models and quality ladders models are largely equivalent, this may not be the case after accounting for spatial effects on the inputs to innovation. If former varieties are destroyed when they are replaced by new versions and the location of economic activity changes, so does the availability of knowledge in different locations. This idea suggests that innovation is a process that is subject to circular causality based on other factors that affect innovation activities. An additional output from innovation is knowledge, which is an input to future innovations. Knowledge is embedded in firms and their locations such that if the location of those firms changes, so too does the source of knowledge for innovation. Therefore, the way knowledge is used, generated, shared or replaced by increasing variety or quality improvement has implications for innovation and growth in future periods.

The NEG offers an avenue to explore these types of circular causal effects within theoretical models. Furthermore, it shares key features with endogenous growth, in particular the use of increasing returns as an incentive, in this case determining the firm's choice of location (Krugman, 1991b). These similarities in modelling techniques make a fusion of endogenous growth theory and NEG an obvious place to start. A number of models have been developed to combine NEG and growth (for example, Martin and Ottaviano, 1999; Martin, 1999; Baldwin and Forslid, 2000; Baldwin et al., 2001, 2003; Yamamoto, 2003; Baldwin and Martin, 2004). Understanding how space affects growth and growth affects space suggests a strong relationship between the location of economic activity and the mobility of knowledge.

A similar characteristic emerges from the innovation literature. If innovations are largely generated in only a few locations, the resulting knowledge that is an input to future innovation activities is also geographically concentrated. Audretsch and Feldman (1996) find that the concentration of innovation and knowledge leads to an uneven spatial distribution of economic activity. Industries will be particularly concentrated if knowledge plays an important role in production (as measured by research and development activity, concentration of skilled labour and the size of the relevant scientific field). Furthermore, they find that the geographic concentration of innovation is strongly related to the geography of knowledge spillovers and not just the concentration of industrial activity. Florida (1995) suggests that regions are concentration points for knowledge creation and learning, such that regions are becoming more important mechanisms for organising global economic and technological activity. These characteristics of innovation are integral to understanding economic growth outcomes.

One new avenue explored in this thesis is how the spatial characteristics of knowledge spillovers affect quality ladders models of growth. Since these require innovation for production in every period, the interaction between location and knowledge spillovers as the main input to innovation is quite different from two-region product variety models of growth. Furthermore, existing models combining NEG and endogenous growth consider only geographic space and not the interaction between different types of technological knowledge. The use of common knowledge by related industries also has an effect on the location and generation of innovation, such that industries with similar knowledge inputs are also likely to cluster in similar locations. There is circular causality such that when firms choose a geographic location, they also need to consider the technological space that determines the usefulness of knowledge spillovers as an input to innovation. Lastly, the thesis discusses a model in which the ability to innovate is dependent on the possibility of firm entry where each sector is only open to a discrete number of firms. Real world markets are much more disaggregated than standard endogenous growth models appreciate. Each of these niche segments are typically served by only a few firms. As a result, the mode and level of competition has a big impact on the ability to enter the market and the incentives for investing in innovation.

1.4 Thesis objectives

The overarching objective of this thesis is to reconsider the understanding of innovation within theoretical models of economic growth with the motivation to consider policy solutions that support economic growth and the long term well-being of residents in peripheral regions.

As a background, Chapter 2 starts by thoroughly exploring the literature in the fields of endogenous growth, economic geography and innovation, noting the similarities and differences in these three areas. Similarities include circular causality, the fundamental use of Dixit-Stiglitz competition and the key role of knowledge spillovers. Building on this foundation, Chapter 3 explores how the NEG has made a contribution to including spatial factors into theories of economic growth. This group of models offer a key understanding of the spatial nature of knowledge spillovers and concludes that the mobility of knowledge is a stabilising feature of economic integration within product variety models of growth.

Providing a new contribution to the literature, Chapter 4 develops a model with quality ladders and spillovers between varieties such that the mobility of knowledge could be considered destabilising. Increasing knowledge spillovers makes individual varieties more mobile and therefore more likely to shift to a large agglomeration and away from the periphery.

By disaggregating the market into many industrial sectors, where the level of knowledge spillovers is based on the technological relatedness of different industries, Chapter 5 obtains a model that implies clustering as a solution for peripheral regions to retain knowledge industries. Balancing the trade-off between sourcing knowledge from a specialised peripheral cluster and a diversified agglomeration allows clusters of firms within the same technological sector to sustainably produce in the periphery and resist the agglomeration, because they are interdependent on the location of other firms in the cluster. In particular, it is sectors where innovation knowledge is more intensively sourced from within the sector that are likely to sustain locating in the periphery compared to less knowledge intensive sectors.

Previous models of growth in the NEG have typically used increasing product variety as the engine of growth. As far as the author is aware, the models presented in Chapters 4 and 5 are the first endogenous growth models with quality improvements or creative destruction to incorporate two regions and spatial externalities.

Chapter 6 considers how market entry can be a prerequisite for innovation within the endogenous growth model such that the opportunity to innovate is only available to discrete entrants. Small markets are a common characteristic of isolated, distant and peripheral locations. By exploring the effect of discrete entry upon innovation and growth, the model finds that both the mode and level of competition are particularly important for determining economic growth outcomes. As a result, small or distant markets require more thorough monitoring and management of competition for the purpose of encouraging innovation and economic growth.

Lastly Chapter 7 revisits the thesis' objectives and summarises the growth and policy implications of these theoretical models and simulations. Notably, the models here also open up additional areas for future research to better understand innovation as the engine of growth. Most importantly, the thesis concludes that research in this area is fundamentally important, because a better understanding of innovation really is the key to understanding economic growth.

Chapter 2

On the similarities and differences between endogenous growth and New Economic Geography

Endogenous growth theory and NEG both use increasing returns as an incentive for economic responses in models of growth (Romer, 1986, 1990, 1994; Aghion and Howitt, 1992) and trade (Krugman, 1979, 1991b). While growth theorists recognise the significance of innovation and technological change to economic growth, both neoclassical (Solow, 1956; Swan, 1956) and endogenous (new) growth theory fail to understand the spatial properties of knowledge spillovers from the innovation and economic geography literatures as a necessary prerequisite to innovation. NEG attempts to explain the distribution of economic activity through incorporating transport costs in production and trade (Krugman, 1991b; Krugman and Venables, 1995). Despite the geographic nature of innovation neither literature discusses the drivers and factors that contribute to the dissemination of innovations, knowledge and ideas across regions and the influence of economic geography upon growth rates of different cities, regions and nations. This chapter develops an understanding of the background literature in growth, NEG and innovation, leading to a discussion of the similarities and differences. With a better understanding of the factors that drive the spatial distribution of innovation, linking it with endogenous growth and the spatial distribution of economic activity, this chapter offers insights for developing growth theory. New models of endogenous growth should recognise the role of economic geography to both trade (as in NEG) and innovation (as in the systems of innovation literature) and subsequently to economic growth. The models developed in chapters 4 to 6 of this thesis respond to this challenge.

2.1 Introduction

The significance of innovation and technological change to economic growth has long been promoted in the economic literature. But why do we care about growth? The economic growth that has been achieved since the industrial revolution has enabled significantly better living standards. Production grows due to increases in labour, capital and technology. Individuals become better off because as the level of technology increases so does the value of the marginal production of their labour.

Technology improves over time due to new innovations. Innovation is the successful implementation of a new idea. It may be a new product, organisational structure, marketing tool or any other idea that creates value for the firm or agent implementing the innovation. However, in order to develop a new idea, agents and firms have to build upon existing knowledge and innovations. Isaac Newton described his own scientific discoveries with "if I have seen further, it is by standing on the shoulders of giants." The key implication for innovation and growth is that new ideas are based on an enormous foundation of existing knowledge. It is these new innovations that are the dominant part of economic growth.

But different regions have grown at quite different rates. Even when the institutions and economic situation appear the same, some regions continue to grow faster than others. Agglomeration and the importance of cities plus the geographic nature of knowledge transfer at least partially explain the divergence in growth rates for some regions (Audretsch and Feldman, 2004; Baldwin and Martin, 2004). Total factor productivity (TFP) is a variable in the production function that is not dependent upon the factors of production. It is the growth in TFP that is related to changes in technology or ways of doing things and that results in real increases in earnings per capita beyond merely increasing the inputs to production. In the models developed in this thesis growth refers to improvements in TFP and it is the geographic nature of changes in TFP that is important to variation in economic outcomes.

In this chapter, the background literature in endogenous growth theory and economic geography is explored, paying particular attention to the role of innovation. Section 2.2 reviews growth models, Section 2.3 reviews models in the NEG and Section 2.4 summarises the similarities and differences between the two fields. Chapter 3 follows up this discussion by examining a range of models that combine endogenous growth with space using insights and techniques from the NEG. Subsequent chapters develop endogenous growth models that incorporate a better understanding of the spatial nature of innovation, the tendency for technologically compatible firms to cluster in the same region and competition effects on innovation to gain key insights into understanding economic growth.

2.1.1 Growth

Both neoclassical growth theory (Solow, 1956, 1957; Swan, 1956) and endogenous growth (Romer, 1986, 1990, 1994; Grossman and Helpman, 1991b, 1994; Aghion and Howitt, 1992) recognise the role of innovation to long-run economic growth. Neoclassical growth treats innovations as exogenous to the economy. While neoclassical models recognise the contribution to growth from labour and capital, Solow (1956) realised the relatively small contribution of these to long-run growth.

Innovation was endogenised in economic growth models by Romer (1986, 1990), Grossman and Helpman (1991b) and Aghion and Howitt (1992). This triggered a wave of models that account for the economic signals which affect innovation activity. Endogenous growth models rely on innovations being built upon existing knowledge. But for new companies to research enhancements to existing knowledge, endogenous growth models usually assume a frictionless spillover of knowledge to all firms and other agents.

The reality is that knowledge is not transferred so effortlessly. While some knowledge can be codified and transferred easily, much knowledge is at least partially tacit and spillovers of tacit knowledge occur over space and time through face-to-face contact (McCann, 2007), interaction and migration (Faggian and McCann, 2009). This restriction of knowledge only being partially able to be codified means knowledge and innovation has space, time and cost characteristics in its spillover between agents and firms. Cities are a conduit for knowledge transfers (Glaeser et al., 1992) because the density of firms and population facilitates interaction. This role of space and time in knowledge spillovers means that economic growth also has space and time characteristics, contributing to a wide variance in economic fortunes between cities, regions and nations.

There is a need for new endogenous growth models that include the local economic geography properties of the region and its firms in the knowledge spillover and innovation process, and therefore the varying growth outcomes for different regions and nations. Growth theory can benefit from even further incorporating other factors that affect innovation activity by participating firms. The incentives and ability for firms to invest in innovation are affected by many kinds of factors, including the availability of knowledge, the ability to transfer knowledge between industries, competition, as well as research and development infrastructure, human capital, liquidity of capital markets and many more. The aim of this thesis is to explore some of these by developing new theoretical growth models that factor in additional complexity of innovation and derive implications for innovation and growth policy for peripheral regions.

2.1.2 Innovation

Schumpeter (1934) provides an early serious examination of innovation by defining five types of innovation: new products, new production processes, new markets, new sources of supply for inputs and new market structures. He also describes three phases of technological change: invention, innovation and dispersion of innovation. While invention is the discovery of an idea, innovation is the actual implementation of the idea. The dispersion of new technologies or ideas and the implementation of these ideas is known as adoption. This is particularly important in the context of general purpose technologies (GPTs) such as information and communication technologies (ICT), where the benefits of technology can be achieved by many industries. Finally, the dispersion of innovation should also include the spread of the knowledge that produced the technology, through which further innovations can contribute to economic growth. Each of these stages has implications for growth. In particular, the dispersion of innovation and knowledge has geographic properties.

An important contribution of Schumpeter is that he recognised the possible destructive nature of innovation when new innovations replace existing ones. New ideas and new technologies make old ideas and technologies obsolete. The growth effect of innovations is therefore dependent on the destructive effect of each innovation. Models which recognise creative destruction "as the norm of economic growth" (McCraw, 2007) realise the sporadic nature of economic growth outcomes. This idea of "creative destruction" is expanded upon with a model by Aghion and Howitt (1992) that includes creative destruction within endogenous growth theory.

2.1.3 Neoclassical growth to endogenous growth

Initial theories of growth were developed in the 1950s. However, economists found that a portion of growth could not be credited to increases in the stock of capital and labour (Solow, 1956). Accordingly, the additional economic growth was credited to technological change or TFP growth (Abramovitz, 1956; Kendrick, 1956; Solow, 1956, 1957; Swan, 1956). However, the drivers of change are exogenous to the models that were developed. Economists agreed that technological progress and innovation made a significant contribution to economic growth but included no drivers for innovation in economic models. Neoclassical growth theories fail to explain what technological progress actually is, the innovation drivers and the incentives for individuals and firms to innovate. It leaves these outside the growth equation.

Endogenous growth solves this problem by making innovation endogenous in the growth process. Romer (1986, 1990), Lucas (1988), Aghion and Howitt (1992) and Barro (1990) treat investment in innovation as generating a type of capital. These investments may be in research and development for new products, higher quality products, human capital, or other initiatives that assist with innovation. Empirical studies also suggest that technological spillovers between regions or industries are an important component of the growth process (Coe and Helpman, 1995). All of these studies of growth have different implications from a modelling perspective. Solow and others (Abramovitz, 1956; Kendrick, 1956) have also highlighted the importance of innovation drivers such as stimulating investment, capital accumulation and workforce participation. Endogenous growth has implications for policy from the perspective of improving public infrastructure (Barro, 1990), research and development (Romer, 1990; Grossman and Helpman, 1991b, 1994; Aghion and Howitt, 1992) and education or human capital (Lucas, 1988).

2.1.4 Economic geography and growth

The problem with endogenous growth theory is that there is no allowance for the role of geography in knowledge spillovers, innovation and subsequent economic growth. Techniques from NEG may offer a solution to model the geographic nature of knowledge spillovers and innovation.

Krugman (1991b) explains how economic activity concentrates in cities and specific locations. This concentration of activity results in more knowledge spillovers between firms in similar locations. Agglomeration economies in innovation are just one example of the geographic nature of knowledge spillovers (Audretsch and Feldman, 1996). Face-to-face interaction, communication and migration are extremely important to the creation and transfer of ideas because of the partially codified nature of knowledge. Ideas from economic geography offer solutions to understanding knowledge spillovers across regions, the diffusion of ideas and knowledge on which new innovations are built, the diffusion of GPTs and consequently spatial implications for economic growth. Models of innovation that include the transfer of knowledge, ideas and technologies across regions, between firms and individuals and between economies or industries could be the solution to better explaining variation of economic growth in different countries, regions, cities, industries and even firms.

The remainder of the chapter is structured as follows. Two sections discuss the literature across the broad themes of endogenous growth and economic geography. Within each broad theme, the role of knowledge and innovation is examined and theories are compared. The sections discuss how these should be brought together to form new models of economic growth and outlines the theoretical background for the thesis. The models developed in this thesis integrate the features of innovation theory into endogenous economic growth and take into account the geographic nature of the diffusion of technology and knowledge spillovers which affect innovation and economic growth.

2.2 Growth models

Theories of economic growth all highlight the importance of technological change as the major contributor to economic growth. Neoclassical theories of economic growth rely on changes to inputs such as capital and labour. Solow (1956) and others (Abramovitz, 1956; Kendrick, 1956) take away the weighted growth rate of capital and labour from the overall economic growth rate. This leaves a large portion of growth that can not be explained by either capital or labour. Solow (1956) describes this residual as the contribution of "technological progress" to economic growth but keeps it exogenous to the growth of production. The contributions of Solow (1956, 1957), Swan (1956), Abramovitz (1956) and Kendrick (1956) to neoclassical growth theory showed the importance of innovation but failed to include innovation within the overall theoretical models.

New growth theory endogenises innovation by adding increasing returns as an incentive for investment in innovation. New developments in economics allowed the macroeconomic phenomenon of economic growth to be explored from an industrial organisation perspective (Dixit and Stiglitz, 1977), because those models allow for increasing returns and the competitive nature of markets. While neoclassical models of growth assume perfect competition, industrial organisation models of oligopoly and monopoly allow firms to earn greater than normal profits. Firms differentiate their products and are able to gain returns over their competitors, because they can price above marginal cost. It is these additional returns that encourage firms to innovate by providing a return on investment in innovation activities.

Schumpeter promoted the idea of creative destruction as the norm of capitalism (McCraw, 2007), moving away from the static economic models which focus on equilibrium. By its very nature, the economy is not static and economic growth relies on the constant change of technology to improve productivity. The destruction part of the concept is that old ideas are destroyed and fail, because they are replaced by new ideas and innovations that prove more productive and create more benefits for users and consumers than the old ideas. Schumpeter (1934) particularly notes that innovators have to battle "resistance to new ways" in order to actually achieve innovation. People are scared to accept new innovations for fear of their own skills becoming redundant. This resistance is inherent in the idea of creative destruction, because those who do not adapt to innovation are harmed as their skills, businesses and products become obsolete.

This is particularly important to the modern study of endogenous growth. Aghion and Howitt (1992) explore endogenous growth theories known as Schumpeterian growth, because they take account of the destructive nature of new innovations. While some may assume modern society is now more adaptable to new technologies and innovation, this may actually not be the case. Modern society has encouraged innovation and invention with incentives such as patents, copyright and tax incentives. However, other factors may restrict research and development, invention and ultimately innovation occurring in the first place. In the United States for example it often takes six to eight years and many hundreds of millions of dollars of investment for the Federal Drug Administration to approve a new drug or treatment (Eaton and Kennedy, 2007). The ability of regions to grasp onto new technologies and ideas also has implications for economic growth and the geography of growth.

2.2.1 What is innovation?

As innovation is integral to endogenous growth models, the meaning of innovation must be thoroughly understood. Innovation needs to be considered distinct from invention. While invention is an idea made in manifest, innovation is an idea that is applied successfully (McKeown, 2008). While many definitions consider high-technology innovation, there are many more forms of innovation including innovations in processes, marketing innovations, service innovations and innovations in low-technology products. The discussion of innovation here develops a broad understanding of innovation in the context of examining growth.

Modern theories and definitions of innovation are greatly influenced by Schumpeter (1934). His work expanded beyond inventions, which he described as "radical" innovations that create major disruptive change, to also include "incremental" innovations. Examples of more radical innovations are GPTs such as computers, while examples of incremental innovations are the increasing speeds and improvements in software that have led to the sustained economic growth contributions of ICT. GPTs are a particular category of innovations that yield growth in other sectors, increasing productivity across the economy. GPTs are explored in Lipsey et al. (1998b,a) where a radical innovation is eventually useful in many other sectors beyond the original invention.

Ideas build upon existing knowledge. As a result of the non-rival nature of knowledge, firms are unable to extract the full value of the new knowledge they create when they develop through innovative activities. These benefits can accrue to other firms developing later ideas or to other agents that are integral in the comercialisation process of innovating (Teece, 1986). These firm level interactions have important implications for the incentives to undertake innovation activities. Industrial organisation theory also considers a firm's market position and the need to innovate to stay ahead of competitors (Tirole, 1988). There are many motives for firms to innovate. These microeconomic foundations are explored extensively in Dosi (1988). Understanding these incentives is important in the study of innovation, but so is understanding how these incentives vary across cities, regions, nations, firms and markets. Although endogenous growth is a macro model, the characteristics of the micro markets within the macro-economy are important for understanding how firms innovate and the overall economic growth outcomes.

Schumpeter treats economic development as a process of improving quality through innovation (Fagerberg, 2005). While initially focused on the role of the entrepreneur in the process of transition from invention to innovation, Schumpeter later focused on innovation within the firm, where there is no specific innovator. Notably Schumpeter recognised that innovations "cluster" within industries, time periods and countries (Fagerberg, 2005). Historical "clusters" of innovation can be identified with economic growth in geographic areas during the industrial revolution in the United Kingdom, then continental Europe and later the United States. Modern clusters have occurred for example in the internet industry and computer technology in Silicon Valley and more recently in biotechnology in Boston. This clustering of innovative activity shows how innovation has strong properties of agglomeration economies. This highlights the importance of cities to the innovative economy. Furthermore, it may be evidence of the spatial properties of knowledge spillovers which are explored in more depth in two of the models developed in this thesis. Firms benefit from knowledge transferred between them when they undertake economic activities in similar locations. The long history of economic thought on innovation shows a need to include the geographic properties of knowledge and technology.

Kline and Rosenberg (1986) point out the continuous nature of innovation and invention. Specifically they note that it is almost impossible to define a specific innovation: inventions go through many significant improvements beyond just the original single innovation. These incremental improvements may prove significantly more influential upon economic growth than the original invention itself. This sustained nature of the innovation process is the driver behind continuing growth. A new innovation can provide growth over a long period because of both the distribution of knowledge and technology over time, and because of continued innovations developing the initial invention. Continued innovation is integral to continued economic growth, yet innovation is a stochastic process which occurs intermittently (Aghion and Howitt, 1992). This intermittent nature of innovation also has an influence upon the decision to invest, produce or research within an endogenous growth model.

Economic studies predominantly focus on product and process innovation. However, other innovations also contribute considerably to economic growth. Marketing theories (Hunt, 1983) look into consumer behaviour and market exchanges. Subsequent research into innovation has led to recent standards (OECD/Eurostat, 2005) which expand the definition of innovation considerably to include marketing and other innovations, including relationship-based innovations. The OECD guidelines for collecting innovation data contained in the third edition of the OECD/Eurostat (2005) manual state: "An innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations". While traditional innovation thinking focuses on technological change, this is a much broader definition that includes service innovations, marketing, business, industry and market structures.

All types of innovations - product and process innovations, marketing, service, institutional and organisational innovations - have geographic implications for economic growth. A marketing innovation is a significant change in design, packaging, placement, promotion or pricing of a product. Examples include new media such as web-based "viral" campaigns or better price differentiation. The objective of a marketing innovation is to increase a firm's sales. Service industries and products can also create economic growth through innovation. A service innovation is the same as a product innovation, but for service products. These types of innovations are often overlooked but are also be a significant part of economic growth, particularly in developed countries that have moved beyond traditional manufacturing. Institutional innovations such as property rights, prizes or auctions also provide significant implications for economic growth. For example, government auctions of radio bandwidth could lead to a more efficient use of the available spectrum by the innovators who can make the most benefit out of its use. An organisational innovation is a change to a firm's business practice, organisation or external relations which improves sales for the company or industry. These types of innovations are about reducing transaction and administration costs (OECD/Eurostat, 2005). All of these definitions have implications for modelling the influence of innovation and geography upon economic growth.

While Schumpeter's definition of innovation applies only to the first time an innovation is implemented, the OECD guidelines define an innovation by each time an innovation is implemented in a new market or context. Schumpeter would have alternatively described this kind of innovator as an imitator (Fagerberg et al., 2005). For the purposes of this thesis, the distinction between an innovator or an imitator is not important. The issue at stake here is how an economy, either regional or national, achieves economic growth through innovation, whether developed new, or imitated from other countries or markets. Therefore, whether an innovation is actually new or not does not matter, what is important is its effect on economic growth. Imitations which borrow ideas from existing markets and thereby achieve economic growth through innovation are still of interest. This broader definition captures all kinds of innovations, and therefore the widest portion of economic growth that can be attributed to innovation.

Some of the growth literature does distinguish between innovation and imitation, such that firms face different costs depending upon their chosen strategy. Each firm's individual circumstance determines whether an imitation or innovation strategy is selected in order to maximise profits. Several theorists in the growth literature explore the implications of imitation or innovation based growth. For example, Aghion et al. (2001) explore a model where lagging firms must first catch up with the technology leader before pursuing an innovation strategy. Similarly, Acemoglu et al. (2006) consider the different strategies of firms in economies with varying technology levels. They find that when economies are lagging behind technologically, firms pursue an imitation (investment) strategy but as the economy catches up firms switch to an innovation strategy. Institutions can be used to encourage the appropriate strategy but economies may be prone to a low growth trap (Nijkamp and Poot, 1993) if institutions support an imitation strategy for too long. In this thesis however, a broader definition of innovation is used which may still capture some of the implications for imitation based growth.

In economic models, a factor, typically A, represents the current level of technology. Innovations in this modelling context are the events which cause improvements in A. They are the events which create improvements in technology and therefore a downstream increase in production or quality. This chapter explains the nature of these models and proposes that these models could be improved by a better understanding of the factors that influence innovation. Understanding the factors that affect the rate of technological change will provide a better understanding of economic growth. The next section discusses neoclassical growth theory and developments leading to endogenous growth.

2.2.2 The Solow-Swan model

Solow (1956) developed the neoclassical growth model for which he won the Nobel Prize in 1987. It is expanded empirically in Solow (1957). At the same time, Swan (1956) also developed a similar growth model. Hence, the model is commonly referred to as the Solow-Swan model. The model emphasises the role of capital accumulation and how increased savings can raise an economy's growth rate. However, it also shows that this increase in growth cannot last indefinitely and that a country's growth rate will return to the exogenous rate of technological change. This highlights the importance of technological change as the major contributing factor to economic growth. Essentially the idea was to explain the effects of changes in capital stock and the supply of labour on economic growth. The remaining growth was assumed to be driven by changes in technology (Solow, 1957). Innovation is the cause of technological change, but the rate is exogenous to the model and not affected by economic forces.

In the Solow model there is only one commodity with a rate of production $Y(t)$, where output is produced with two factors of production: capital (K) and labour (L):

$$Y(t) = F(K(t), L(t)). \tag{2.1}$$

Assuming that capital and labour are supplied inelastically; there is no unemployment or underemployment of capital or labour. The function has diminishing returns to capital (holding labour constant). It is captured by the conditions that the marginal product of capital is positive but strictly decreasing in the stock of capital. $F^{'}(K) > 0$ and $F^{''}(K) < 0$ for all K and the Inada conditions $\lim_{K,L\to\infty} F^{'}(K) \equiv 0$ and $\lim_{K,L\to 0} F^{'}(K) \equiv \infty$. These define the ends of the production curve such that diminishing returns diminishes towards zero and operates over all positive levels of production.

Over time, a portion of output is consumed and the remaining portion is saved and invested. The portion that is saved is a constant s so that the rate of savings is $sY(t)$. The stock of capital is the total amount invested and it grows at the savings rate: The increase in the capital stock includes a depreciation rate, δ, on existing capital. Solow takes depreciation away from net output but for ease of understanding, here it is included it in the differential:

$$\frac{dK}{dt} = sY - \delta K. \tag{2.1a}$$

The aggregate production function and the capital accumulation equation can be combined because production is a function of capital:

$$\frac{dK}{dt} = sF(K, L) - \delta K. \tag{2.1b}$$

This final equation is fundamental to neoclassical growth theory. As capital accumulates so does national income and the net amount of depreciation increases. However, because of diminishing marginal returns, the amount of additional saving will diminish. Eventually, the steady state is where savings equals depreciation. The reinvestment is equal to the fall in value of existing capital. Thus the model reaches a steady state where there is no further growth. Growth beyond changes in capital and human capital is assumed to be due to technological change, exogenous to the model.

If growth in the labour supply is included or population growth with a constant labour force participation rate, there is a similar outcome. In the equation below L represents the labour supply. If full employment is assumed and the supply of labour increases at a constant relative rate n, then

$$L(t) \equiv L_0 e^{nt}. \tag{2.2}$$

Capital accumulation with growth in the supply of labour is

$$\frac{dK}{dt} \equiv sF(K, L_0 e^{nt}) - \delta K. \tag{2.2a}$$

Since labour is supplied inelastically, the labour supply curve is a vertical line which shifts to the right in time as the population grows. As both labour and capital increase, wage rates and rates of return on capital adjust to ensure full employment. An additional assumption is included that aggregate production exhibits constant returns to scale. This assumption does not change any of the other assumptions. There are still diminishing returns in each capital and labour when the other is held constant. In the special case of this aggregate production function, for a single unit of labour where k is the capital per person, the per capita production function is given by:

$$Y/L = F(K, L)/L = F(K/L, 1) = F(k, 1) = f(k) = y. \tag{2.2b}$$

The rate of capital accumulation per capita is then given by

$$\frac{dk}{dt} \equiv sf(k) - (n + \delta)k. \tag{2.2c}$$

Diminishing returns will impose an upper limit on capital accumulation and growth in per capita income will be zero. The steady state k^* will be such that

$$sf(k^*) = (n + \delta)k^*. \tag{2.2d}$$

The level of capital per worker will be steady at k^* and the income per capita steady at $y^* = f(k^*)$. This can be seen in Figure 2.1 below, reproduced from Aghion and Howitt (2009). In the short run the growth rate can be increased by increasing the savings rate or lowering population growth or the depreciation rate. However, this is short term, as higher levels of depreciation at a higher level of capital consumes more capital and the growth rate per worker would fall to the steady state growth rate. A higher savings rate will nonetheless yield a higher income per capita in the steady state.

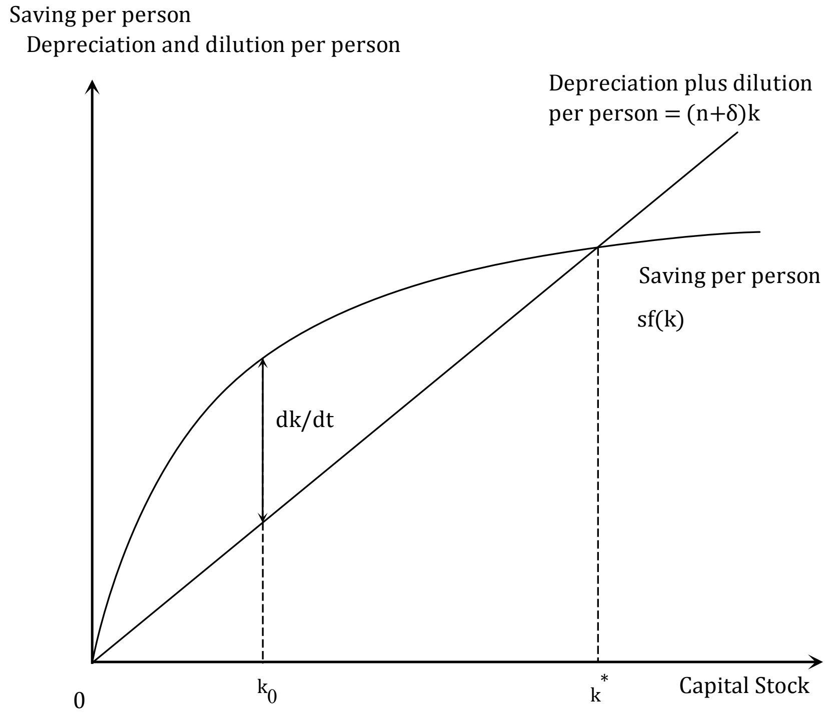

Figure 2.1: The relationship between capital stock and savings, depreciation and dilution

2.2.3 The AK model

To take account of technological change, Solow (1956) extends the model described above by adding the parameter $A(t)$ to represent technology improving over time:

$$Y(t) = A(t)F(K(t), L(t)). \tag{2.3}$$

In the steady state, the growth in per capita incomes and capital is simply the growth rate of $A(t)$. The economy approaches a steady state whereby depreciation exactly offsets capital accumulation because of diminishing returns. At this point, the only effect on the growth rate is the rate of technological progress, exogenous to the model.

The weakness of neoclassical growth theory is that it deliberately excluded technological change from economic incentives, yet it claimed innovation was the greatest contributor to long run economic growth. Arrow (1962) and Frankel (1962) attempted to solve this problem by assuming technological development was a by-product of investing in new capital goods. Arrow (1962) named the phenomenon "learning by doing" in which knowledge was developed in the process of production. Technological progress was external to the firm and therefore all firms benefitted from the collective "learning", independent of their own production. Here the chapter follows the models described in Aghion and Howitt (1998, 2009). Firstly the

Harrod-Domar model (Harrod, 1939; Domar, 1946) is briefly discussed and secondly, the learning by doing model (Frankel, 1962).

The Harrod-Domar model assumes aggregate production has fixed technological coefficients A and B:

$$Y \equiv F(K, L) \equiv \min(AK, BL). \tag{2.4}$$

In this function, production is limited by either capital or labour, and the other factor of production will be in surplus. When $AK < BL$ production is $Y = AK$ and the number employed is $\left(\frac{1}{B}\right) Y = \left(\frac{1}{B}\right) AK < L$. The capital stock grows at the same rate as before $\frac{dK}{dt} = sF(K) - \delta K$, but with the addition of the fixed coefficient $\frac{dK}{dt} = sAK - \delta K$. Therefore, the growth rate of capital is $g = \frac{dK}{dt}/K = sA - \delta$ which is also the growth rate of output since output is proportional to capital.

Two cases are considered. Firstly, if output per person is growing, an increase in the savings rate will temporarily increase the growth rate, but cannot be sustained just as in Solow (1956) and Swan (1956). Eventually capital will grow such that labour becomes the binding constraint and there is no continued growth in output per person. Alternatively, if output per person is falling, the increase in savings will reduce the contraction in per capita output permanently. Diminishing returns does not set in.

The Harrod-Domar model cannot account for sustained growth other than technological change. Alternatively, following Arrow (1962), learning by doing is viewed as a by-product of capital accumulation. In Frankel (1962), knowledge is accumulated when firms accumulate capital. His model assumed that each firm j has the production function $y_j = \bar{A} k_j^{\alpha} l_j^{(1-\alpha)}$. In this function $\bar{A}$ is aggregate productivity which is dependent upon the amount of capital accumulated by all N firms $\bar{A} \equiv A_0 \left(\sum_{j=1}^{N} k_j\right)^{\eta}$, where η represents knowledge externalities among firms but k_j and l_j represent the capital and labour respectively employed by firm j. If Y, K and L denote aggregate production, capital and labour respectively and aggregate labour is given by $L = \sum_{j=1}^{N} l_j = 1$, there is an equilibrium such that $\bar{A} \equiv A_0 K^{\eta}$, individual production is $y_j = A_0 K^{\eta} (K/N)^{\alpha}$ and aggregate production is $Y \equiv N A_0 K^{\eta} (K/N)^{\alpha} \equiv AK^{(\eta+\alpha)}$ where $A = A_0 N^{(1-\alpha)}$. This is known as the AK model. Assuming a constant savings rate $\frac{dK}{dt} = sAK^{(\eta+\alpha)} - \delta K$, growth of capital is

$$g = \frac{dK}{dt}/K = sAK^{(\eta+\alpha-1)} - \delta. \tag{2.5}$$

There are three possibilities for growth. If

$$\eta + \alpha < 1, \tag{2.5a}$$

long-run growth is zero, because the knowledge spillover externality is not enough to counteract the $1 - \alpha$ decreasing returns effect on capital accumulation to each individual.

As in the Solow-Swan and Harrod-Domar models, there is a steady state level where the growth rate of capital is zero. If

$$\eta + \alpha > 1, \tag{2.5b}$$

long-run growth would continually increase. If the capital stock is above the steady state described above, it would keep on growing at an increasing rate, because growth is an increasing function of capital. Lastly if

$$\eta + \alpha = 1, \tag{2.5c}$$

the externalities of learning by doing are exactly equal to the decreasing returns of capital accumulation to the individual. Aggregate production becomes $Y = AK$. Growth is as in the Harrod-Domar model, where $g = \frac{dK}{dt}/K = sA - \delta$. As capital increases, output increases in proportion, because technology also improves and counteracts decreasing returns. However, the difference in this case is that an increase in the savings rate does permanently increase the growth rate because of the additional knowledge externalities of capital accumulation.

2.2.4 Limitations of neoclassical growth

The AK model is fundamental to the early models of endogenous growth. Growth is endogenised through technology being a bi-product of capital accumulation. Lucas (1988) developed an AK model where growth is endogenised in human capital accumulation. Rebelo (1991) considers how different growth rates may be the result of differences in economic policy. The problem with both the neoclassical model and the AK model, even with learning by doing, is that economic signals do not have any influence upon technological change. The economic incentives are for capital accumulation and all firms benefit from the learning by doing externality. Firms in these models have no incentives to invest in research and development or other innovation activities.

Endogenous growth models attempt to solve this problem by making investment in innovation responsive to economic signals. These models are of particular interest, because they leave a significant role for policy due to the positive externalities of research and education. Romer (1990) and (Grossman and Helpman, 1994) propose that technological change is the result of actors, typically responding to market incentives, who invest in research and development to develop new ways of using raw materials. The model relies on monopolistic competition rather than price-taking behaviour, whereby those who invest in research and development can extract additional rents beyond that of a price taker, so that there are economic incentives for developing new products. This group of models are known as “product variety models”, because growth comes from a continually increasing number of available varieties and consumers’ love of variety.

Alternatively Grossman and Helpman (1991b) and Aghion and Howitt (1992) model quality-improving innovations. In these models, profit-seeking actors invest in research and development to develop better-quality versions of existing products. In Grossman and Helpman (1991b) every variety is a quality-leading product with monopolistic competition while in Aghion and Howitt (1992) the quality-leading firm has a monopoly on the intermediate good until a better-quality version produced by another firm replaces it.

The characteristic of technology that is important to endogenous growth is that the result of research and development is knowledge and essentially it has little or no economic cost to knowledge being reused to continue combining raw materials in the same way, or to being used as an input to future innovation activities. In these models, improvements in technology are no longer just an externality, but are motivated by profit-seeking behaviour as the main driver of economic growth.

2.2.5 Product variety models

This section describes a simple version of the product variety model of growth (Romer, 1990; Grossman and Helpman, 1994) in which profit-maximising agents make intentional decisions to invest in research and development and develop new products. The sources of growth are both new intermediate goods and increased specialisation. It is this group of models that has been combined with the NEG as explored in Chapter 3. As a background to these NEG models, this section briefly describes the endogenous growth model from Romer (1990).

As in the neoclassical models, the model has the basic inputs of capital (K), labour (L), and technology (or knowledge, A). Capital is measured in units of different types of durable intermediate goods (x_i). Among the various types of capital goods, it is possible that some pairs of inputs are substitutes, while others are complements or somewhere in between. For simplicity, the definition used in the model here implies output is an additively separable function of all the different types of capital goods such that they are perfect substitutes. Knowledge is considered a non-rival, yet partially excludable good. It is an input to innovation but it is also a measure of the productivity of the optimal combination of currently available intermediates. In this model, there can be an additional variable, human capital, which is the cumulative effect of training and education (including on the job training and learning by doing), beyond the simple employment of an individual and considered separately from traditional capital. H is the rival component of knowledge available only via learning, whereas A represents the non-rival component of knowledge that is part of "technology". Technology is measured by the number of designs of intermediate goods available but should also be considered a broad term to include the processes and knowledge that can be traded.

The economy has three sectors: a research sector, a capital (or durable) good sector and a final good sector. The research sector uses labour and knowledge to produce new designs for durable goods. The durables sector uses the designs and some portion of final output (in foregone consumption) to produce a variety of durable goods that can be used in final production at any time in multiple periods. The final goods sector uses the durables, along with unused labour to produce the final output (Y).

Monopolistic competition of differentiated varieties provides an incentive to research new technologies. Output is either consumed as final goods or saved as new capital. New capital is really foregone consumption where final goods are not produced and the inputs are instead allocated to research (i.e. the cost of new capital is the opportunity cost of producing new designs for durables). For simplicity, population is assumed constant.

Utility is given by the function

$$U(c) = \frac{c^{1-\varepsilon} - 1}{1-\varepsilon}, \qquad \varepsilon > 0, \tag{2.6}$$

discounted by the constant rate of time preference ρ. c is the consumption of final goods. In the steady state, the growth rate and interest rate satisfy the Euler equation

$$g = \frac{r-\rho}{\varepsilon}. \tag{2.7}$$

For a proof of the Euler equation as applied to growth models with consumer optimisation see Barro and i Martin (2003, Ch. 2). Production is a Cobb-Douglas function:

$$Y_t(L_M, x_i) = L_M^{1-\alpha} \sum_{i=1}^{\infty} x_i^{\alpha} \qquad 0 < \alpha < 1, \tag{2.8}$$

where x_i represents the continuous range of possible durable goods such that capital is given by $K = \eta \sum_{i=1}^{\infty} x_i$ and L_M is the labour supply used in manufacturing final goods. However, at time t only a finite number of durable goods A_t have been invented such that for all $i \geq A_t$, $x_i = 0$ and $K = \eta \sum_{i=1}^{A_t} x_i$. Product variety in intermediate goods increases the productivity for final goods. If there are more varieties available (A_t), more of the final good can be produced. The final good is used for consumption, and investment in designs for durables.

A firm that has a design has an exclusive use over that design forever from the period when the design was developed. Each firm i produces a single durable intermediate variety i. Once a firm has come up with a design it can convert units of the final good into variety i on a one to one basis. When a firm produces a capital good it can rent it to the final good sector for a rental rate of p_i. In this model there is no depreciation. The value of each durable good is the value of the discounted future revenues it generates.

To avoid integer constraints, the production function of types of durable goods is instead given by the integral:

$$Y_t\left(L_M, x_i\right) = L_M^{1-\alpha} \int_0^\infty x_i^\alpha di \qquad 0 < \alpha < 1. \tag{2.8a}$$

The level of capital is the level of foregone output (used in production of durables) in all previous periods. The rental price per period of each intermediate durable is its marginal product in the production of final goods.

$$p_i = \alpha L_M^{1-\alpha} x_i^{\alpha-1}. \tag{2.8b}$$

The profit flow to each intermediate monopolist is given by

$$\begin{aligned} \pi_i &= \max\left(p_i x_i - x_i\right) \\ &= \max\left(\alpha L_M^{1-\alpha} x_i^\alpha - x_i\right), \end{aligned} \tag{2.8c}$$

which results in a profit-maximising output of

$$x = L_M \alpha^{\frac{2}{1-\alpha}} \tag{2.8d}$$

and profit of

$$\pi_i = \frac{1-\alpha}{\alpha} L_M \alpha^{\frac{2}{1-\alpha}}. \tag{2.8e}$$

New varieties of durables arrive through investments in research using some portion of labour and knowledge of existing varieties.

$$\frac{dA_t}{dt} = \lambda A_t L_R, \tag{2.9}$$

where λ represents the researcher's productivity and labour employed in research is given by L_R. The rate of new designs can be increased by allocating more workers to research, but this also has the effect of reducing output of final goods as these workers would be otherwise occupied with production. The equation embodies the idea of spillovers in research such that accumulated knowledge A_t is used to produce new ideas.

The profit is the value of each new variety where all value is paid to researchers in wages:

$$\frac{\pi_i}{r} \lambda A_t L_R - w_t L_R = 0. \tag{2.9a}$$

Solving for r,

$$r = \frac{\lambda A_t \pi_i}{w_t}, \tag{2.9b}$$

shows the rate of interest must equal the flow of profit the entrepreneur receives from investing labour in research. Since labour is used in either manufacturing or research, the research wage must equal the marginal product of labour in the manufacturing sector. The production function (Equation 2.8a) implies that total output is given by:

$$Y_t = L_M^{1-\alpha} A_t x^\alpha \tag{2.10}$$

such that wages are given by:

$$\begin{aligned} w_t = \tfrac{dY_t}{dL_M} &= (1-\alpha)\, L_M^{-\alpha} A_t x^\alpha \\ &= (1-\alpha)\, A_t \alpha^{\frac{2\alpha}{1-\alpha}}. \end{aligned} \tag{2.10a}$$

Returning to the research arbitrage equation (Equation 2.9a), substituting profits, wages, the interest rate equality and solving for r gives

$$r = L_M \alpha \lambda. \tag{2.11}$$

Growth in technology is given by:

$$\begin{gathered} g = \tfrac{1}{A_t}\tfrac{dA_t}{dt} = \lambda L_R = \lambda\,(L - L_M) \\ \lambda L_M = \lambda L - g \end{gathered} \tag{2.12}$$

such that

$$r = (\lambda L - g)\,\alpha. \tag{2.12a}$$

Substituting into the Euler equation (given by equation 2.7) and solving for g gives

$$g = \frac{L\alpha\lambda - \rho}{\varepsilon + \alpha}. \tag{2.12b}$$

Growth is a function of the productivity of research activity, the size of the economy and the rate of time preference. Since knowledge can be reapplied to the production of future varieties, the full benefits of new knowledge are shared with future innovators. As a result, the equilibrium growth rate is always less than the socially optimum growth rate. The socially optimal growth rate would be the equilibrium level of investment in innovation if investors were to receive all the benefits of the knowledge they create. Since innovators do not receive the full benefit of their research efforts, they under-invest in innovation. By definition, the equilibrium growth rate is always less than the socially optimal rate.

While the dynamics of the product variety model are not the purpose of this chapter, the above simple model gives a clear outline of modelling techniques in product variety models of endogenous growth, including the role of spillovers and the market failure of innovation.

In particular, existing knowledge is an important input into knowledge production, there are increasing returns to research and markets for durables have monopolistic competition. These provide the incentives for investment in research and drive economic growth. Monopolistic competition is an important characteristic, which can also be found in Aghion and Howitt's (1992) endogenous growth model with creative destruction and Krugman's (1991b) core-periphery model. In each case, there are fundamental outcomes for economic activity and the similarity of modelling techniques suggests the two can be combined in a spatial endogenous growth model such as in Baldwin et al. (2001).

2.2.6 Creative destruction and quality ladders models

Models with creative destruction (Aghion and Howitt, 1992) or quality ladders (Grossman and Helpman, 1991b) introduce obsolescence to the endogenous growth model. The idea recognises that new designs replace old designs. The replacement of existing technology is described by Aghion and Howitt (1992) as Schumpeter's (1934) concept of "creative destruction". In this sense, these types of models are often described as Schumpeterian growth models. In the creative destruction model, a new durable replaces the existing monopoly (or monopolist if there is more than one intermediate variety) durable good in the production process. That is, an owner of an existing design can no longer rent out the old durable, because producers prefer to use the newer design. Rather than there being a continuum of machine types as in Romer's (1990) model, there are quality improvements in a single sector. The quality ladders model retains the continuum of machine types but each variety is continually replaced by higher quality versions in much the same way as in the creative destruction model.

In the simple model adapted from Aghion and Howitt (1992) and Aghion and Howitt (2005), time is defined by discrete periods (t). Innovations arrive after random intervals where the expected number of periods between innovations is dependent on the inputs to innovation. The population of workers/consumers ($L = 1$) supply one unit of labour inelastically and maximise expected consumption. Labour can be employed in manufacturing final goods (L_M) or research (L_R) such that $L_M + L_R = 1$. Prices and quantities are constant within the interval, but may vary between intervals with new technology. The payoff for an innovation is the monopoly profits in the next period. As in the product variety model above labour can be used for research or manufacturing. There are three types of tradables: labour, consumption goods and intermediate goods (comparable to Romer's durable goods). Infinitely-lived individuals have a constant rate of time preference $r > 0$ to discount future consumption, where r is also the rate of interest.

There is only one consumption good, the final good. Similar to Romer (1990), final goods production is a function of productivity and inputs

$$Y_t = (L_M A_t)^{1-\alpha} x^{\alpha} \qquad 0 < \alpha < 1, \tag{2.13}$$

where x is the flow of the current intermediate input good, L_M is labour used in manufacturing and A_t is the current level of productivity of this intermediate good. The wage is the numéraire and parameters are chosen such that consequently the price of the final good is equal to the wage. The intermediate good is manufactured using final goods on a one to one basis such that $x_t = sY_t$, where s is foregone consumption. The inverse demand curve for intermediate goods is the marginal product:

$$P = (L_M A_t)^{1-\alpha} \alpha x^{\alpha-1}, \tag{2.13a}$$

where P is the price of the intermediate good relative to the price of the final consumption good Y_t. The current holder of the patent on intermediate goods is a monopolist that maximises profits:

$$\pi = \left((L_M A_t)^{1-\alpha} \alpha x^{\alpha-1} - w \right) x. \tag{2.13b}$$

The monopolist takes technology as given. Differentiating profit

$$\frac{\partial \pi}{\partial x} = (L_M A_t)^{1-\alpha} \alpha x^{\alpha-1} + (\alpha - 1)(L_M A_t)^{1-\alpha} \alpha x^{\alpha-1} - w = 0 \tag{2.14}$$

and solving for output gives

$$x = L_M A_t \alpha^{\frac{2}{1-\alpha}}. \tag{2.14a}$$

The price of the current version of intermediate goods is equal to

$$P = \frac{1}{\alpha}. \tag{2.14b}$$

Monopoly profits will be equal to

$$\pi = \left(\frac{1-\alpha}{\alpha} \right) x. \tag{2.14c}$$

Research produces a random flow of quality-improving innovations that arrive with a probability given by $\lambda \phi (L_R)$ where L_R is the amount of labour employed in research by a future intermediates producer and the technology of research determines the constant λ and the functional form of ϕ. In this simple model, there is only one research firm that shares the payoff from innovation between its research workers. L_R can also be thought of as foregone output as it is the portion of labour unable to be used to manufacture final goods. This produces a distribution of the interval length between innovations. If the input to research is greater, the length of time between innovations is expected to be shorter. Labour is an essential factor, required for development of new innovations. If the labour allocated to research is zero, there is no growth.

All improvements in producing intermediate and final goods are reflected in the final goods sector parameter A_t. New intermediate goods increase productivity in production by a factor of γ such that $A_t = A_0 \gamma^t$ where A_0 is the historical level of technology. There is an assumption of immediate diffusion of technology in that the new intermediate good is used by all final goods producers. Aghion and Howitt (1992) note that diffusion could be introduced so that the productivity parameter gradually increases towards A_t over the period between new inventions. The inventor of the intermediate good has an infinitely-lived patent on the design and gains a monopoly on intermediate goods production until the next innovation. Final goods markets are still perfectly competitive, but pay the monopoly price for intermediate inputs. At any time, the monopolist takes the current amount of research as given and therefore also takes the length of each period as given.

The objective of the entrant firm is to choose the amount of labour to invest in research L_R independently of other firms, to maximise the flow of expected profits

$$\pi_{i,t} = \lambda L_R V_{t+1} - L_{R,t} \tag{2.15}$$

where V_{t+1} is the value of the $t+1^{st}$ innovation. The value of the innovation is the expected present value of the monopoly profits (π_{t+1}) of the new innovation over the period of length exponentially distributed with parameter $\lambda\phi\left(L_{R,t+1}\right)$:

$$V_{t+1} = \frac{\pi_{t+1}}{r + \lambda L_{R,t+1}}. \tag{2.15a}$$

All research is conducted by the potential entrant rather than the incumbent monopolist. An incumbent firm does not want to make its existing technology obsolete by developing the next innovation as they would choose to extend this period by as long as possible. That is, the value of research to the monopolist is $V_{t+1} - V_t$, which is strictly less than the value of research to an outside firm. The model here considers only drastic innovations. In the alternate case with incremental innovations, perhaps such as the innovations in the model developed in Chapter 6, the incumbent receives a higher payoff than an outside firm because of the rent-dissipation effect (Tirole, 1988). That is, a firm can extend its monopolist period of production by incrementally innovating but a potential entrant's incremental innovation has to compete with the incumbent. In this way, incremental innovations are typically developed by incumbents while more drastic innovations are developed by new entrants. Similarly Reinganum (1983) reconciles stochastic and deterministic approaches to growth by suggesting the deterministic model is most appropriate for incremental innovations where the incumbent can minimise risk by aiming for incremental innovations with a high probability.

An innovation raises productivity forever. New innovations raise A_t by the multiple $\gamma > 1$ with the probability $\lambda\phi\left(L_R\right)$. Each additional innovation raises the now higher A_t. Each innovator captures the monopoly rents for all previous innovations, but only for one period. Therefore, an innovating firm is only able to capture a portion of the total payoff created by its productivity gain while the rest of the payoff is captured by future innovators who build upon the existing innovation. This knowledge spillover is similar to the intertemporal knowledge spillovers in Romer (1990) where all researchers benefit from existing knowledge when developing new innovations. The knowledge is non-rival and only partially excludable.

Maximising expected return to the entrepreneur, the equation

$$\frac{\partial \pi_i}{\partial L_R} = \lambda V_{t+1} - 1 = 0 \tag{2.16}$$

can be rearranged to

$$\lambda V_{t+1} = 1. \tag{2.16a}$$

The right hand side of the equation is the marginal benefit of labour (the wage is equal to the numéraire) and the left hand side is the marginal benefit of research. In equilibrium, these equalise such that

$$\frac{1-\alpha-\alpha\lambda r}{1-\alpha+\alpha\lambda} = L_{R,t+1}. \tag{2.16b}$$

The important outcome of the model is not the level of research, but the growth in real output of consumption goods. Output in period t is given by $Y_t = A_t^{1-\alpha}x^{\alpha}$ and output in the subsequent period is $Y_{t+1} = \gamma A_t^{1-\alpha}x^{\alpha}$. Growth in Y_t per period is given by

$$g_{Y_t} = \frac{Y_{t+1}-Y_t}{Y_t} = \frac{\left(\gamma-1\right)A_t^{1-\alpha}x^{\alpha}}{A_t^{1-\alpha}x^{\alpha}} = \gamma - 1. \tag{2.17}$$

Since the length of each period is random with probability of ending in period t equal to $L_{R,t+1}$, the expected growth rate in every period is

$$g_Y = \left(\gamma-1\right)\frac{1-\alpha-\alpha\lambda r}{1-\alpha+\alpha\lambda}. \tag{2.17a}$$

The parameters that increase growth are increases in arrival parameter λ, the size of innovations γ, and for the Cobb-Douglas example, the level of market power α, while increases in the interest rate decrease the growth rate. Furthermore, each of these parameters have the same effect on variance. For a full proof and discussion of similar models see Aghion and Howitt (1992, Ch. 4) and Aghion and Howitt (1998, Ch. 2). It is particularly important to note the necessity of market power to growth as it provides the incentive for investment in innovation. The Cobb-Douglas example relies on a minimal degree of market power, while every endogenous growth model relies on some level of imperfect competition to encourage research and development investment in the first place.

There are a number of extensions to the model offered by Aghion and Howitt (1992, 2005, 2009). These extensions are not addressed in formal models here as it is beyond the scope of this chapter but a few are listed and their impacts noted. These extensions answer some interesting questions for endogenous growth theory.

Firstly, many models disaggregate labour such that skilled labour is an integral input to innovation. This extension offers an additional policy lever to increase the growth rate by increasing the amount of skilled labour. Secondly, models are adapted for non-drastic innovations. Under non-drastic innovations, the new innovator is constrained by competition from the previous monopolist. Therefore, the new intermediate manufacturer does not inherit a monopoly income stream, because it is forced to price at a level which takes the market from the incumbent, a level below the otherwise monopoly price. The analysis also assumes that the former monopolist cannot contract with the new innovator not to compete, for example by selling the patent to the innovator, allowing them a full monopoly even with non-drastic innovations. If this were allowed, the outcome depends upon the bargaining power of the former incumbent.

Thirdly, the strategic monopsony effect is where the intermediate monopolist adjusts its production to take account of its influence on current research. By producing more of the intermediate good, they can increase the wage rate for skilled workers, increasing the cost of research and reducing the amount of research conducted by its competitors. This in turn lengthens the time between innovations. If an intermediate monopolist is acting strategically in the labour market, they must balance the additional profits from a longer stream of monopoly profits with the additional cost of its own skilled labour.

The model also works when there are multiple intermediate sectors. This version of the model describes each innovator as gaining a niche monopoly, until such time as another innovation, within that niche sector, replaces it. This occurs in all sectors with some interaction between sectors on wages. The outcomes of the model are the same, except there is more stable growth. With many varieties, the outcome is similar to a deterministic model where the length of time between innovations is fixed, but the level of quality improvement is endogenised, determined by investment in innovation.

Alternatively models could have a stochastic size of innovation per period where only the largest innovator takes the niche monopoly in each period. In all of these models, larger innovations (or more frequent innovations in the stochastic model) require a higher research expenditure. The typical result of all of these models is that there is some market failure of innovation such that the level of research is below the socially optimal level. This is because the inter-temporal spillover of knowledge prevents an innovating firm from receiving the full benefits of iss research investment.

2.2.7 No-scale-effects models

Both the product variety and quality ladders models predict scale effects. That is, as the number of researchers increases so too does the growth rate. On the other hand, empirical literature has found that large economies do not grow faster than small ones (Jones, 1995b; Dinopoulos and Thompson, 1999). In response several models are developed that avoid the scale effect (Jones, 1995a, 1999; Segerstrom, 1998; Peretto, 1998; Young, 1998). These models use an elegant approach of incorporating two dimensions to research and development: both product variety and quality ladders. The insight of Young (1998) is that a larger population results in quality improving research and development being spread more thinly across more varieties.

Jones (1999) suggests that there is a critical implicit parameter that is assumed to be one such that the scale increase in population exactly matches the additional number of firms thereby dissipating the scale effect. If this factor is greater than one, there are scale effects while if it is less than one, there is an inverse scale effect. It may well be that this factor is greater or less than one for individual industries at different times in their product cycle and therefore the average would tend towards one. Therefore in aggregate growth models, it is sensible and elegant to pursue this approach. The models developed in Chapters 4, 5 and 6 in this thesis also avoid scale effects such that any scale implication is a result of other factors related to the spillover of knowledge between co-locating firms.

2.3 New Economic Geography

Economic geography deals with how economic activity is arranged over space. New Economic Geography brings location and space to the mathematical economic model allowing an opportunity to extend other areas of economics (including economic growth) by examining the role of space. In these models, firms and people make interdependent spatial decisions that change the nature and performance of traditional models. While most micro and macroeconomic models assume everything occurs, figuratively speaking, "on a pinhead" (Isard, 1956), Krugman (1979) brings spatial factors to the topic of international trade and the location of manufacturing (Krugman, 1991b). While spatial factors for localisation have been explored in the economics literature as long ago as Marshall (1890), but also Hoover (1948) and Lichtenberg (1960), economics typically ignores the role of space.

Marshall (1890) identifies three reasons why firms choose to locate together. These reasons are labour market pooling, benefits in subsidiary trade and technological spillovers.

While Marshall's observation would suggest the relevant factor for innovation is technological spillovers, it is possible that all three causes of localisation involve the spatial transfer of knowledge. Firstly, through labour market pooling, firms benefit in a number of ways. The reason presented in Krugman (1991b) is a stylised story which shows how labour availability allows firms in the same locality to benefit more during good times than firms in different locations. Similarly, employees are less worse off during a firm's bad times, because they have alternative firms available for employment, which may be experiencing good times. These are all factors that relate to trade and location. In terms of subsidiary trade, the shared use of suppliers has benefits for trade from increased specialisation, economies of scale, interaction with many suppliers, and competitors' suppliers. In addition, these factors are also important not just for reducing transaction and transport costs, but they facilitate additional knowledge spillovers between similar firms. Finally, Marshall identifies technology spillovers. In terms of innovation, proximity facilitates the transfer of knowledge embedded in technology, from which nearby firms can innovate in future periods and contribute to economic growth. There are also knowledge spillover and innovation benefits to labour market pooling. Firms can benefit from workers' experience at other firms. This sharing of labour between firms creates a knowledge spillover as workers bring their knowledge and human capital to other firms, from which firms and workers can innovate and contribute to economic growth. With localisation, workers facilitate these knowledge spillovers.

Economic geography identifies the knowledge embedded in traded technology as benefits from localisation, but the knowledge spillovers created by exchange of employees and interaction between firms are also benefits from localisation, and contribute to economic growth. The NEG combines formal economic models with the observations of economic geography. By adding transport costs using the iceberg specification (Samuelson, 1952), Krugman's (1991b) core-periphery model has some unique insights to firm and labour location which were not previously available in formal economic models. Under an iceberg model, some of the good is consumed during transportation, that is, only a portion of what is shipped actually arrives at the destination and another portion "melts" away. As a result of increasing returns, workers and firms have an incentive to choose a location that minimises transport costs. This section explores the core-periphery model and the benefits of using a transport costs approach to recognising the role of distance. Chapter 3 discusses how this type of trade and transaction cost approach has also been applied to the role of distance in knowledge spillovers as an input to economic growth.

Many of the geography factors of knowledge spillovers and innovations have been explored in some depth. Urban and regional economics has existed for some time, though on the fringe of traditional economics. Krugman (1991b) argues for economic geography to be accepted as a major field within economics, encompassing international trade. To some extent, his Nobel

Prize in 2008 has brought the study of economic geography to mainstream economics through the topic of international trade. In a similar way the economics of innovation has benefited from considering the role of space. Asheim and Gertler (2005) note the increasing spatial concentration of knowledge-intensive industries despite the increasing use of ICT. They argue that geography is fundamental to studying innovation itself. Varga (2009) explores regional economic growth as knowledge and innovations are transferred from universities to the regional economy. Cappellin and Wink (2009) look at knowledge and innovation transfers in medium-technology sectors in Europe. Acs (2002) explores a range of literature on innovation and growth in cities. Hall (2005) discusses the determinants of knowledge diffusion and their effects on the rate that innovations spread out across society and space. These types of studies are not a new phenomenon either. Rosenberg (1972) also explores the diffusion of innovations and remarks at the overall slow rate society accepts innovations.

It is also not a new trend to investigate the spatial or geographic variation in economic growth. For example, Nijkamp and Poot (1998) explore endogenous technological change and interdependent regions in a model that demonstrates how space could be examined with endogenous growth. Baldwin et al. (2001) consider a NEG model similar to Krugman's core-periphery model, incorporating endogenous growth. This section describes some models of the NEG and then explores how economic geography and innovation are linked. Lastly, this section considers how the literature on innovation that has benefited from considering spatial factors can now be incorporated into the endogenous growth literature to better understand the nature of economic growth.

2.3.1 The core-periphery model

The core-periphery model (Krugman, 1991b) aims to explain the location of manufacturing using techniques from industrial organisation and an iceberg specification of transport costs. In particular, it aims to explain why economic activity becomes concentrated in some regions while leaving other regions relatively undeveloped, even if the regions have similar underlying characteristics.

The model is a variation on the Dixit and Stiglitz (1977) model of monopolistic competition with the addition of increasing returns, trade costs and choice of location for firms and individuals. Manufacturers locate near their markets in order to minimise trade costs and consumers locate near manufacturers to minimise their costs of living. This localisation story has been told many times to explain concentration of economic activity. However, Krugman's (1991b) model is able to more rigorously formalise the observation into an economic model. The method in developing the core-periphery model is to assume increasing returns to scale for manufacturers and to include trade costs for both workers and firms. This creates a circular causation mechanism where manufacturers locate where there is a concentration of demand (market access effect) while workers locate near manufacturers, because it reduces the cost of purchasing manufactured goods (cost of living effect). These are both centripetal effects attracting economic activity to concentrate and the process continues to feed on itself, further concentrating economic activity.

At the same time, greater competition (market crowding effect) within the concentrated region discourages concentration of economic activity. This centrifugal effect causes firms to locate in the smaller region where they benefit from lower competition and higher prices. When trade costs (both transaction and transport costs) are high, the market crowding effect is greater than the market access and cost of living effects. This results in dispersed economic activity as firms locate separately from each other and serve a captive local market. However, as transport costs fall, the strength of each force changes. At some point, manufacturers can serve a wider market (whether there are competitors or not) from a single location using economies of scale. This is a centripetal force, known as the market access effect that encourages firms to locate in the largest market in order to minimise transport costs. Similarly, the cost of living effect encourages workers to locate in the region with the most suppliers as this minimises the transport cost component of consumption. The market crowding effect declines with lower transport costs as manufacturers are better able to compete with each other despite transport costs from operating in different locations to their customers. While the cost of living effect has also fallen, the market access effect has increased by enough that these centripetal effects dominate the centrifugal market crowding effect. The overall result is workers and firms become concentrated at a single point if transport costs decline enough.

For simplicity, the model described here is based on the Krugman (1991b) core-periphery model and a similar version described in Krugman (1991a, Ap. A). The simplest form of the model has two regions: North and South. Each region is assumed to have the same initial endowment of attributes. All goods are produced using only labour (and land for agricultural goods, though in this model land is really about the immobility of this factor of production). As in Dixit and Stiglitz (1977), two types of goods are produced: agricultural goods, C_A, which are produced in a perfectly competitive market and which depend upon the land for their production, while manufactured goods, C_M, are produced in a monopolistically competitive market and require a significantly smaller amount of land (if any).

However, there are a few key differences between Krugman (1991b) and Dixit and Stiglitz (1977). Firstly, in Krugman (1991b), manufacturing is less reliant on land, so firms can locate in either region whereas agriculture must locate where there is available land so it's factor of production is distributed equally between both regions. Secondly, in Krugman (1991b), there are two specific types of workers: agricultural workers L_A and industrial workers L_M. Agricultural workers are spread equally between regions and are immobile.

Immobility is a proxy for a land use requirement which means land can more be ignored as a factor of production in the model. Industrial workers on the other hand are completely mobile and can locate in either region. Thirdly, in Krugman (1991b), transport and trade costs are included using Samuelson's (1952) iceberg framework. Finally, manufacturing firms experience increasing returns which encourages a tradeoff in manufacturing between having a single location (to minimise manufacturing costs) and multiple locations (to minimise transport costs). Furthermore, when firms limit their production to one location, they choose the location where there is the largest market to also minimise transport costs.

The total population is set to one, $L = L_M + L_A = 1$. All individuals, share the utility function

$$U = C_M^{\mu} C_A^{1-\mu}, \tag{2.18}$$

where C_M is the quantity of manufactured goods consumed and C_A is the quantity of agricultural goods. Manufactured goods are the aggregate of a variety of differentiated manufactured goods

$$C_M = \left(\sum_{i=1}^{n} c_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{2.18a}$$

where n is the total number of products, c_i is the quantity of each manufactured good i and $\sigma > 1$ is the constant elasticity of substitution (CES) between different varieties. While the Dixit-Stiglitz approach to production (or consumer preferences) may be unrealistic by assuming that all firms have the same production function (or all consumers have the same preferences), the approach is particularly useful as an aggregation tool and performs well in models where consumers and workers are mobile. Although other models could also incorporate a love-of-variety, the Dixit-Stiglitz structure is chosen as the preferred modelling approach in this thesis due to its tractability and the model's prominence in both endogenous growth theory and the NEG.

The labour requirement per unit of each good is assumed equal to 1, the supply of agricultural workers is set to $(1-\mu)/2$ in each region, while the total portion of manufacturing workers is set to $\mu < 1$, where these workers can locate in either region. Therefore, world manufacturing labour is given by $L_M = L_N + L_S = \mu$, where L_N refers to the number of industrial workers in the northern region and L_S refers to the number of industrial workers in the southern region. Production of manufactured goods includes a fixed cost and a variable cost, creating economies of scale, such that the labour requirement for each manufacturer is $L_{M,i} = \alpha + \beta c_i$.

As there are two regions, transport costs are incurred for manufactured goods traded between regions. Agricultural goods are assumed to have transportation costs of zero as a simplification for analytical convenience that ensures there is a constant agricultural price (and wage) between regions, which is used as the numéraire. If there were a positive agricultural transport cost, local agricultural wages and prices would be lower in the smaller region resulting in a reduced cost of living effect for skilled workers locating in the larger region. That is, there would be some lessening of the agglomeration effects because agricultural goods would be cheaper in the periphery (although manufactured goods still relatively more expensive).

To avoid this additional complexity, Krugman (1991b) assumes zero transport costs for agricultural goods, which is also the approach taken by the models in this thesis for a similar reason. Since agricultural goods are considered homogeneous, regions will either be exporters or importers of agricultural goods, but never both. Manufactured goods have iceberg transport costs such that a portion of the good "melts" away while the good is being transported. The fraction that arrives $\tau < 1$ is an inverse index of transportation (and transaction) costs. If transport costs are low, then τ is close to 1 (it is exactly 1 if costs are zero) but if transport costs are high, then τ is close to zero.

A large number of manufacturing firms each produce their own product. As in the Dixit-Stiglitz model, the constant elasticity of demand facing each firm is σ. Therefore, the profit-maximising price in the region for each firm is set according to $p_i = \left(\frac{\sigma}{\sigma-1}\right) w_j$, where w_j represents the wage in that firm's region. As there is free entry into manufacturing, profits are driven to zero. Therefore, revenue less production costs is equal to the fixed costs incurred for market entry, $(p_i - \beta w_j)\, c_i = \alpha w_j$, which implies that output by each firm is given by $c_1 = c_2 = \frac{\alpha(\sigma-1)}{\beta}$. Output per firm is the same in each region. However, smaller regions will have fewer manufactured products. Further, the number of manufactured goods is proportional to the amount of labour $\frac{n_1}{n_2} \equiv \frac{L_1}{L_2}$. Equilibrium prices, wages and quantities are found initially taking the distribution of workers between regions as given. Workers then have the incentive to move between regions if they can achieve a higher real wage level, where real wages are a measure of the utility of workers. A steady state is where the distribution of workers is constant, such that no worker can be made better off by migrating to a different region.

First, consider demand in each region for products from the two regions. Let c_{NN} be the consumption in the northern region of a representative manufactured good from the northern region, and c_{NS} be the consumption in the northern region of a representative manufactured good from the southern region. The price of a good produced in the same region is simply the free-on-board price $p_i = \left(\frac{\sigma}{\sigma-1}\right) w_N$. Conversely the price of a product from the other region is the price including a transport component, $\frac{p_S}{\tau} = \left(\frac{\sigma}{\sigma-1}\right) \frac{w_S}{\tau}$, where τ is as specified above. Therefore, relative demand for these two goods in the northern region is

$$\frac{c_{NN}}{c_{NS}} \equiv \left(\frac{p_N \tau}{p_S}\right)^{-\sigma} \equiv \left(\frac{w_N \tau}{w_S}\right)^{-\sigma}. \tag{2.19}$$

The ratio of expenditure on local products to products from the other region is defined as

$$z_{NN} \equiv \left(\frac{n_N}{n_S}\right)\left(\frac{p_N\tau}{p_S}\right)\left(\frac{c_{NN}}{c_{NS}}\right) \equiv \left(\frac{L_N}{L_S}\right)\left(\frac{w_N\tau}{w_S}\right)^{-(\sigma-1)}, \qquad (2.19a)$$

There is a similar ratio for southern region expenditure on northern region products to southern region products

$$z_{NS} \equiv \left(\frac{n_N}{n_S}\right)\left(\frac{p_N\tau}{p_S}\right)\left(\frac{c_{NN}}{c_{NS}}\right) \equiv \left(\frac{L_N}{L_S}\right)\left(\frac{w_N}{w_S\tau}\right)^{-(\sigma-1)}. \qquad (2.19b)$$

The total income for workers is the total spending on goods in both regions, i.e. workers spend all their income. The total income of northern region workers is

$$w_N L_N = \mu\left[\left(\frac{z_{NN}}{1+z_{NN}}\right)Y_N + \left(\frac{z_{NS}}{1+z_{NS}}\right)Y_S\right]. \qquad (2.19c)$$

Similarly the income for workers from the southern region is

$$w_S L_S = \mu\left[\left(\frac{1}{1+z_{NN}}\right)Y_N + \left(\frac{1}{1+z_{NS}}\right)Y_S\right], \qquad (2.19d)$$

where Y_N and Y_S are total regional incomes (including the wages of agricultural workers) in the northern and southern regions respectively. The regional incomes depend upon the distribution of workers and wages. Noting the agricultural wage is numéraire, regional incomes are

$$\begin{aligned} Y_N &= \tfrac{1-\mu}{2} + w_N L_N \qquad \textit{and} \\ Y_S &= \tfrac{1-\mu}{2} + w_S L_S. \end{aligned} \qquad (2.20)$$

These equations can then be used to determine wages, income and expenditure on individual goods. As a result, there are a number of possible steady states.

Firstly, if $L_N = L_S$, then $w_N = w_S$, and there is also equal expenditure, an equal number of goods produced in each region and an equal amount of trade between regions. However, if labour shifts to the northern region, relative wages $\frac{w_N}{w_S}$ can move either up or down depending upon the parameters of the model. There is a “home market” effect where wages tend to be higher in the larger market, as firms are also attracted to the larger market to minimise transport costs.

Workers are interested in real wages such that they want to move to the region which offers them the highest level of utility. Workers in the larger region have a higher portion of manufactured goods produced within their own region and face a lower price for manufactured goods, because they spend a higher portion of their income on goods produced within their own region as these do not face transport costs. Defining the portion of the manufacturing labour force that locates in the northern region as $f = \frac{L_N}{\mu}$, the true price index of manufactured goods for consumers in the northern region is

$$P_N = \left[f w_N^{-(\sigma-1)} + (1-f) \left(\frac{w_S}{\tau} \right)^{-(\sigma-1)} \right]^{-\frac{1}{\sigma-1}}, \tag{2.20a}$$

and for consumers in the southern region

$$P_S = \left[f \left(\frac{w_N}{\tau} \right)^{-(\sigma-1)} + (1-f)\, w_S^{-(\sigma-1)} \right]^{-\frac{1}{\sigma-1}}. \tag{2.20b}$$

Therefore, real wages in each region are given by

$$\omega_N = w_N P_N^{-\mu} \qquad and \qquad \omega_S = w_S P_S^{-\mu}. \tag{2.20c}$$

The manufacturing price index is taken to a power of $-\mu$ consumers also have a taste for agricultural goods. An aggricultural price index is not required because the price of aggricultrual goods is the numéraire. From the price index equations, if there is a shift of workers to the northern region, this lowers the northern price index and therefore increases the real wages of northern workers, relative to workers in the southern region.

When the distribution of workers is equal between regions, then $\omega_N = \omega_S$. This is a stable equilibrium or steady state if ω_N/ω_S decreases with f, because workers are attracted to the region with fewer workers. If ω_N/ω_S increases with f, then workers are attracted to the region with more workers and there is divergence in real wages and regional incomes. Under certain conditions, the stable equilibrium is an equal distribution of workers. However, under some conditions, the larger region will dominate, and all manufacturing will concentrate within this region. In the real world, the selection of the region that dominates can be because of an initially higher selection of resources attracting more workers or access to natural transport routes. The choice of region that agglomerates can simply be a historical "accident".

A complete characterisation of the model's behaviour can be found in Krugman (1991b) with a simple description also available in Krugman (1991a). Instead, to characterise these types of models more intuitively Baldwin et al. (2003) use a tomahawk diagram to describe the possible steady states for a range of transport costs. Figure 2.2 shows a tomahawk diagram to describe the northern region's share of skilled workers (or capital or otherwise mobile economic activity in alternate versions of the core-periphery model) in relation to transport costs for all possible steady states. The various stable equilibria depend upon the the elasticity of substitution between goods σ, the portion of skilled labour (and of agricultural labour) and the cost of transport. The solid lines describe stable equilibria of the northern region's share of skilled workers.

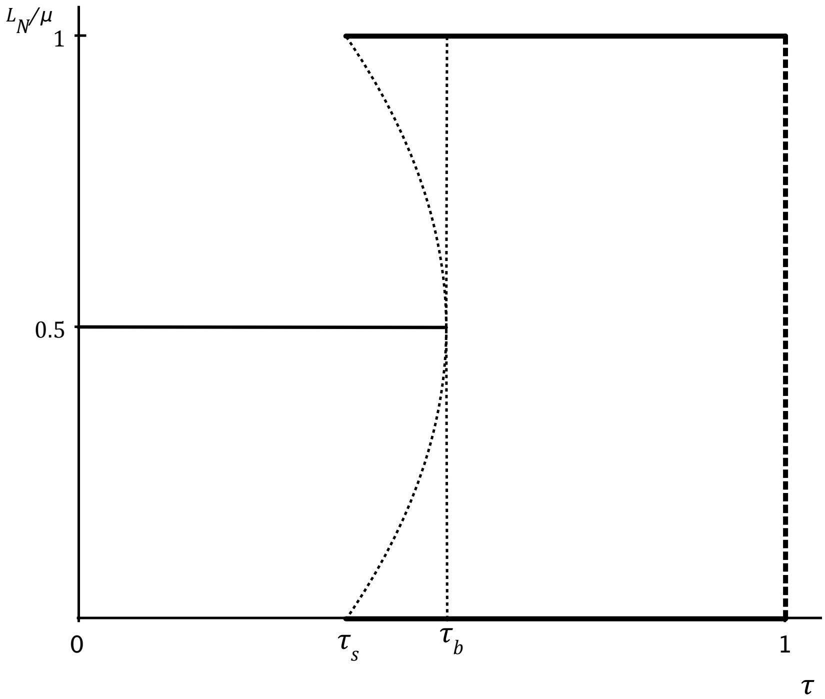

Figure 2.2: Tomahawk diagram of the core-periphery model
Source: Reproduced from Baldwin et al. (2003)

As transport costs fall (τ increases), at some critical value of τ called the *break point* (indicated by τ_b), manufacturing firms will find it profitable to concentrate in a single region. Eventually, when τ reaches 1, the distribution of labour and manufacturing is arbitrary (indicated by the dashed line), since transport costs have reached zero. When transport costs are high and only a small share of manufacturing is mobile, the distribution of skilled workers and manufacturing firms depends on the distribution of agricultural workers (in this case 50:50). When there are lower costs, a higher share of manufactured goods or stronger economies of scale, manufacturing concentrates in a single region (either by a historical "accident" that creates a head start or an advantage in resources).

Similarly starting with economic activity concentrated in a single region, because transport costs are sufficiently low. As transport costs increase (τ declines), at some critical level of τ called the *sustain point* (indicated by τ_s), transport costs are sufficiently high that the market crowding effect dominates the cost of living and home market effects such that there is a dispersion of economic activity between regions.

The assumption of zero transport costs in the agricultural good is false, but assists analytical tractability. Since agricultural transport costs would clearly be a force for dispersion it would appear to counteract the home market effect. Davis (1998) even finds that the home market effect disappears entirely in a model with immobile labour. The importance of agricultural transport costs was explored in Fujita et al. (1999, ch. 7). Put simply, there is an additional dispersion force that varies with transport costs just like the home market, cost of living and market crowding effects discussed above. As a result of agricultural transport costs, the steady state distribution of economic activity is now a balance of four location forces. Agricultural transport costs reduce the range of parameters for which agglomeration occurs but under reasonable conditions, similar effects occur to those discussed in the model explored above. In reality there are many more drivers of location choices that determine the steady state distribution of economic activity (such as congestion costs or land rents). Spatial economic models can only sensibly model a few of these in order to discuss the trade-offs faced by economic agents. Assuming zero transport costs in agricultural goods in order to maintain tractability appears a sensible approach. The focus of Chapters 4 and 5 of this thesis are the effects of knowledge spillovers upon location choices. In particular, Chapter 5 assumes even zero transport costs in the differentiated good such that there is no home market effect and the model can focus solely on these knowledge-based location trade-offs.

Krugman and Venables (1995) explore another forming of the core-periphery outcome. The new model includes trade in intermediate goods. The interaction between transport costs and intermediate goods creates country-specific external economies that for some countries leads to agglomeration. These externalities are similar to the labour and trade cost interaction in the earlier Krugman (1991b) model. However, they occur without labour mobility. It is therefore the linkages between firms supplying and receiving goods from one another that create the agglomeration, rather than a worker/customer relationship between labour and firms. Furthermore, if labour is immobile, peripheral wages are lower than in the core. As a result, when transport costs fall, there is a re-industrialisation of the periphery region and a convergence of wage rates when manufacturing savings from a lower wage rate in the periphery become greater than the disadvantage of being remote from markets and suppliers. Krugman and Venables believe the implications of this for economic geography are a clear distinction between interregional and international economics, where labour is less mobile between nations.

In contrast to Krugman and Venables (1995), the integral role of skilled workers in innovation and the international mobility of very high skilled workers may mean this final result is not transferable to the study of economic growth. Skilled labour in particular is far more mobile than most industrial labour. While there may be a peripheral re-industrialisation for low to moderately skilled industries, the labour market may be more international for the very high skilled workers in research and development, suggesting there would be significantly less agglomeration in the periphery for these industries. For this reason, the two-region models developed in Chapters 4 and 5 of this thesis retain mobility of skilled labour as these workers play a key role in innovation and are highly mobile. The mobility of skilled workers may be a key part of the explanation for the phenomenon that is the motivation for this thesis, that is, low innovation rates in peripheral countries such as Australia and New Zealand. The mobility of skilled workers may partially explain the agglomeration effects of modern industrial location in a similar way to Krugman's (1991b) model above. With the addition of knowledge as an input to innovation and growth as modelled in this thesis, it also helps explain the disparity of regional and international rates of innovation and the clustering of firms in similar industries.

2.3.2 Economic geography and systems of innovation

Knowledge and innovation both have spatial features. The study of innovation has greatly benefitted from introducing the concept of space. Therefore, endogenous growth models should be extended to take account of these economic geography factors.

It is possible that under certain circumstances, knowledge spillovers could be an even greater contributor to economic growth than knowledge development. While it would not be sustainable for every country to free-ride on each other's knowledge development, new knowledge may be difficult to develop in places where an industry does not have the scale, competitive pressures or right incentives to develop knowledge. This is the opportunity taken in Chapters 4 and 5 of this thesis to examine the spatial nature of knowledge spillovers and growth. Models that incorporate the spatial nature of knowledge and innovation may show how smaller economies rely on knowledge spillovers from larger foreign economies for economic growth, while larger economies have enough economies of scale to be at the cutting edge of knowledge development.

One of the major barriers to knowledge spillovers is distance. This effect occurs because knowledge is only partially codified, as discussed above. By definition, tacit knowledge requires some form of interaction (such as face-to-face or other interactive communication) in order to be fully diffused . In New Zealand for example, this could be by far the most important factor hindering knowledge spillovers. Geographically, New Zealand is both isolated and sparsely populated. In addition, many centres are also isolated due to the nature of New Zealand's terrain and the relative lack of infrastructure in comparison to more densely populated and developed nations. Furthermore, technology and knowledge diffusion appears to be greater within countries than between countries (Eaton and Kortum, 1999; Jaffe et al., 1993), so the sheer small scale of the economy is a further barrier to the transfer of knowledge.

While small but dense economies, such as Singapore or the Netherlands, will benefit from the close proximity of firms, the transfer of knowledge from other economies depends upon the interaction of firms with workers and firms in those other economies as well as international geographic distance. Keller (2002) suggests the decay of technology diffusion over distance is so great that a distance of 1,200 km leads to a decrease by 50% in diffusion. While this diffusion is likely to be specific to the type of technology and knowledge involved, Keller (2004) suggests the effect of distance is such a significant barrier, that economic growth in Australia for example would benefit very little from technology developed in the G-5 countries. Trade may facilitate interaction between workers, but the combination of distance and a relative lack of large near-by trading partners mean technological knowledge takes too long to reach Australia to be successfully incorporated into innovation activities. Distance is a barrier to trade, capital and the flow of people, and therefore is a barrier to the knowledge spillovers that occur within these flows.

A country cannot shift its geographic location, but policies can facilitate ways to artificially reduce the distance to the rest of the world and the spatial transaction costs of interaction. Improving international connectivity and networking, such as skilled immigration, improved international telecommunications networks and international free trade agreements, are some of the mechanisms that can reduce the size of the distance barrier. The former "closeness" of New Zealand with the UK is an example of reducing the geographic isolation of New Zealand's economy. The effect of migration, both into and out of the country will improve ties with foreign countries, building international knowledge networks and assist the introduction of new ideas and working practices. Each mechanism described essentially results in a reduction of the transaction cost of knowledge spillovers and improves the effectiveness and number of knowledge spillovers to an isolated region.

In addition to international knowledge spillovers, models of innovation must also consider local and regional interactions and knowledge diffusion within the country. Knowledge dissemination around an economy is important for growth in other regions. Models of endogenous growth do not have to be limited to national boundaries. These boundaries can be arbitrary and do not necessarily reflect modern integrated economies. Modelling may be more effective when representing regions or cities within a country or continent and examining regional knowledge spillovers and regional economic growth.

Agglomeration refers to the tendency for economic activity to cluster together in close proximity. This has the effect of both economies and diseconomies of agglomeration, with the relative magnitude of each, determining the overall distribution of economic activity. As cities get larger, so does the cost of living and other externalities to a point where the costs of research and development, for example, force some of these activities to move elsewhere (Acs, 2009). In terms of innovation, agglomeration provides benefits through knowledge spillovers or thicker labour markets for specialised skills. Agglomeration effects on innovation includes the knowledge spillovers that occur through co-locating firms that share the same knowledge base (Feldman and Audretsch, 1999) or a similar skilled workforce (Audretsch and Feldman, 1996).

Both tacit and codified knowledge are transferred between firms by interaction and cooperation. Firms that are clustered together in larger cities will benefit from more interaction and cooperation. Furthermore, knowledge transfer is affected by the distribution of human capital between different firms and regions (Simonen and McCann, 2008). The models in Chapters 4 and 5 include parameters which show the effect of localisation and clustering respectively, as a result of these spillovers between firms, industries and regions.

Cities are an important facilitator of face-to-face contact. Face-to-face contact is imperative to bringing together economic actors, despite reductions in transport costs and the rise in the complexity of information. Storper and Venables (2004) and McCann (2007) find that face-to-face contact and interaction are the key method for knowledge transfer. In particular, Storper and Venables (2004) argue that the three main forces of urbanisation are backward and forward linkages of firms such as access to intermediate and consumer markets, the clustering of workers (both as in the core-periphery model), and localised interactions which promote technological innovation (as in the systems of innovation literature), but that they must be studied in terms of face-to-face contact as the key mechanism.

The findings in favour of face-to-face contact have strong implications for understanding how innovation functions in different places and the subsequent results for growth. This draws the conclusion for economic policy to be location based and take account of local conditions regarding economic geography, industrial clustering and the mode and scale of competition.

2.4 Similarities and differences

This section summarises the similarities and differences between economic geography and endogenous growth, particularly examining the role of innovation. While many aspects have already been described in the previous sections, the purpose of this section is to bring the three areas together as a foundation for developing the models presented in the rest of the thesis.

2.4.1 Similarities

The techniques used in endogenous growth models and NEG have many similarities. In endogenous growth models with horizontal innovation, consumers have a love for variety (Baldwin et al., 2003). Similarly, both models have monopolistic competition and increasing returns. This gives firms and agents incentives about how to invest in innovative activities or where to locate. This eventually led to the development of endogenous growth models that combine features of the NEG. The fundamental use of Dixit-Stiglitz competition means modelling approaches are particularly compatible.[1] Product variety models (horizontal innovation) work well with core-periphery models in the NEG and growth literature (Baldwin et al., 2003). Chapter 3 describes how endogenous growth theory has been combined with the NEG in this way. These similarities suggest there are further opportunities for combining the two areas of literature.

The systems of innovation literature also has similarities with both endogenous growth and NEG. Both systems of innovation and endogenous growth models recognise that the source of growth is innovation. They also recognise the circular causality that innovation develops knowledge which is used in developing further innovations. This is referred to in the literature on economic dynamics as a positive feedback loop. This recognition of innovation within endogenous growth models suggests future models could benefit further from the insights of the innovation systems literature. While some models just assume that innovation occurs with a proxy for knowledge as an input, models should also take account of how knowledge is used, created and geographically constrained.

Similarly, the innovation and economic geography literatures recognise regional differences, agglomeration and the importance of cities. Modelling techniques, such as those used with other spatial phenomenon like transport costs, could be similarly applied to innovation inputs. As such, endogenous growth models can benefit from findings in both the innovation and economic geography literatures. For example, some endogenous growth models include human capital and economic geography also considers the spatial distribution of this production factor. While NEG models include trade and transport costs these can also be brought into endogenous growth models as is shown in Chapter 3.

2.4.2 Differences

However, there are also key differences that cannot be so easily combined. Endogenous growth models with vertical innovations assume firms with monopoly power. However, models with horizontal innovation and models in the NEG use monopolistic competition and have not previously accounted for the Schumpeterian effect of creative destruction. Up until now there has been no progress to combine Schumpeterian or quality ladders models (vertical innovation) with the core-periphery model. As far as the author is aware, the models presented in Chapters 4 and 5 are the first endogenous growth models with quality improvements or creative destruction to incorporate two regions and spatial externalities. In this type of model, creative destruction allows the innovating firm to achieve a higher profit than their competitors with depreciated varieties and the innovator would have a preference for production in a larger region.

[1] A full exploration of the contributions that followed from the publication of Dixit and Stiglitz (1977) can be found in Brakman and Heijdra (2001).

Economic geography offers implications for innovation and economic growth. Martin and Sunley (1998) note the limitations of endogenous growth theory when applied to regional development. Human capital, technological progress and economic growth are all uneven in spatial economics. Future growth theories need to include the implications of economic geography in both the growth outcomes and the inputs to technological progress. The dissemination of innovations, technology and knowledge across regions, between countries and throughout the economy is affected by the geographic characteristics of those economies.

Core-periphery models have difficulty including geographical mechanisms for knowledge spillovers. Inequality between regions in these models is largely because of transport costs and does not recognise the other advantages of cities and agglomeration. Future models should take account of the implications offered by the systems of innovation literature to geographically model knowledge spillover mechanisms. To develop a mechanism for knowledge spillovers is challenging. Understanding that knowledge is largely tacit, knowledge spillovers are highly localised. Knowledge transfer is highly dependent on interaction and face-to-face contact as in McCann (2007) which presents a mechanism for showing how firms and agents choose location and their level of innovation through the intensity of face-to-face contact.

Furthermore, innovation as a concept is difficult to model within endogenous growth theories due to its nature. It is a broad concept generally referring to some knowledge inputs and potentially a profit motive. But the many kinds of innovations, the different ways inputs develop into innovations, and the multi-stage process from basic science and invention to its application as an innovation and its development as a market, are far too complex for endogenous growth theory to fully account for in an elegant and tractable model. Introducing some complexity requires the use of simulations to understand the effects on growth from a better understanding of innovation.

Despite this complexity, the models presented in this thesis in Chapters 4 to 6 add new understandings of innovation to endogenous growth theory in three elegant and mostly tractable theoretical models. Where necessary to aid understanding, the results of simple simulations are included along with analysis of the implications for economic growth and policy.

2.5 Conclusion: Incorporating a better understanding of innovation into endogenous growth

The thesis develops new models of endogenous growth including factors affecting innovation, such as space, clustering and competition. In the style of Aghion and Howitt (1992), Romer (1986, 1990, 1994) and Lucas (1988), knowledge and innovation are endogenised in the growth process. However, the model that is developed in Chapter 4 includes parameters to represent the transaction costs of inter-regional and inter-varietal knowledge transfers. The model is extended in Chapter 5 to account for a related technology dimension of knowledge spillovers. Lastly, Chapter 6 considers the effect on growth of barriers to entry to each of many industrial sectors, as well as the level and mode of competition in each sector.

The models include several extensions of Young (1998) by incorporating additional manufacturing sectors (either agricultural goods, multiple manufacturing sectors or both), knowledge spillovers between sectors and lastly characteristics of market entry and competition. The models include the forward and backward linkages of NEG models which lead research and development, skilled migrants, and manufacturing to cluster in cities. Peripheral regions are not able to support the higher wages required for skilled mobile workers. But by adding mechanisms for industrial clustering in Chapter 5, it is possible for a peripheral region to sustain a specialised industry with higher wages. Hence, there are unique implications for growth policy in isolated or distant regions.

The thesis builds on previous related research. Nijkamp and Poot (1998) explore models of endogenous growth in interdependent regions. Acs and Varga (2002) and Acs (2002) ask similar questions and also suggest combining NEG, endogenous growth and innovation as a method of developing a model of technology-led economic development. The models in Chapters 4 and 5 of this thesis go one step further, and consider a knowledge-led spatial model of endogenous economic growth, where knowledge spillovers are the prerequisite to innovation-led growth but are restricted by regional economic geography.

The models are not limited to national growth and should actually be of more use to represent regions, cities, or areas where there is some localisation of industry. In this sense, the models focus on a spatial unit where transaction costs to knowledge transfers within the region are relatively low but significantly higher between regions. For mathematical convenience and elegance, models are reduced to two regions and a number of assumptions remove any unnecessary complications. Elegance and tractability is maintained as much as possible and implications derived regarding the mechanism of economic growth and appropriate policy responses. Each chapter considers how the results of these models translate to policy implications in the case of n regions, continuous space or real world situations. These models move the debate regarding growth beyond that of endogenous growth models to one that involves location as a strategic choice in the innovation process as well as the production process.

Chapter 3

Incorporating space in the theory of endogenous growth: Contributions from the New Economic Geography[1]

This chapter describes how endogenous growth theory has now incorporated spatial factors. Both product variety models of endogenous growth and NEG models typically use Dixit-Stiglitz competition. Increasing returns provide an incentive for innovation in endogenous growth theory and, in combination with transport costs, increasing returns provide an incentive for firm location decisions in the NEG. Since innovation is the engine of endogenous growth and knowledge spillovers are a key input to innovation, the chapter also explores how innovation and knowledge have distinctly spatial characteristics. The reader is guided through the modelling of space in endogenous growth theory via the NEG. When space is included, growth is enhanced by agglomeration because of the presence of localised technology spillovers. The chapter considers some of the many other spatial factors included in models of space and growth explores the spatial effects on economic growth demonstrated by these theoretical models. Lastly, policy implications for integration are considered beyond lowering trade costs and how lowering the cost of trading knowledge is a stabilising force and is growth enhancing.

[1] An earlier version of this chapter is published as Bond-Smith and McCann (2014). Incorporating space in the theory of endogenous growth: Contributions from the new economic geography. In Fischer, M. M. and Nijkamp, P., editors, *Handbook of Regional Science*, pages 213–236. Springer Berlin Heidelberg.

3.1 Introduction

Theoretical models of endogenous growth explain the engine of economic growth with intentional investments in innovation motivated by monopolistic competition. But these theories have typically ignored space. Endogenous growth and NEG have grown quite separately despite similarities in modelling using Dixit-Stiglitz (1977) preferences. Within the literature on innovation, contributions on systems of innovation and the geography of innovation (Audretsch and Feldman, 1996) have the potential for a number of spatial aspects to also be incorporated into the theory of growth. More recently, endogenous growth theory has been combined with the NEG and provided insights on how geographic space can influence economic growth.

There are persistent differences in growth rates and incomes between even highly integrated regions, such as the European Union or the United States. Endogenous growth theory offers some explanations for varying growth rates. Firms invest in research and development to design new innovations, whereby knowledge of existing products is an integral input to research and development. Profits provide an incentive for investment and are protected by patents. The theory implies that varying rates of economic growth may be caused by regions specialising in different sectors with varying rates of productivity or rates of innovation and by differing institutions that protect patents. The theory fails to provide an adequate explanation of varying growth rates, because it does not explain differences in levels of innovation when regions have similar institutions, or innovations that are not protected by institutions (for example, process innovations, firm structure, etc.). Spatial factors offer some explanation for these differences, but economic growth theory typically ignores the role of space in determining economic growth outcomes.

Kaldor (1970) explains how trade can drive apart even identical regions as industry agglomerates in a single location. Some contributions to endogenous growth theory include this trade mechanism (Lucas, 1988; Grossman and Helpman, 1991c, 1995) but still ignore the role of space (distance-related factors) in economic growth. Despite the increasing use of space in economic theory through developments in new trade theory (Krugman, 1979), NEG (Krugman, 1991b) and similarities in modelling, it has only been a recent development to incorporate geographic space into growth theory to create spatial models of endogenous growth (Martin and Ottaviano, 1999). Developments from the NEG have now led to the incorporation of spatial factors related to both production and knowledge into theoretical growth models. These types of models may help explain varying growth rates between even highly integrated regions with similar institutions. For example, McCann (2009) suggests an economic geography perspective of New Zealand might help explain the difference in growth rates with Australia.

Hence, the NEG and growth literature incorporates space into the theory of growth by combining endogenous growth theory with the NEG. The chapter starts by describing the basic theory of endogenous growth (Romer, 1990; Grossman and Helpman, 1991a) followed by a typical NEG and growth approach where the theory accounts for the spatial factors of transport costs, migration and imperfect knowledge spillovers. Lastly the chapter reviews the contribution of these types of spatial models, variations in the use of spatial parameters and discuss the consequences for regional growth policy.

3.2 A simple model of endogenous growth

Endogenous growth theory uses increasing returns as an incentive for firms to make intentional investments to develop innovations. In all theoretical models of growth, the accumulation of capital (physical and human capital) is the engine of growth. Romer (1986, 1990), Lucas (1988) and Aghion and Howitt (1992) treat investment in innovation as investment in an additional type of capital, with increasing returns. While accumulation of physical and human capital suffers from diminishing returns, returns to investment in innovation are not diminishing and growth is sustained in the long run. These theoretical models are separated into two groups: Grossman-Helpman-Romer models (Romer, 1990; Grossman and Helpman, 1991a) use a love of variety with Dixit-Stiglitz (1977) competition and an increasing number of varieties as the source of growth. Alternatively, Schumpeterian growth models (Aghion and Howitt, 1992) use creative destruction or quality ladders (Grossman and Helpman, 1991b) where higher quality products replace existing varieties.

This section presents a simple product variety model of endogenous growth through research and development (Romer, 1990; Grossman and Helpman, 1991a, Ch. 3). Subsequent sections explore how space is added to this model of endogenous growth. To focus on innovation, the models here overlook factor accumulation (such as investments in physical and human capital) so that all investment is in the form of creating new technologies (innovations). In the product variety model, growth comes from an expanding variety of goods. In the model here, these are treated as final goods using Dixit-Stiglitz preferences as in Grossman and Helpman (1991a, Ch. 3). In contrast, Romer (1990) has an expanding variety of intermediate goods which are used to make a final good with a Dixit-Stiglitz production function. Grossman and Helpman (1991a, Ch. 3) acknowledge the alternative Dixit-Stiglitz specification of the production function rather than the utility function. Romer's (1994) model was briefly described in Subsection 2.2.5 and the outcomes of the model are essentially the same. In this section the final goods version is used for consistency with the global and local spillovers models in Baldwin et al. (2003) and the models developed in later chapters of the thesis.

There are two sectors – final goods and the research and development sector. Labour is either employed in producing final goods or in research and development which produces new designs. Workers are free to choose the sector in which they are employed and supply their labour inelastically. Consumers have a taste for diversity and are made better off by an expanding number of varieties. For each new good there is a sunk cost of innovation that occurs once, when the product is developed. Each firm must first obtain a design from the research and development sector, but once the design is obtained the firm can produce that variety forever at a constant marginal cost. The presence of fixed costs leads to monopolistically competitive markets. The up-front cost is financed by monopoly profits that are later earned in sales.

3.2.1 Demand

The representative consumer is infinitely lived and has intertemporal preferences:

$$U = \int_{t=0}^{\infty} e^{-\rho t} \ln C_t dt, \tag{3.1}$$

where C_t is the consumption index of goods, ρ is the rate of time preference and time is indexed by t (for simplicity, the subscript t will be dropped hereafter where the time dimension is clear). Consumers have CES preferences over the continuum of final goods $[0, K]$:

$$C = \left[\int_0^K c_i^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1, \tag{3.1a}$$

where c_i is the consumption of variety i and σ is the constant elasticity of substitution. Consumers have a taste for diversity over an infinite set of products $i \epsilon [0, \infty]$, where at any point in time a subset K is available in the marketplace. Consumers allocate income between consumption and savings, distributing consumption across available varieties. Intertemporal utility optimisation implies that expenditure changes over time according to the Euler equation $\frac{\dot{E}}{E} = r - \rho$ where E is consumer expenditure, $\dot{E}$ is expenditure differentiated with respect to time, and r is the risk free rate of return on savings. In equilibrium, $r = \rho \, \forall t$. Subject to the budget constraint

$$\int_0^K c_i p_i di \leq E, \tag{3.1b}$$

consumers allocate expenditure across varieties to maximise utility. With aggregate consumption defined as $C = \left[\int_0^K c_i^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}}$ in the utility function, the price index is defined as

$$P = \left[\int_0^K p_i^{1-\sigma} di \right]^{\frac{1}{1-\sigma}}, \tag{3.1c}$$

where p_i is the price of variety i, so that $E = CP$. It can be shown (for example Appendix A derives the demand functions in Chapter 4 and Appendix 2.A of Baldwin et al. (2003) derives demand functions similar to those here) that the demand function facing an individual firm is:

$$c_i = E p_i^{-\sigma} P^{\sigma-1}. \tag{3.1d}$$

Hence, demand is inversely related to relative price.

3.2.2 Production

There are L workers/consumers. Each worker produces one unit of labour per period and supplies its labour inelastically. It is assumed that each firm takes other firms' prices as given and, with a large number of varieties, firms ignore the effect of their own price on the price index. Some portion of labour $L_M < L$ is employed in the manufacturing sector. Remaining labour is employed in the innovation sector L_I such that $L_I + L_M = L$. Firms choose optimal prices in order to maximise profit

$$\pi = (p_i - \beta w)\, c_i, \tag{3.2}$$

where β describes the marginal units of labour per unit of production and w describes the wage, such that βw is the marginal cost of production. This implies that optimal prices are a constant mark-up over marginal cost:

$$p_i = \frac{\sigma}{\sigma - 1}\beta w, \qquad \forall i \epsilon \left[0, K\right]. \tag{3.2a}$$

It is possible to impose a normalisation $\beta = \frac{\sigma-1}{\sigma}$, which implies that $p_i = w$. This normalisation is not done here so that the reader can identify in the formulas how the productivity of labour in both manufacturing and research and development is important for the allocation of labour between sectors. As there are symmetric demands for all varieties, every firm in the manufacturing sector yields the same price, production and operating profit per period. Expenditure is normalised such that $E = 1\, \forall t$. With this normalisation and all firms being otherwise identical, individual firm profit in each period is given by:

$$\pi_i = \frac{E}{K\sigma} = \frac{1}{K\sigma}. \tag{3.2b}$$

3.2.3 Research and development

A manufacturing firm has a one-off fixed cost to develop the patent for the good (or purchase it from an entrepreneur) in the research and development sector which generates designs for new varieties of final goods. Each variety requires one unit of knowledge capital produced by the research and development sector. Individual firms face an innovation cost of a_I units of labour per unit of capital produced. The model follows Grossman and Helpman (1991a, Ch. 3) and Romer (1990) by using a learning curve approach such that the marginal cost of new knowledge capital a_I declines as cumulative knowledge output increases. Romer (1990) rationalises this assumption by referring to the non-rival nature of knowledge, emphasising the role of knowledge spillovers. Labour and the stock of knowledge (equal to the number of varieties) are used to develop new innovations. Each innovation adds to the stock of knowledge that can be used for developing future innovations. Innovation production is given by:

$$\dot{K} = \frac{L_I}{a_I}, \qquad F = wa_i, \qquad a_I = \frac{1}{K_t}, \tag{3.3}$$

where $\dot{K}$ is knowledge capital differentiated over time, L_I is the labour employed in the innovation sector, F is the fixed cost of innovation to develop a new variety in the research and development sector and $a_I = \frac{1}{K_t}$ describes the productivity of the research and development sector increasing with cumulative output (i.e. the fixed cost of each innovation decreases over time). The model described is based on Grossman and Helpman (1991a), but for consistency the functional form used here is adapted from Baldwin and Forslid (2000). From Equation 3.3 it follows that the rate of growth in the number of varieties is equal to L_I, which may be scaled by a constant to calibrate the model as described in Grossman and Helpman (1991a, Ch. 3).

An entrepreneur seeks funding for upfront costs (wages for L_I) from credit markets (or provides that credit in foregone wages). It is assumed there are no frictions in credit markets and no aggregate uncertainty so the purchasing of a patent can also be thought of as the entrepreneur issuing debt or equity (or some combination). Once a patent has been obtained, the manufacturer has monopoly rights to produce variety i forever at constant marginal cost. Equity owners are paid the infinite stream of profits from the firm. Free entry into the research sector implies that labour is hired such that the wage of workers in the research sector equals its marginal product (in manufacturing). At time t and with constant interest rates, the present value of the future stream of profits, v_t, is:

$$v_t = \int_{s=t}^{\infty} \pi_t e^{-r(s-t)} ds. \tag{3.4}$$

Differentiating and rearranging the “no arbitrage” condition,

$$\dot{v}_t = -e^{-r(s-t)}\pi_t + rv_t \tag{3.4a}$$

is obtained. This can also be written as a rate of return,

$$\frac{\dot{v}_t}{v_t} + \frac{\pi_t}{v_t} = r. \tag{3.4b}$$

The "no arbitrage" condition describes that in the interval between t and $t + dt$ the owners of the patent (equity holders) receive a return $\frac{\dot{v}_t}{v_t}$ (made up of the profit rate $\frac{\pi_t}{v_t}$ and the rate of capital gain or loss) equal to the yield on a risk-less loan. In other words, for a manufacturing firm to purchase a patent (or investors to hold equity/debt) the payoff must exceed the opportunity cost.

The cost of research that yields $\dot{K} = \frac{l_I}{a_I}$ incremental varieties is wa_I and has the value $v_t\dot{K} = v_t\frac{l_I}{a_I}$ where l_I is the labour input by a typical entrepreneur. Continuous growth, $\dot{K} > 0$, involves an active research sector and free entry requires the research costs to be equal to the value of research for all t. If the costs of research are greater than the value of innovation, no research would occur in equilibrium. A situation where the cost of research is less than the value of research and development will never occur in equilibrium, because it would cause an unbounded demand for research labour. Equilibrium therefore requires $v_t \leq wa_I$ with equality when $\dot{K} > 0$.

3.2.4 Equilibrium

Rather than deriving the full equilibrium this section discusses the equations that govern the steady state. For a full discussion of steady state conditions for similar models see Grossman and Helpman (1991a), Baldwin and Forslid (2000) or Baldwin et al. (2003). In equilibrium, there is a flow of new innovations:

$$\dot{K} = \begin{cases} \frac{L}{a_i} - \frac{\beta}{v} & \text{for } v > \frac{\sigma-1}{\sigma}\frac{a_I}{L} \\ \\ 0 & \text{for } v \leq \frac{\sigma-1}{\sigma}\frac{a_I}{L}, \end{cases} \tag{3.5}$$

with $L = L_I + L_M$, i.e. total employment is the sum of research and development employment and manufacturing employment. Substituting the interest rate $r = \rho$ and the profit rate $\pi = \frac{1}{K\sigma}$ into the "no arbitrage" condition, the change in firm value is a function of the value of a firm and the number of firms:

$$\dot{v} = \rho v - \frac{1}{K\sigma}. \tag{3.6}$$

These two differential equations (Equations 3.5 and 3.6) describe the dynamic equilibria.

3.2.5 Balanced growth

If conditions allow for employment in research and development, there are an increasing number of varieties. As firms compete for a fixed supply of labour, the output per firm and the value of a firm goes down over time. Research into new varieties remains profitable since the cost of innovation decreases as the number of varieties increases. The steady growth rate of the number of varieties, $\frac{\dot{K}}{K}$, is denoted by g_K. Defining a new variable, $V = \frac{1}{Kv}$, that represents the inverse of the economy's aggregate equity value, the growth rate is given by:

$$g_K = \frac{\dot{K}}{K} = \begin{cases} L - \beta v & \text{for } V < \frac{\sigma}{\sigma-1}\frac{L}{a_I} \\ \\ 0 & \text{for } V \geq \frac{\sigma}{\sigma-1}\frac{L}{a_I}. \end{cases} \tag{3.7}$$

These definitions also imply that $\frac{\dot{V}}{V} = -g_K - \frac{\dot{v}}{v}$. Substitution of Equation 3.6 obtains

$$\frac{\dot{V}}{V} = \frac{1}{\sigma}V - g_K - \rho. \tag{3.7a}$$

The model is reduced to one differential equation and the condition for growth given by Equation 3.7. The steady state rate of innovation is calculated by setting $\dot{V} = 0$:

$$g_K = \frac{L}{\sigma} - \beta\rho. \tag{3.7b}$$

This is positive, so long as $L > \rho(\sigma - 1)$, otherwise growth is zero. Growth is positively related to the scale of the economy (L), which is a common property of these models. Innovation (and incentives for research and development investment) is sustained, because there are offsetting forces of declining profits due to expanding varieties and falling product development costs due to research externalities.

This is not the overall growth rate of the economy. Macroeconomic growth is described by the growth rate of the consumption index, $C = \left[\int_0^K c_i^{\frac{\sigma-1}{\sigma}} di\right]^{\frac{\sigma}{\sigma-1}}$. Since $E = CP = 1 \forall t$, growth is also the rate at which the price index $P = \left[\int_0^K p_i^{1-\sigma} di\right]^{\frac{1}{1-\sigma}}$ declines. The growth rate of consumption g_C can be shown to be $g_C = \frac{g_K}{\sigma-1}$. This is also not the growth rate of gross domestic product (GDP). GDP is defined as the value added in both manufacturing and research and development. GDP grows at a rate equal to a weighted average of the growth rates of the index of manufacturing output/consumption and of research output. Since research and development is usually only a small percentage of a country's GDP, the difference is negligible. See Grossman and Helpman (1991a, p. 63) for a discussion.

3.3 A two-region model of growth

Virtually all endogenous growth models rely on technical externalities such as knowledge spillovers and production externalities. Endogenous growth models usually assume a frictionless spillover of knowledge. The reality is that knowledge is not transferred so effortlessly. While some knowledge can be codified and transferred easily, much knowledge is at least partially tacit. Spillovers of tacit knowledge occur over space and time through face-to-face contact (McCann, 2007) and migration (Faggian and McCann, 2009). Eaton and Kortum (1999) show that knowledge spillover and production externalities are related to the geographic distribution of manufacturing and research and development. A better understanding of the economics of innovation (Nelson, 1993) and its geographic characteristics (Audretsch and Feldman, 1996) significantly improves our understanding of economic growth.

Innovation is a predominantly local event and is now included in economic geography. Acs and Varga (2002) note the similarities between modelling techniques of endogenous growth theory and the NEG, suggesting a new model of technology-led regional economic development that combines the two fields with insights from the economics of innovation. Knowledge and innovation also have space, time and cost characteristics in their spillover between locations. This role of space and time in knowledge spillovers means economic growth also has spatial characteristics.

Given this understanding of innovation, the concentration of economic activity also results in greater knowledge spillovers between firms in concentrated locations. In the endogenous growth literature, there is recognition of partial international knowledge spillovers. Grossman and Helpman (1991c) model foreign knowledge as an innovation input in a small economy where the availability of foreign knowledge is dependent on the level of trade, yet these models ignore the role of space. Spatial characteristics incorporate domestic knowledge transfers, shared infrastructure, institutions, education and other methods of knowledge spillovers that might be unrelated to international trade. Space can be added to the theory of growth by including spatial characteristics in knowledge spillover inputs to innovation production.

Transport costs are also a key spatial parameter typically ignored in endogenous growth models. The new trade theory (Krugman, 1979) and the NEG (Krugman, 1991b) include transport costs and have Dixit-Stiglitz competition in common with many theoretical endogenous growth models. Transport costs can therefore be included easily within endogenous growth. The result of transport costs is the concentration of production in specific locations, when transport costs reach some low threshold. This is known as the core-periphery model. With low enough transport costs, firms choose to locate close to their customers to reduce transport costs. When models also allow for migration, workers choose to locate near producers to reduce their cost of living. These transport-cost-related phenomena are known as the home market effect, because it causes the concentration of firms and people.

Higher transport costs may induce firms to seek locations where there are fewer firms to compete with. This is known as the market crowding effect. It is the balance of these two effects that determines equilibrium and the steady state. Concentration occurs at low transport costs when the home market and cost of living effects dominate the market crowding effect, while dispersion occurs at higher transport costs, where market crowding dominates. The NEG suggests that imperfect integration may create regional winners and losers (Krugman, 1991b; Krugman and Venables, 1995). A particularly interesting characteristic is that the economic conditions of two regions can be exactly the same, yet yield dramatically different economic outcomes.

3.3.1 Incorporating space in the theory of growth

Models of NEG and growth combine horizontal innovations à la Grossman-Helpman-Romer with the NEG (e.g., Martin and Ottaviano, 1999; Baldwin and Forslid, 2000; Baldwin et al., 2001; Fujita and Thisse, 2003) predominantly due to the fundamental use of Dixit-Stiglitz competition. Different NEG models of growth vary assumptions on the mobility of capital, labour and industry or consumer demand to influence the forward and backward linkages. This section describes a typical NEG approach to modelling growth (Baldwin and Forslid, 2000) that includes the spatial factors:

- location,
- migration,
- transport costs,
- local knowledge spillovers and
- imperfect global knowledge spillovers.

The model has two regions that trade. There is a traditional goods sector with perfect competition that employs immobile unskilled workers L_T. Consumers have a taste for traditional goods such that $C = C_M^{\mu} C_T^{1-\mu}$ where C_M is the index of manufactured goods (similar to C in the previous section) and C_T is the traditional goods sector. Foreign region variables are denoted by a tilde $(\tilde{})$. The representative consumer is infinitely lived and has intertemporal preferences:

$$U = \int_{t=0}^{\infty} e^{-\rho t} \ln \left[C_M^{\mu} C_T^{1-\mu} \right] dt. \tag{3.8}$$

In what follows the time subscripts will again be suppressed for simplicity. Transport costs are zero in the traditional goods sector and workers in this sector cannot migrate between regions. In the real world, workers in the traditional goods sector are not necessarily unskilled or immobile. The important feature here is that the factor of production for traditional goods is immobile and "unskilled" is the commonly used term in these models.

The purpose of the additional sector in this model is that some residual demand remains in the periphery, even when there is full agglomeration, so that regions continue to trade.

Skilled workers (L_K) are employed in either manufacturing or innovation (similar to workers in the previous section with subscript K since they work in the knowledge sectors of manufacturing or innovation). The world population of skilled and unskilled workers is normalised to one such that $L = L_K + L_T = 1$. Skilled workers and manufacturing firms have a choice of location. Skilled workers respond to wage pressure when making a decision to migrate between regions. If there are differences in real wages there will be migration. The perfect price index describes the price index of utility and therefore includes traditional goods such that $P = P_T^{1-\mu} P_M^{\frac{-\mu}{\sigma-1}}$. The change in skilled workers in the home region is given by the ad hoc migration equation in Fujita et al. (1999):

$$L_K = \left(\omega_K - \tilde{\omega}_K\right) s_H \left(1 - s_H\right), \tag{3.9}$$

$$s_H = \frac{L_K}{L_K + \tilde{L}_K}, \qquad \omega_K = \frac{w}{P}, \qquad \tilde{\omega}_K = \frac{\hat{w}}{\hat{P}}, \tag{3.9a}$$

where $\dot{L}_K$ is skilled labour in the home region differentiated over time, s_H is the share of skilled workers in the home region and ω_K is the real wage of skilled workers in the home region. Since the real wage is defined by means of the perfect price index, workers migrate to the region that provides the highest level of utility.

Manufactured goods transported between regions incur transport costs that take Samuelson's "iceberg" form (Samuelson, 1952) where transport costs are incurred in the good itself. The manufacturer produces more of the good than actually arrives, because some portion of the good "melts" in transit. If τ represents the proportion of the final good that arrives at the destination, the remaining portion is used up during transportation. Hence $\tau < 1$ is a measure of the freeness of trade or an index of the inverse of transport costs. Transport costs for the traditional goods sector are assumed zero ($\tau = 1$). Firms are incentivised to locate in the largest market to minimise transport costs. From the migration equation above, skilled workers try to locate in the region with more firms, as this reduces their cost of living (since they have a taste for diversity) by increasing real wages.

So far this chapter has added space with migration and transport costs which affect manufacturing but there are also spatial effects in the production of innovations. Knowledge does not transfer completely between regions so not all knowledge is available to entrepreneurs when manufacturing is shared between regions. Innovation is included in the manufacturing sector the same as in the endogenous growth model of Section 3.2, but now with partial spillovers of knowledge between regions. Individual firms face the innovation cost of a_I units of labour for each unit of knowledge capital produced. Innovation production in the home region is given by:

$$\dot{K} = \frac{L_I}{a_I}, \qquad F = wa_i, \qquad a_I = \frac{1}{K_t + \lambda \tilde{K}_t}, \qquad 0 \leq \lambda \leq 1, \tag{3.10}$$

where $\dot{K}$ is knowledge capital differentiated over time, L_I is the skilled labour employed in the innovation sector, λ is the ability for foreign knowledge to be used in the home region and $a_I = \frac{1}{K_t + \lambda \tilde{K}_t}$ describes how productivity of the research and development sector increases with cumulative output and varies between locations due to the geographic nature of knowledge spillovers. Hence, the model assumes perfect local knowledge spillovers but imperfect spillovers between regions. The parameter λ represents how space affects knowledge production such that firms choose a location that considers how existing knowledge can be used for innovation. In this way, firms are attracted to regions where other firms are located, because the cost of innovation is lower.

3.3.2 Model description

Reconsider the product variety model of Section 3.2 together with these additional spatial factors. Again, world expenditure is normalised, $E_w = 1 \forall t$. Subject to the budget constraint, consumers allocate expenditure across varieties to maximise utility. Hence in the home region $P_M C_M + P_T C_T \leq E$, where P_M is the local price index of manufactured goods (the world equivalent are weighted average manufacturing price and consumption indices such that $\overline{P_M C_M} + P_T C_T \leq E_w$) and P_T is the price of traditional goods. Consumers spend a constant portion (μ) of their expenditure on manufactured goods and the rest on traditional goods:

$$P_M C_M = \mu E, \qquad P_T C_T \left(1 - \mu\right) E. \tag{3.11}$$

The traditional goods sector is perfectly competitive, with 1:1 technology (one unit of unskilled labour input yields one unit of traditional goods output) and constant returns to scale. Total production of traditional goods is shared across both regions. Let L_T and $\tilde{L}_T$ be the supply of unskilled workers in the home and foreign regions respectively. The model follows Krugman (1991b) by setting the worldwide stock of skilled workers to μ and the stock of unskilled workers to $1 - \mu$, shared equally between regions:

$$L_T = \frac{1 - \mu}{2}, \qquad \tilde{L}_T = \frac{1 - \mu}{2}. \tag{3.12}$$

The choice of units ($1 - \mu$ unskilled workers and μ skilled workers) follows Krugman (1991b) and ensures that prices and wages in the traditional goods sector are the numéraire, and that the nominal wage rate of skilled workers equals that of unskilled workers in the steady state. If the number of skilled workers were specified differently, the wages of skilled workers are a constant multiple of the wage rate of unskilled workers. Simplicity is maintained by avoiding this additional multiple. A scaling factor could also be used to calibrate the model to any arbitrary growth or wage rate.

Unskilled workers provide one unit of production per period, i.e. $\int_0^{L_T} C_T + \int_0^{\hat{L}_T} \tilde{C}_T = 1 - \mu$. Free trade ensures the same nominal price of traditional goods and equal nominal wages in the two regions. With full employment of $1-\mu$ unskilled workers and 1:1 technology, the traditional goods sector is the numéraire,

$$\begin{aligned} w_T \left(\int_0^{L_T} C_T + \int_0^{\hat{L}_T} \tilde{C}_T \right) &= w_T (1-\mu) \\ &= P_T \left(\int_0^{L_T} C_T + \int_0^{\hat{L}_T} \tilde{C}_T \right) \\ &= P_T (1-\mu) = E_W (1-\mu) , \end{aligned} \tag{3.13}$$

$$w_T = P_T = \tilde{w}_T = \tilde{P}_T = 1. \tag{3.13a}$$

The remainder of the analysis focuses on the manufacturing sector. The home region produces K manufactured varieties and the foreign region produces $\tilde{K}$ varieties. Consumers have CES preferences over the continuum of manufactured goods $\left[0, K + \tilde{K}\right]$, such that

$$C_M = \left[\int_0^{K+\tilde{K}} c_i^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}}, \qquad \sigma > 1, \tag{3.14}$$

where c_i is the consumption of variety i and σ is the constant elasticity of substitution. Defining the local price index of manufactured goods as in the model of Section 3.2, $P_M = \left[\int_0^{K+\tilde{K}} p_i^{1-\sigma} di \right]^{\frac{1}{1-\sigma}}$, where p_i is the price paid by local consumers, the demand function in the home region facing an individual manufacturer is

$$c_i = \mu E p_i^{-\sigma} P_M^{\sigma-1} \tag{3.15}$$

and the equivalent demand function exists in the foreign region with the foreign region's price index.

Manufacturing firms in each region face the same optimisation problem as in the endogenous growth model: $\max_{p_i} = [p_i - \beta w_M] c_i$, where βw_M is the marginal cost of production. Firms ignore the effect of their own price on the index. Once again optimal prices for home region manufacturers are a constant mark-up over marginal cost and transport costs are passed on directly to consumers:

$$p_i = \frac{\sigma}{\sigma-1} \beta w_M, \qquad \tilde{p}_i = \frac{\sigma}{\sigma-1} \beta \tilde{w}_M, \qquad \forall i \epsilon [0, K] , \tag{3.16}$$

where p_i and $\hat{p}_i$ are the local and export prices of a home manufacturer. A foreign manufacturer has analogous prices, with transport costs on goods exported to the home region. Here it is also possible to impose the same normalisation $\beta = \frac{\sigma-1}{\sigma}$ such that $p_i = w_M$ and $\tilde{p}_i = \frac{w_M}{\tau} = \frac{p_i}{\tau}$. While its distribution is subject to worker migration, by following Krugman's (1991b) choice of units where the worldwide stock of skilled workers is μ, nominal skilled wages in equilibrium are $w_M = 1$ or $\tilde{w}_M = 1$ for the core-periphery outcome and $w_M = \tilde{w}_M = 1$ in the symmetric outcome.

3.3.3 Long-run location

The long run is characterised by a "steady state": defined by an unchanging growth rate in the number of manufactured varieties, its regional division, as well as the prices and quantities defined by the short-run equilibrium above. Migration of skilled workers due to spatial inequality of real wages leads to the long-run equilibrium. With the migration equation above, and particularly the role of the perfect price index in this equation, it can be seen that real wages will only be unequal when one region has a larger share of manufacturing. As in Section 2.3.1, this is the cost of living effect such that consumers minimise the cost of living by locating in the region with the most manufacturers. When one region is larger than the other, the larger region is also the lowest cost location for innovation to occur, because of greater knowledge spillovers. This is described as the innovation agglomeration effect. Furthermore, at low levels of transport costs there are higher profits in the larger region, due to a combination of the market access and innovation agglomeration effects. At high levels of transport costs, and only a slightly unequal equilibrium, there may be higher margins in the smaller region due to the market crowding effect which dominates the market access, cost of living and innovation agglomeration effects and would return the system to the equal distribution outcome. The balance of these forces for agglomeration and dispersion imply that there are two long run types of steady states:

- the equal distribution outcome and
- the core-periphery outcome.

See Baldwin and Forslid (2000) for a more formal discussion of the conditions of the steady state in the NEG model of growth here and Baldwin et al. (2003) for a discussion of other NEG and growth models. The equal distribution outcome is where both regions have half the skilled workers, half the manufacturing and half the traditional goods production. The other steady state is the core-periphery outcome where all manufacturing concentrates in a single region (either home or foreign) known as the core and only unskilled workers (the traditional goods sector) remain in the other region known as the periphery. Traditional goods production is split equally between regions.

If there are asymmetric transport costs, it is not inevitable that the region with the lowest transport costs will be the core. The core region will be the one which has the higher share of varieties and where the difference in the number of varieties is large enough to trigger a switch from the equal distribution outcome to the core-periphery outcome. This could be for several reasons. Since every variety has a patent forever, hysteresis plays a large role in determining which region is the core. For example, an initial higher endowment of resources might lead to a greater number of manufacturers and innovators, or greater infrastructure investment at some stage (and temporarily freer trade) might also trigger agglomeration. Similarly, temporarily different policy settings between regions where one region has favourable policies for research and development could lead to initially higher rates of innovation, a greater share of varieties and agglomeration. While not included in typical NEG models of growth, stochastic effects could mean one region gets "lucky". In the model here, innovations are simply costs where each firm has to employ a certain amount of skilled labour in research and development in order to achieve an innovation. In reality, successful innovations are not so guaranteed. The inclusion of probabilistic outcomes in the research and development sector could mean one region achieves a higher rate of innovation by luck, resulting in it becoming the core.

Figure 3.1, reproduced from Baldwin and Forslid (2000), but with a different measure of trade freeness, describes the possible equilibria with different combinations of trade freeness and knowledge diffusion. As the level of trade freeness increases (i.e. transport costs decline), the break point τ_B describes the level of trade freeness where the equal distribution outcome is no longer a steady state. The sustain point τ_S describes the level of trade freeness at which the distribution of firms and workers switches from the core-periphery outcome to the equal distribution outcome when trade freeness is declining (transport costs increase). The values of trade freeness between the sustain and break points represent situations in which both the potential equilibria outcomes are stable. As the level of knowledge spillovers (λ) varies, so do both the break and sustain points. Figure 3.1 describes how the break and sustain points increase as knowledge spillovers increase.

Alternatively, Figure 3.1 describes the combinations of knowledge spillovers and trade freeness that result in stable (and unstable) equilibria for both the equal distribution outcome and the core-periphery outcome. There are three sections within the knowledge spillover (λ) and trade freeness (τ) space. In the top-left corner the core-periphery equilibrium is unstable and the equal distribution is stable. In this situation, trade freeness is sufficiently low (high transport costs) that the market crowding effect means firms make greater profit by locating away from other firms. There is very little trade (if any) between regions. Closer to the curve, regions will trade, but the market crowding effect always dominates the home market, cost of living and innovation cost effects. In the middle section, both the equal distribution and core-periphery equilibrium are stable. If there is an equal distribution, regions will trade, but it is possible that with an external shock, the home market, cost

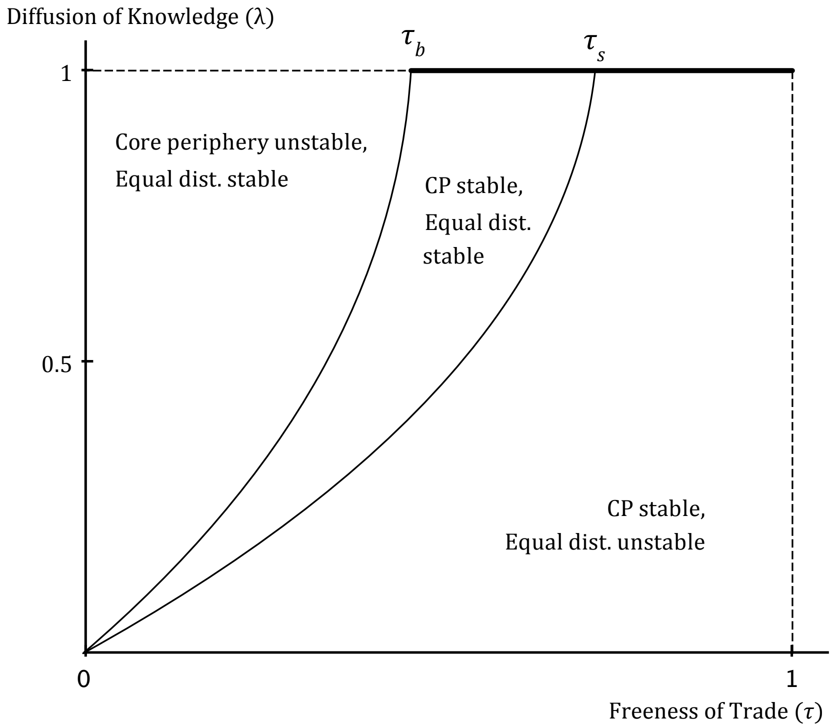

Figure 3.1: Core-periphery and symmetric equilibrium stability map
Source: Reproduced from Baldwin and Forslid (2000)

of living and innovation cost effects could dominate the market crowding effect and the system would switch to the core-periphery outcome. Similarly, if there is a core-periphery equilibrium, an external shock to the distribution could lead to the market crowding effect dominating the home market, cost of living and innovation cost effects causing a switch to the equal distribution outcome. Lastly, the bottom right section describes combinations of knowledge spillovers and trade freeness where only the core-periphery outcome is stable. In this situation, the home market, cost of living and innovation cost effects always dominate the market crowding effect.

3.4 Spatial consequences for economic growth

The incorporation of space in the theory of growth means the model recognises the role of space through transport costs and through knowledge not transferring perfectly between locations. This section considers world and regional growth in both the short and long run in the two possible types of equilibria: core-periphery and equal distribution. Because regions are able to trade, even a periphery region benefits from growth in the number of varieties produced in the core. Over time the price index for manufactured varieties falls as more varieties are invented and producers of traditional goods experience growth in real income, because they trade for manufactured goods. Growth is considered both in terms of the number of manufactured varieties, but also growth in terms of the consumption bundles available to all consumers.

3.4.1 Market integration

While traditional conceptions of market integration refer to lowering of the cost of trading goods, Figure 3.1 shows that incorporating space and growth gives a more detailed view where integration can also be viewed as lowering the cost of trading ideas. Integration policies which focus solely on free trade may be destabilising and result in a deindustrialisation of the periphery region. That is, when integration reduces trade costs alone, the region that emerges as the periphery suffers relative to the region that emerges as the core. Integration policies that also focus on knowledge spillovers (or entirely on knowledge spillovers) will be growth-enhancing for both regions. The model shows how this form of integration is stabilising while pure trade cost integration can be destabilising.

3.4.2 Growth in varieties

The number of manufactured varieties worldwide evolves according to:

$$\dot{K} + \dot{\tilde{K}} = \frac{L_I}{a_I} + \frac{\hat{L}_I}{\hat{a}_I}, \qquad a_I = \frac{1}{K + \lambda \tilde{K}}, \qquad \hat{a}_I = \frac{1}{\lambda K + \tilde{K}}. \tag{3.17}$$

For simplicity, the subscript i is removed from p_i, because home firms are symmetric and prices are equal for all home firms. Once a blueprint or variety is invented, manufacturers require β marginal units of labour per unit of production so aggregate demand for labour in the manufacturing sector in the home region is $\frac{\beta}{p}$.

As in the endogenous growth model without space, equilibrium in the skilled labour market in the home region requires $L_K = L_I + L_M = a_I \dot{K} + \frac{\beta}{p}$. In the equal distribution outcome, prices are higher than the core-periphery outcome because of the additional cost to transport goods between regions. A larger share of skilled labour is used in manufacturing, because each producer has to produce a larger amount to cover the cost of transport. In other words, the cost of transport increases the marginal cost of production such that some labour is no longer available for innovation. When freeness of trade is greater, i.e. the cost of transport is lower, more labour is available for growth. As such, incorporating space in the theory of growth shows how trade liberalisation and agglomeration are growth-enhancing for world growth.

Turning to regional growth, the number of manufactured varieties in the home region evolves according to:

$$\dot{K} = \frac{L_I}{a_I}, \qquad a_I = \frac{1}{K + \lambda \tilde{K}}. \tag{3.18}$$

Trade liberalisation and agglomeration (in the home region) are growth-enhancing, because they reduce the cost of transport. However, if transport costs induce the core-periphery outcome, there is no manufacturing in the periphery and therefore no growth in varieties produced by that region. That is, reducing transport costs means growth in varieties may be limited to a specific region(s). Therefore, trade liberalisation is not growth-enhancing for growth in varieties for the region that emerges as the periphery.

For both world and regional growth the inclusion of space means firms face an innovation cost that is dependent upon location. The output of skilled workers in the innovation sector is greater when knowledge is more easily available. With s being the home region's share of manufacturing, the rate of growth is $\frac{L_I^W}{K+\tilde{K}}\left[s\left(K+\lambda\tilde{K}\right)+(1-s)\left(\lambda K+\tilde{K}\right)\right]$. That is, when λ is greater, both world and regional growth increase, unless there is a change in the distribution of economic activity (refer to Figure 3.1). Including space in the theory of growth shows how closer economic integration is growth-enhancing for world and regional growth in varieties.

Similarly, when one region has a greater share of manufacturing than the other region, growth increases for the agglomerated region. Agglomeration in either region is growth enhancing for world growth and for regional growth in the region where agglomeration occurs. In the core-periphery outcome, there is zero growth in the number of varieties in the periphery region as no varieties are manufactured there, no skilled workers are employed and no innovation occurs. However, it is possible that a change in λ could induce a dispersion of economic activity between regions and therefore reduce agglomeration. If the reduction in innovation from moving to the equal distribution outcome is greater than the increases in innovation from a higher λ, then growth could potentially be lower despite a higher λ. So while λ is growth enhancing, the locational effects of a change in λ could still be growth diminishing, depending upon whether it induces a change in the distribution of economic activity.

3.4.3 Consumption growth

While so far the chapter has described the effect of space on the growth rate of the number of varieties, this is not the overall growth rate, because consumers also have a preference for traditional goods. The growth rate of the overall economy, should reflect the growth rate of what people actually consume. In other words growth describes the rate of increase in what the income to workers allows those consumers in each region to purchase, which is measured by the growth rate of the consumption index, C, where $E = CP = 1$. This best describes how the well-being of consumers increases over time. While there is no growth in the number of varieties in the periphery, as the region only produces traditional goods, the ability to trade for manufactured goods allows the unskilled workers to benefit from innovations and increases in the number of varieties produced in the core.

Since $E = CP$ the rate at which the consumption index grows is the rate at which the perfect price index declines. In the endogenous growth model of Section 3.2, the growth rate of consumption g_C was shown to be $g_C = \frac{g_K}{\sigma-1}$. With the addition of the traditional goods sector, the overall perfect price index is to a power of $\frac{\sigma-1}{\mu}$. The perfect price index is falling at a rate of $g_C = \frac{\mu g_K}{\sigma-1}$. Notably, the growth rate of consumption is the same in both regions, whether it is the symmetric outcome or a manufacturing concentration outcome. This is because the price index for both regions falls at the same rate, since consumers in both regions still spend the same portion of their earnings on traditional goods – in the steady state the growth rate of consumption is equal in both regions. The inclusion of space does not explain the differences in growth rates between locations in the long run. Instead space affects the world rate of growth and the share of wealth/earnings in each location.

In the short run, however, there can be different growth rates between locations if the regions are in transition between steady states. Given $\tau < 1$, the price index will be permanently lower in a core location, because core location consumers do not pay transport costs for manufactured goods. If the economies are shifting from an equal distribution to the core-periphery outcome, growth rates in the periphery will be temporarily lower (or even negative) as periphery consumers transition to paying transport costs on a greater share of the manufactured goods they consume (eventually all goods). Consumers in the core gradually pay transport costs on a smaller share of manufactured goods and the core will have higher growth rates.

3.4.4 Agglomeration and freeness of trade

Agglomeration is growth-enhancing in the long run through both transport costs and knowledge spillovers. Agglomeration minimises the total cost of transport if all manufacturing and the majority of consumption is in one location. Agglomeration is also growth-enhancing, because it increases knowledge spillovers if all research and development occurs in one location. Increased freeness of trade is growth-enhancing in the long run, but in the short run the outcome is ambiguous. Increased freeness of trade is always growth-enhancing if there is no change in the distribution of economic activity. However, as described in Figure 3.1, increased freeness of trade can lead to a switch from the equal distribution outcome to the core-periphery outcome. While this is significantly growth-enhancing for the region that becomes the core, it is temporarily growth-diminishing for the periphery while the two regions transition to the new equilibrium.

3.4.5 Impact of knowledge spillovers upon economic growth

Knowledge spillovers are generally growth-enhancing. Increased knowledge spillovers mean firms have a lower cost of innovating and therefore there is a greater growth rate in varieties and consumption. In the core-periphery equilibrium, increasing knowledge spillovers has no effect on growth, because knowledge is unaffected by space, since all production is in a single location.

However, as with agglomeration, the effect is ambiguous if there is a change in the steady state. A large enough increase in knowledge spillovers could lead to a switch from the core-periphery outcome to the equal distribution outcome (see Figure 3.1). With a change in the location of production from one region to multiple regions, the knowledge spillover parameter now has an effect on growth when there was previously no effect. That is, firms initially had access to all knowledge, because all manufacturing was in the same region, but in the new steady state foreign knowledge is only partially available. While knowledge spillovers are generally growth-enhancing, there is the possibility of knowledge spillovers being growth-diminishing in the former core region if it brings about the sharing of manufacturing.

In the steady state where production is shared between locations, knowledge spillovers are growth-enhancing. Furthermore, knowledge spillovers also make production in the equal distribution outcome more stable. That is, increasing knowledge spillovers means changes in trade costs are less likely to lead to a switch to the core-periphery outcome (see Figure 3.1). With greater knowledge spillovers, production in both regions is a stable equilibrium for a greater range of trade freeness.

3.5 Variations for incorporating space in the theory of growth

In the NEG and growth literature, there are many variations of the model presented here. These include differences in the mobility of labour or capital, the inclusion of intermediate goods, heterogeneous firms, multiple labour types and heterogeneous skill levels. Other areas of economics also incorporate space by using continuous space (rather than discrete regions), by defining location on an interval, by incorporating land as a factor of production and by introducing congestion costs. All of these variations have different effects on the role of space, location and geography on growth, but in general incorporating space in the theory of growth has similar effects to those presented here.

3.5.1 Mobility of labour and capital

The model here describes the typical approach by NEG scholars to incorporating space in the theory of growth with the inclusion of migration of skilled labour. The effect of footloose skilled labour can lead to catastrophic agglomeration which means the model is unable to show other unequal internal steady states. The model in this chapter includes skilled worker migration to demonstrate the role of firm and worker location choices and how migration influences innovation. Highly skilled workers and innovators are internationally mobile so it is important to consider how this affects the location of innovation and subsequently economic growth.

Capital mobility is the ability for capital to shift between locations. In all endogenous growth models, growth comes from the accumulation of capital. Capital can come in a number of forms: human capital, physical capital or knowledge capital. Labour and education can be thought of as human capital, which is able to migrate between locations in the model above. Physical capital is the equipment used in production such as machinery and production plants. This has been excluded from the model above. Knowledge capital is the ideas generated in the innovation sector which are marketable and tradable through patents. This is the type of capital commonly modelled in the endogenous growth literature.

There are two options for the mobility of knowledge capital. With mobile capital, the owners of capital can decide where to locate production. If knowledge capital is mobile, the number of innovations produced (and owned) by one region may be different from the number of firms actually producing in that region. That is, the developer of a patent may choose to produce in a region other than her own. In this situation, the decision to accumulate capital is the same in all locations: the mobility of capital eliminates demand-linked causality such that the shifting of production does not shift the location of consumption or the earnings from owning a manufacturing firm. Alternatively with immobile capital, the owners of capital are only able to produce within the region where they are located. With immobile capital, any shift that favours production in one location leads to new capital in that region. Since owners are local this also leads to expenditure shifting and further production shifting via the home market effect.

In many NEG models of growth, such as in Martin and Ottaviano (1999), Baldwin et al. (2001) and Baldwin and Martin (2004), migration is not allowed. In these models, workers are instead completely mobile between traditional, manufacturing and innovation sectors but not between regions. These models require an extra assumption that a single country's labour endowment must not be enough to meet global demand for traditional goods, to avoid complete specialisation in manufacturing goods only.

In models with labour immobility and capital mobility, the owners of capital are indifferent between producing in either region in the steady state because the steady state has equalised rates of return in either region. However, with localised knowledge spillovers, innovators prefer to be located in the region with the highest level of manufacturing. Despite the differences, these models reach similar steady states to the model presented above. In particular, space has the same effects on growth, because space is included using the same mechanisms with localised knowledge spillovers and transport costs. Agglomeration is growth-enhancing due to localised knowledge spillovers and knowledge spillovers are growth enhancing, because they reduce the cost of innovation.

In models without labour or capital mobility, agglomeration is enabled by either vertical linkages in production or the spatial influence on knowledge creation and transfer. If NEG models of growth have immobile capital and mobile labour these models have the same catastrophic agglomeration described by the model above (and most NEG models), because innovation occurs at a faster rate in a region with greater capital and this is self-reinforcing as all new firms prefer to innovate in the location with the largest share of manufacturing. Whenever labour is mobile, agglomeration is catastrophic.

However, models with immobile capital and immobile labour offer an alternative advantage of unequal internal solutions. That is, a range of transport costs and knowledge spillovers that yields steady states where one region has a larger share of manufacturing (but not all) than the other. As there is no migration, this means the region with the larger share of manufacturing has a share of traditional goods production smaller than the other region's share of traditional goods production. Even though these models ignore the role of migration in economic activity and growth, it allows the effect on knowledge spillovers and growth to be considered even if there are unequal levels of agglomeration.

The result is very similar to the core-periphery outcome. Growth rates are equal in both regions, because consumers in the low manufacturing region still benefit from innovations made in the high manufacturing region because of trade. Similarly, the growth rate in varieties is greater in the high manufacturing region because of localised knowledge spillovers. Real wages are also higher in a high manufacturing region. Without migration, there is no mechanism to equalise real wages between regions.

This is explored more closely in the trade literature and although this literature does not explicitly model space, the analysis could take a spatial interpretation. For example Grossman and Helpman (1991c) suggest knowledge spillovers are related to the volume of trade between two locations. As a result, policies that encourage trade between locations are growth enhancing. However, they also note the limitations of this approach and that knowledge and learning are much more complex than either a trade volume (Grossman and Helpman, 1991c) or a parameter driven approach (Baldwin et al., 2003). In Chapters 4 and 5 this thesis explores the implications of a more complex form of knowledge spillovers.

Another advantage of modelling with labour immobility between regions is there is no need for the modelling trick of the Krugman (1991b) core-periphery model which fixes the share of skilled and unskilled workers. Instead labour mobility between sectors equalises real wages between manufacturing and traditional sectors within each region and zero transport costs in the traditional goods sector equalises nominal wages. While some of these features are mathematically elegant, the model described in this chapter includes skilled labour migration, to demonstrate more of the many spatial features that can be incorporated into spatial models of endogenous growth.

3.5.2 Vertically-linked industry

Other types of NEG and growth models have vertically-linked industry following the practice of some NEG models (Krugman and Venables, 1996; Venables, 1996). This is where goods are a factor of production. For example, final goods may be produced from a variety of manufactured intermediate goods (Yamamoto, 2003), manufactured goods may be produced using a variety of manufactured goods (which have not been consumed) and/or the innovation sector could use manufactured goods as a factor of production (Martin and Ottaviano, 1999).

If the vertical linkage is in the innovation sector, this generates a feedback between growth and agglomeration with a similar result to localised knowledge spillovers. Martin and Ottaviano (1999) do not use the localised knowledge spillover mechanism demonstrated here and, instead, the innovation sector uses manufacturing goods as an innovation input such that the location of manufacturing affects the cost of innovation through trade costs. Similarly Yamamoto (2003) describes a model where final goods and innovation are produced using manufactured intermediate goods. This vertical linkage creates a circular causation mechanism in growth and agglomeration because of the vertical linkages between intermediates and innovation.

3.5.3 Other characteristics

There are many different factors which affect firm location decisions and subsequently space, innovation and economic growth. This chapter explored how these are dealt with in models that combine the NEG with endogenous growth by recognising localised knowledge spillovers. But there are many more modelling choices for spatial factors which influence growth. For example, by studying heterogeneous firms, Baldwin and Forslid (2010) help describe the characteristics of firms that choose to locate in core or lagging regions. Other models include land requirements and continuous space (Desmet and Rossi-Hansberg, 2009), whereby every firm is in a different location but willing to pay higher land rents to access more valuable locations. All of these have some influence on location choices for firms but ultimately demonstrate the same role of space in growth – that space is a barrier to knowledge transfer and technology diffusion which are inputs to innovation – and that policies or decisions by firms that reduce these spatial costs are growth-enhancing.

3.6 Conclusions

This chapter describes how NEG and growth models incorporate space into the Grossman and Helpman (1991a, Ch. 3) product variety model of endogenous growth. Incorporating space into endogenous growth increases the complexity of these theoretical models. In all of these models with full local knowledge spillovers and partial global knowledge spillovers, space affects growth, and growth affects location. The circular causality reinforces the core-periphery outcome of the NEG models. This chapter shows that integration between regions is more complex than is described by international trade models. In particular, the cost of transferring knowledge between locations is important for firm location, stability, innovation and growth.

From the discussion of the effect of space on growth through freeness of trade, agglomeration and knowledge spillovers, there are a number of implications for economic policy in different locations. Agglomeration, freeness of trade and knowledge spillovers are generally growth-enhancing. The natural conclusion is that closer integration of economies will lead to increased growth rates. However, in these spatial models of growth, integration has two dimensions: trade costs and knowledge spillovers.

While traditional conceptions of integration refer to lowering of the cost of trading goods, Baldwin and Forslid (2000) show that combining theories of growth and space produce a more subtle view of integration where integration can also be viewed as lowering the cost of trading information. Integration policies which focus solely on free trade may be destabilising and result in a deindustrialisation of the periphery region. Alternatively, integration policies that also focus on knowledge spillovers, (or entirely on knowledge spillovers) will be growth-enhancing for both regions. The model here shows how this form of integration is stabilising, while pure trade cost integration can be destabilising.

While lowering trade costs induces uneven development, it also results in higher rates of economic growth. Alternatively, policies that improve knowledge spillovers improve stability of the location of economic activity. Growth policies should consider the effect of trade, knowledge spillovers, labour and capital market integration.

The approach demonstrated in this chapter is typical of two region growth models that incorporate space into the theory of growth by using increasing product variety as the engine of growth, as in the models in Baldwin et al. (2003). The similarity of modelling techniques means the approach is a natural progression, but it comes with certain limitations, particularly in distinguishing between sources of knowledge or characterising the nature of competition. As far as the author is aware, prior to this thesis, endogenous growth models with quality improvements or creative destruction have not incorporated spatial externalities in two region models. This thesis addresses the defecit in the literature with new models presented in Chapters 4 and 5 that offer insights about the spatial and growth features of different countries or regions based on their economic geography and implications for innovation and growth policy. In this type of model, creative destruction allows innovating firms to achieve a higher profit taking the market from their competitors with depreciated varieties. Furthermore, the innovator has a preference for production in a larger region but also prefers to choose a location alongside other producers with compatible knowledge inputs to innovation.

Chapter 4

A regional model of growth with creative destruction

This chapter develops a two-region growth model where innovations are produced in the form of quality improvements building on available knowledge from all industries. Improved varieties replace obselete versions, such that there is creative destruction of the value in old varieties. Entrepreneurs prefer a location with greater availability of knowledge to maximise the productivity of research and development and maintain or capture a niche monopoly by developing a quality-improving innovation. The partial nature of spillovers between locations causes an additional force for agglomeration: the clustering effect where firms prefer locations with greater access to technological knowledge spillovers. The ability to transfer knowledge between locations is found to be destabilising, because it increases the mobility of firms or varieties to shift between locations.

4.1 Introduction

There are many dimensions of market integration considered by the economic literature, including trade costs, labour mobility, firm location, efficiency and specialisation. Economic growth and prosperity is high on a list of priorities justifying market integration policies and research has found that market integration has consequences for growth through firm location decisions (Baldwin and Forslid, 2000). Technology levels are a key determinant in firm location, yet have been left out of theoretical geography and growth models. This chapter explicitly models technology as a determinant of location choice to maximise the productivity of research and development by using a quality ladders or creative destruction approach to innovation and growth resulting in quite different conclusions to the two region expanding variety growth models of Baldwin et al. (2003) as described in Chapter 3.

Innovations build on existing knowledge and given the evidence that technical externalities such as knowledge spillovers are related to the spatial distribution of economic activity (Audretsch and Feldman, 1996), by following the same Marshallian logic, it is reasonable to assume that firm location is similarly influenced by the technology level of nearby manufacturers. While some knowledge and technology can be codified and transferred easily, much knowledge is at least partially tacit and spillovers of tacit knowledge occur over space and time through face-to-face contact (McCann, 2007), interaction and migration (Faggian and McCann, 2009). This role of space and time in technology transfer means that innovation and subsequently growth also have space and time characteristics, contributing to a wide variance in innovation activity and growth between different cities, regions and nations. Empirical evidence also shows that productivity growth is related to the spatial distribution of productivity levels (Coe and Helpman, 1995) and research and development (Keller, 2002; Jacobs et al., 2002). Yet, existing spatial models of growth rely on increasing product variety. Romer (1990) and Grossman and Helpman (1991a, Ch. 3) offer no avenue to consider technology or quality as a factor in firm location decisions other than increasing productivity due to a love for variety.

While Helpman (1992) finds increasing-variety models and quality ladders models are largely equivalent, this chapter shows that this is not necessarily the case when modelling location decisions that are dependent on the ability for knowledge to transfer between locations and on the productivity of innovation activity. Increasing-variety models of growth have been extended by applying spatial factors to knowledge spillovers, resulting in models where growth affects location and location affects growth (Martin and Ottaviano, 2001). This chapter adds to the literature by using a quality ladders approach to growth (Grossman and Helpman, 1991b; Aghion and Howitt, 1992; Young, 1998) within a two-region model and including related variety effects so the mobility of firms and entrepreneurs is now related to the ability for technology to transfer between locations and varieties. Product variety models suggest location as the only difference between sources of knowledge. A quality ladders model provides a unique opportunity to distinguish between three mechanisms for knowledge spillovers: inter-temporal, spatial and varietal knowledge spillovers. This chapter rectifies the problem that existing models can not describe the effects of integration on vertical innovation.

With the market contestable via quality improvements the model here suggests increasing the ability to transfer knowledge between locations is destabilising, because it makes technology (and entrepreneurs) more mobile by reducing the opportunity cost of switching from the agglomeration to the periphery, but has no opportunity cost when switching between equivalent locations. Agglomerated firms prefer to remain in their incumbent location and this preference decreases with increases in regional knowledge spillovers, but there is no equivalent impact on the equal distribution outcome. This is in contrast to increasing variety models which conclude that increasing spatial knowledge spillovers is stabilising, because increasing the ability for knowledge to transfer between locations has a greater impact on decreasing innovation costs when manufacturing is dispersed. The model here implies alternative impacts on the steady state from changes in the ability for knowledge to transfer between locations.

The chapter is organised as follows. Section 4.2 and 4.3 describe the mechanics of the model, equilibria and forces for firm location decisions. Section 4.4 explores the effect on growth of knowledge spillovers and agglomeration, developing policy implications for growth and integration of different regions. Section 4.5 provides a brief summary of the conclusions drawn from earlier sections.

4.2 A model of growth with partial knowledge spillovers

Consider a model with two production sectors: one for differentiated manufactured varieties and the other for traditional goods. There exists an research and development sector which produces quality improvements in manufactured varieties using skilled labour and technological knowledge spillovers that are subject to spatial, temporal and varietal characteristics. A firm must obtain an innovation produced by the research and development sector with its associated quality level, in every period, prior to the period of production at that quality level.

4.2.1 Model structure

There are two types of labour: unskilled and skilled labour. Following Young (1998) and Grossman and Helpman (1991b), manufacturing firms employ skilled workers in the research and development sector to develop a quality improvement in a selected variety for which the firm is then granted a patent. In the subsequent period, firms employ unskilled workers to produce each firm's variety with the patented quality level. To enhance tractability, the labour market is specified acording to the footloose entrepreneur model (Forslid and Ottaviano, 2003). Skilled labour is employed in the research and development sector only, with unskilled labour used in the production of manufactured goods and traditional goods. Unskilled labour is immobile between regions but mobile between manufacturing and traditional sectors while skilled labour migrates between regions in response to differences in real wages. Workers in the manufacturing and traditional goods sectors are not necessarily unskilled. The important property is that unskilled labour as a factor of production is immobile between regions. The two regions are referenced by home and foreign where foreign variables are noted by a tilde $(\tilde{\ })$.

This labour market specification is used here rather than the specification in Krugman (1991b) because of its tractability and that it does not require Krugman's (1991b) modelling trick. The alternate specification is used in Chapter 5 where skilled labour is used in both manufacturing and innovation while unskilled labour produces traditional goods only. The "trick" is to specify the quantity of labour as the same quantity demanded for the goods that each labour type produces. With 1:1 productivity for traditional goods, the specification results in equal nominal wages for skilled and unskilled workers.

The research and development sector uses knowledge of existing quality levels of all varieties to develop quality improvements. However, existing technological knowledge does not transfer perfectly between varieties or locations. Firms are attracted to locate alongside other firms in order to have greater access to local technological knowledge. Firms compete for market share where consumers have Dixit-Stiglitz (1977) preferences, prefer higher quality products and have inter-temporal preferences. With $\rho > 0$ as the discount rate, the representative consumer has inter-temporal utility given by:

$$U = \sum_{t=0}^{\infty} \alpha^t \ln Q_t, \qquad \alpha = \frac{1}{1+\rho}. \tag{4.1}$$

Q_t is the consumption of both traditional and manufactured goods in period t given by:

$$Q_t = C_{M,t}^{\mu} C_{T,t}^{1-\mu}, \qquad 0 < \mu < 1, \tag{4.1a}$$

where $C_{T,t}$ is the consumption in period t of the traditional good, $C_{M,t}$ is the consumption of n and $\tilde{n}$ differentiated varieties of manufactured goods from the home and foreign regions respectively:

$$C_M = \left[\sum_{i=1}^{n} (A_i c_i)^{\frac{\sigma-1}{\sigma}} + \sum_{i=1}^{\tilde{n}} \left(\tilde{A}_i \tilde{c}_i \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \qquad \sigma > 1. \tag{4.1b}$$

In Equation 4.1b, c_i is the quantity consumed of variety i, manufactured in the home region, A_i is the quality level of variety i and σ represents the elasticity of substitution between any two varieties of equal quality. This is the standard CES function with the additional factor A_i, which represents the quality level of variety i. Where possible, notation is kept simple by suppressing the time variable t where all variables are in the same time period.

Inter-temporal utility optimisation implies the transversality condition and the Euler equation:

$$\frac{E_t}{E_{t-1}} = \frac{1+r_t}{1+\rho}, \tag{4.2}$$

where E_t is consumer expenditure in period t, ρ is the rate of time preference and r_t is the rate of return on savings between periods $t-1$ and t. Rearranging the first order condition for growth of consumer expenditure (Equation 4.2) gives

$$\frac{E_{t+1}}{1+r_t} = \frac{E_t}{1+\rho} = \alpha E_t. \tag{4.2a}$$

The short-run equilibrium requires the economy move to a steady state level of consumer expenditure with a constant interest rate $1 + r_t = 1 + \rho = \frac{1}{\alpha}$.

As in Section 2.3 and Chapter 3, manufactured varieties transported between locations incur transport costs that take Samuelson's (1952) "iceberg" form where transport costs are incurred in the good itself. That is, the manufacturer sells more of the good than actually arrives, because some portion "melts" in transit. If $\tau \leq 1$ represents the proportion of the variety that arrives at the destination, the remaining portion is used up during transportation. τ is a measure of the freeness of trade or an inverse index of transport costs. Transport costs for the traditional goods sector are assumed zero, ($\tau = 1$).[1]

The cost of innovation is dependent on the knowledge inputs, the targeted quality level and the wage rate of skilled labour as in Young (1998). The model allows for multiple products and endogenous growth through creative destruction without scale effects. The approach of using a model without scale effects is used throughout the remainder of the thesis, because scale would naturally encourage agglomeration. The models presented in Chapters 4 and 5 demonstrate agglomeration effects explicitly due to the technical and pecuniary externalities of knowledge spillovers and transport costs. Chapter 6 also uses this model of growth without scale effects to demonstrate the effects of innovation on competition and fairly compare firms and industries of different sizes and technology levels.

In each period, a manufacturing firm produces for its niche monopoly and conducts research and development to ensure a quality improvement large enough to maintain its niche monopoly (or capture a different niche variety) for the following period. A firm must conduct enough research and development to ensure a quality improvement great enough to enter the market in subsequent periods. Knowledge is assumed to be embedded in a manufacturing facility, where this knowledge can be used in producing quality improvements in either region, subject to spatial externalities. When a new quality improvement is developed, it replaces the former version of that variety. Mobility of skilled labour between locations equalises wages for skilled workers. The skilled labour requirement in the previous period, $t-1$, to achieve the targeted quality level $A_{i,t}$ and the fixed cost of manufacturing in the subsequent period t is:

$$F_i\left(A_{i,t}, \overline{A_{i,t-1}}\right) = \begin{cases} \gamma e^{\eta \frac{A_{i,t}}{\overline{A_{i,t-1}}}} & \text{if } A_{i,t} \geq \overline{A_{i,t-1}} \quad \text{and} \\ \gamma e^{\eta} & \text{otherwise,} \end{cases} \tag{4.3}$$

[1]Other models (Baldwin et al., 2003) use a different measure of transport costs where T units must be shipped in order for one unit to arrive at the destination. They use a corresponding freeness of trade measure $\phi = T^{\sigma-1}$. For comparison with the model here, $T = \frac{1}{\tau}$ and $\phi = \tau^{1-\sigma}$.

where γ and η are fixed parameters for calibration and $\overline{A_{i,t-1}}$ is an index of technological opportunity representing the knowledge spillover that is an input to innovation production.

The model accounts for geographic factors by taking the approach of Baldwin and Forslid (2000) where knowledge transfers imperfectly between firms that are geographically separated. The portion of knowledge spillovers that is sourced from geographically separate location is weighted by the regional knowledge spillover parameter λ_R. As a result, the location of manufacturing becomes a factor in a firm choosing its location. This approach is grounded by empirical literature measuring spatial barriers to knowledge transfer between geographically separated locations (Eaton and Kortum, 1999; Keller, 2002; Keller and Yeaple, 2013).

While in Young (1998) and Grossman and Helpman (1991b) the research and development sector uses skilled labour and only its own variety's technological knowledge to develop the next quality improvement, the index can easily be modified so skilled workers can also use some knowledge from other varieties. There is an opportunity to differentiate between the externalities present in different sources of knowledge spillovers. The innovation production function used here distinguishes between the externality effects on knowledge spillovers depending on their source and separates the technical externality in research and development between variety and spatial effects. The knowledge spillover input to innovation is made up of two components: the knowledge associated with the quality level of the firm's own variety and a portion of the knowledge associated with the quality level of all other varieties. The knowledge input from all manufactured varieties is assumed to be weighted according to a related-variety approach (Boschma, 2005; Frenken and Boschma, 2007), where the relatedness of products describes how useful knowledge is to innovation in the firm's own variety. The knowledge of a firm's own variety is fully understood and carries a weighting of one, while knowledge from other varieties is weighted by its relatedness λ_V. The relatedness of knowledge is assumed constant for all pairs of varieties $\frac{\lambda_V}{n+\tilde{n}}$. Just as manufacturers or innovators can be thought of as being separated by geographic space as in the previous chapter, it's also possible to think of manufacturers of different varieties as separated by a varietal knowledge space reflected in the parameter λ_V.

$\overline{A_{w,t-1}}$ is defined as the observed worldwide weighted average of technological knowledge of all manufactured varieties in the period prior to production:

$$\overline{A_{w,t-1}} = \frac{1}{n_{t-1} + \tilde{n}_{t-1}} \left[\sum_{j=1}^{n} A_{j,t-1} + \lambda_R \sum_{j=1}^{\tilde{n}} \tilde{A}_{j,t-1} \right]. \tag{4.3a}$$

where $A_{j,t-1}$ represents the period $t-1$ quality level of each firm's variety. If the variety has been produced before, the index of technological opportunity is the firm's own quality level plus a portion of knowledge from other varieties:

$$\overline{A_{i,t-1}} = \max\left(A_{i,t-1}, \lambda_R \tilde{A}_{i,t-1}\right) + \lambda_v \overline{A_{w,t-1}}, \tag{4.3b}$$

where λ_V represents the weight of the knowledge spillover from all other varieties. If the variety has never been produced before the index of technological opportunity is given by which ever is greater from an average of local quality levels or the geographically weighted average of foreign quality levels:

$$\overline{A_{i,t-1}} = \max\left(\frac{\sum_{i=1}^{n} A_{i,t-1}}{n_{t-1}}, \frac{\lambda_R \sum_{i=1}^{\tilde{n}} \tilde{A}_{i,t-1}}{\tilde{n}_{t-1}}\right) + \lambda_v \overline{A_{w,t-1}}. \tag{4.3c}$$

Free entry means any firm can develop a quality improvement with the relevant knowledge spillover to produce an existing variety in the following period. For modelling elegance and tractiability, it is assumed that the quality levels of each variety are equal or “symmetric”. A new variety can only be produced if the firm can profitably enter the market. The number of varieties is endogenous, so if an existing variety does not achieve the symmetric quality target, it will be the lowest quality variety available, so it would be the marginal firm which exits the market. Firms develop quality improvements in existing varieties and if existing varieties achieve less than the symmetric quality target they exit the market such that a new variety can enter with technological opportunity given by (Equation 4.3c). This additional index of technological opportunity facilitates free entry and maintains symmetry such that each firm has the same entry cost equal to the cost of innovation for any incumbent firm. The second maximum option in each index is where a variety is being transferred between regions. By assuming symmetry, this second option only applies if the new location of a variety, currently has no manufacturing capability. This is because if there were already a manufacturing presence in the region, symmetry implies that $A_i = \tilde{A}_i$ so the local quality level would determine this portion of the index of technological opportunity for a replacement variety. With no manufacturing capability, foreign knowledge must be the source of knowledge for new manufacturing. After the investment in research and development during period $t-1$, firms may produce any quantity at a constant marginal cost of unskilled wages (w_U) multiplied by the marginal cost of producing one more unit of variety i (β) which is the per unit unskilled labour requirement.

4.2.2 Short-run equilibrium

Workers provide one unit of labour per period. Let L_U and $\tilde{L}_U$ be the number of unskilled workers in the home and foreign regions respectively. The worldwide stock of unskilled workers is assumed to be shared equally between regions such that $L_U = \tilde{L}_U$ The traditional goods sector is perfectly competitive with 1:1 technology and constant returns to scale. By the nature of Cobb-Douglas preferences, μ is the share of expenditure spent on manufactured varieties and $1-\mu$ is the share spent on traditional goods. The value of total expenditure on traditional goods in both regions is $P_T\left(C_T+\tilde{C}_T\right)=(1-\mu)\left(E+\tilde{E}\right)$. Expenditure on manufactured varieties by home region consumers is given by:

$$P_M C_M=\sum_{i=1}^{n} P_i c_i+\sum_{j=1}^{\tilde{n}} \frac{\tilde{P}_i}{\tau}\tilde{c}_j=\mu E, \tag{4.4}$$

where P_i and $\tilde{P}_j$ are the domestic prices of home and foreign manufactured varieties respectively and c_i and $\tilde{c}_j$ are consumption of domestically produced and imported manufactured varieties. Similar equations exist for consumers in the foreign region with transport costs applied to imported manufactured varieties. Free trade ensures the same nominal price of traditional goods in both locations and it is assumed that a single region cannot supply economy-wide demand for traditional goods to ensure equal wages for unskilled workers, such that $\mu<\frac{\sigma}{2\sigma-1}$. Wages are equal to the marginal revenue and price of traditional goods ($w_U=P_T=\tilde{w}_U=\tilde{P}_T$) because the marginal product of labour has been set to one, which suggests making traditional goods the numéraire:

$$\begin{gathered}
w_U\left(C_T+\tilde{C}_T\right)=P_T\left(C_T+\tilde{C}_T\right)=\tilde{w}_U\left(C_T+\tilde{C}_T\right)=\tilde{P}_T\left(C_T+\tilde{C}_T\right)=\left(E+\tilde{E}\right)(1-\mu)\\
w_U=P_T=\tilde{w}_U=\tilde{P}_T=1\\
\left(C_T+\tilde{C}_T\right)=\left(E+\tilde{E}\right)(1-\mu).
\end{gathered} \tag{4.4a}$$

Skilled workers may move between locations. Let L_K and $\tilde{L}_K$ be the supply of skilled workers in the home and foreign regions respectively. Regional income is therefore $L_U+L_K w_K$, where w_K is the wage of skilled workers. As consumers in either region have the same rate of time preference, they spend the same portion of income on consumption, i.e. $\frac{E}{Y}=\frac{\tilde{E}}{\tilde{Y}}$.

In the manufacturing sector, maximisation of consumer utility finds demand functions for home consumers of individual local and imported manufactured varieties:

$$c_i=\mu E A_i^{\sigma-1}P_i^{-\sigma}P_M^{\sigma-1}, \qquad \tilde{c}_i=\mu E\tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{-\sigma}P_M^{\sigma-1}, \tag{4.5}$$

where P_M is the price quality index in the home region:

$$P_M=\left[\sum_{i=1}^{n} A_i^{\sigma-1}P_i^{1-\sigma}+\sum_{j=1}^{\tilde{n}}\tilde{A}_j^{\sigma-1}\left(\frac{\tilde{P}_j}{\tau}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{4.5a}$$

The full derivation of these demand functions and the price index can be found in Appendix A. Similar demand functions exist for consumers in the foreign region with local isomorphic prices, quality levels, expenditure and price quality indices. All varieties are symmetric such that in equilibrium, quality is the same in all varieties.

In the period prior to production, potential investors/firms choose whether to enter, and then select a product and quality improvement target. No firm can appropriate the intertemporal knowledge spillover and a fixed cost investment must be made in the period prior to each production period. It is assumed that no two firms can choose to produce the same variety and there are no economies of scope. Each manufacturing firm is infinitesimally small and there is no strategic interaction.

Manufacturing firms invest in research and development, paid for by future sales revenues. Firms choose price and quality in order to maximise discounted profits:

$$\max_{P_{i,t},A_{i,t}} \pi = \frac{(P_{i,t}-\beta)\left(c_{i,t}+\frac{1}{\tau}\tilde{c}_{i,t}\right)}{1+r_t} - F_{i,t}\left(A_{i,t},\overline{A_{i,t-1}}\right)w_{K,t-1}, \tag{4.6}$$

where $\tilde{c}_i = \mu\tilde{E}A_i^{\sigma-1}\tilde{P}_i^{-\sigma}\tilde{P}_M^{\sigma-1}$ is the demand function for foreign consumers and exported production,[2] c_i is demand from domestic consumers given by the first expression in Equation 4.5, the unskilled wage rate is equal to the numéraire, $w_{K,t-1}$ is the skilled labour wage rate in the previous period and π is firm profit. The form of competition does not matter; prices are the same under both Bertrand and Cournot competition if there is a sufficiently large number of varieties.

Each firm selects in period $t-1$ a period t quality improvement with its cost of innovation and its period t price so as to maximise profit (Equation 4.6). It is assumed that each firm takes price-setting behaviour of other firms as given and ignores the effects of its own pricing or quality decisions on the price index, i.e. the price quality indexes P_M and $\tilde{P}_M$ are treated as fixed constants when differentiating. This assumption is plausible with a sufficiently large number of firms (or a continuous specification of product variety). Similarly, firms have rational expectations of quality increases in other varieties and the effect on price. Differentiating, the first order conditions are given by:

$$\frac{\partial\pi}{\partial P_{i,t}} = c_{i,t} + (P_{i,t}-\beta)\frac{\partial c_{i,t}}{\partial P_{i,t}} + \frac{1}{\tau}\tilde{c}_{i,t} + (P_{i,t}-\beta)\frac{1}{\tau}\frac{\partial\tilde{c}_{i,t}}{\partial P_{i,t}} = 0, \qquad \text{and} \tag{4.6a}$$

$$\frac{\partial\pi}{\partial A_{i,t}} = \frac{(P_{i,t}-\beta)}{1+r_t}\frac{\partial c_{i,t}}{\partial A_{i,t}} + \frac{(P_{i,t}-\beta)\frac{1}{\tau}}{1+r_t}\frac{\partial\tilde{c}_{i,t}}{\partial A_{i,t}} - \frac{\partial\left(F_{i,t}w_{K,t-1}\right)}{\partial A_{i,t}} = 0. \tag{4.6b}$$

Free entry is assumed such that any firm could hire skilled workers to develop the next quality improvement for that variety. Free entry means that in equilibrium the marginal firm just breaks even, i.e. profits will be zero, because if profits were positive there is an opportunity for a firm to produce either an additional variety or greater quality

[2] The notation here is different from above. The ambiguity is required so the reader can recognise that these consumers are foreign. In this part of the analysis, the tilde ($\tilde{}$) refers to exported varieties and transport costs are treated as a cost of manufacturing varieties for export. Throughout the rest of the chapter, tilde otherwise refers to foreign produced varieties.

improvements and take the market from the incumbent firm. The free entry condition is:

$$\frac{(P_{i,t}-\beta)\left(c_{i,t}+\frac{1}{\tau}\tilde{c}_{i,t}\right)}{1+r_t}=F_{i,t}\left(A_{i,t},\overline{A_{i,t-1}}\right)w_{K,t-1}. \tag{4.7}$$

From Equation 4.6a and the demand functions expected pricing behaviour in domestic and export regions is found to be:

$$P_{i,t}=\frac{\sigma}{\sigma-1}\beta w_U, \qquad \tilde{P}_{i,t}=\frac{\sigma}{\sigma-1}\frac{\beta w_U}{\tau}=\frac{P_{i,t}}{\tau}, \tag{4.8}$$

where unskilled wages (which are the numéraire) are still notated for illustrative purposes only. Prices can be normalised to one using the assumption $\beta=\frac{\sigma-1}{\sigma}$. This normalisation is not necessary, but merely has the effect of altering the units of measurement for production (to achieve a price of one), so the model here retains β for illustrative purposes. The price in an export market reflects a direct pass through of transport costs. Even if prices of goods are different (due to transport costs), all varieties will still be purchased as consumers have a love for variety, but higher priced goods will be purchased with lower quantity. With these prices and symmetry, the price quality index (Equation 4.5a) can be simplified to a function of the number of firms producing in each region:

$$P_M=\left(\frac{\sigma}{\sigma-1}\right)\beta A_j^{-1}\left[n+\tilde{n}\left(\frac{1}{\tau}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{4.8a}$$

Isomorphic prices can be found in a similar way for foreign firms. The price index highlights the cost of living effect (as already mentioned in Sections 2.3.1 and 3.3.3) on the consumer's location choice. Differentiating the price index with respect to the number of varieties gives a negative value, meaning the cost of living declines with a greater number of varieties in that region. Location effects for the model in this chapter are discussed further in Section 4.3.1.

Dividing Equation 4.6b by the free entry condition (Equation 4.7) and rearranging finds equilibrium behaviour in the innovation sector. Firms choose a quality improvements to equate the elasticity of the research cost ($F_i w_K$) with respect to quality with the elasticity of demand with respect to quality. For constant elasticity demand:

$$\begin{aligned}\varepsilon_{A_{i,t}}^{c_{i,t}+\tilde{c}_{i,t}} &= \varepsilon_{A_{i,t}}^{F_i w_K}\\ \sigma-1 &= \eta\frac{A_{i,t}}{\overline{A_{i,t-1}}}\end{aligned} \tag{4.9}$$

where ε_y^x represents the elasticity of the x variable with respect to the y variable. Substituting $\overline{A_{i,t-1}}=\max\left(A_{i,t-1},\lambda_R\tilde{A}_{i,t-1}\right)+\lambda_v\overline{A_{w,t-1}}$ and rearranging, obtains:

$$A_{i,t} = \frac{\sigma - 1}{\eta} \left(\max \left(A_{i,t-1}, \lambda_R \tilde{A}_{i,t-1} \right) + \lambda_v \overline{A_{w,t-1}} \right). \tag{4.9a}$$

Equation 4.9a describes the quality target firms prefer. If one region is larger than the other, the preferred quality target is lower in the smaller (peripheral) location because of the parameter λ_R. But since a firm producing a new variety has a knowledge input to innovation given by Equation 4.3c, all firms must meet the higher target to ensure market entry or they will be pushed out of the market by a new variety. That is, spatial externalities may reflect different levels of knowledge inputs to innovation, but the targeted quality level reflects the highest level that is attainable in either region. Assuming the home location determines the preferred quality target, the fixed cost of production is the innovation cost $F_i \left(A_{i,t}, \overline{A_{i,t-1}} \right) w_K = w_K \gamma e^{\sigma - 1}$. Firms in the smaller region achieve the higher target by hiring more skilled workers but to avoid non-negative profits, these firms offer lower wages to skilled workers. The quality target required for market entry is therefore:

$$A_{i,t} = \frac{\sigma - 1}{\eta} \max \left(A_{i,t-1} + \lambda_v \overline{A_{w,t-1}}, \tilde{A}_{i,t-1} + \lambda_v \overline{\tilde{A}_{w,t-1}} \right). \tag{4.9b}$$

By symmetry the quality improvement multiplier is equal to:

$$\frac{A_{i,t}}{A_{i,t-1}} = \frac{\sigma - 1}{\eta} \left(1 + \lambda_V \frac{\max \left(\overline{A_{w,t-1}}, \overline{\tilde{A}_{w,t-1}} \right)}{A_{i,t-1}} \right). \tag{4.9c}$$

An isomorphic quality improvement multiplier can be found for foreign firms. Parameters are specified that assume Equation 4.9c is greater than one to ensure there are always quality improvements in equilibrium. With symmetry, $\frac{\overline{A_{w,t}}}{A_{i,t-1}} = \frac{n + \lambda_R \tilde{n}}{n + \tilde{n}}$ so the condition that ensures quality improvements reduces to:

$$\frac{\sigma - 1}{\eta} \left(1 + \lambda_V \frac{\max \left(n + \lambda_R \tilde{n}, \lambda_R n + \tilde{n} \right)}{n + \tilde{n}} \right) > 1. \tag{4.9d}$$

Quality improvement includes an additional multiplier to Young (1998), $\left(1 + \lambda_V \frac{\max(n + \lambda_R \tilde{n}, \lambda_R n + \tilde{n})}{n + \tilde{n}} \right)$, due to the inter-varietal knowledge spillover. Intuitively, quality improvement has two parts: it is the firm's willingness to pay for research and development plus the quality improvement required for entry due to the maximum technological spillover. Otherwise world growth is entirely determined by parameters of the research and preference functions, η and σ respectively.

4.2.3 Labour market clearing and endogenous variety

Full employment is assumed such that all skilled labour is used in research. The quality level for entry is always determined by the larger region because of knowledge spillovers.

The targeted quality level is the highest level that either the firm prefers or an entrant could capture the market: $A_{i,t} = \frac{\sigma-1}{\eta} \max\left(A_i + \lambda_v \overline{A_{w,t-1}}, \tilde{A}_i + \lambda_v \overline{\tilde{A}_{w,t-1}}\right)$. Assuming the home region has equal or a greater share of firms/workers ($L_{K,t} \geq \tilde{L}_{K,t}$), the fixed input requirement of skilled workers for firms in the home region is $\gamma e^{\sigma-1}$ and for firms in the foreign region the fixed input requirement is $\gamma e^{(\sigma-1)\frac{A_{i,t-1}+\lambda_V \tilde{A}_{w,t-1}}{\tilde{A}_{i,t-1}+\lambda_V A_{w,t-1}}}$. Isomorphic results would apply if the foreign region had the greater share of firms and workers. With symmetry, the skilled worker input is a function of the distribution of firms in the period prior to production: $\gamma e^{(\sigma-1)\frac{n_{t-1}+\tilde{n}_{t-1}+\lambda_V(n_{t-1}+\lambda_R \tilde{n}_{t-1})}{n_{t-1}+\tilde{n}_{t-1}+\lambda_V(\lambda_R n_{t-1}+\tilde{n}_{t-1})}}$. With $\chi = \frac{n_{t-1}+\tilde{n}_{t-1}+\lambda_V(n_{t-1}+\lambda_R \tilde{n}_{t-1})}{n_{t-1}+\tilde{n}_{t-1}+\lambda_V(\lambda_R n_{t-1}+\tilde{n}_{t-1})}$ representing the ratio of knowledge available in alternate locations the skilled worker input simplifies to $\gamma e^{(\sigma-1)\chi}$

Labour market clearing in period t requires $L_{K,t} = n_{t+1}\gamma e^{\sigma-1}$ and $\tilde{L}_{K,t} = \tilde{n}_{t+1}\gamma e^{(\sigma-1)\chi}$. Rearranging, the number of firms in each location is determined by the amount of skilled labour:

$$n_{t+1} = \frac{L_{K,t}}{\gamma e^{\sigma-1}}, \qquad \tilde{n}_{t+1} = \frac{\tilde{L}_{K,t}}{\gamma e^{(\sigma-1)\chi}}. \tag{4.10}$$

By the free entry condition there are no profits in equilibrium as given by Equation 5.8c. Given prices (Equation 4.8), wages of unskilled workers (numéraire) and the number of skilled workers required for a firm to develop a quality improvement, the free entry condition can be rearranged to obtain:

$$\begin{aligned} w_{K,t} &= \frac{\beta\left(c_{i,t+1}+\frac{1}{\tau}\tilde{c}_{i,t+1}\right)}{(\sigma-1)(1+r_t)\gamma e^{\sigma-1}}, \qquad \text{and} \\ \tilde{w}_{K,t} &= \frac{\beta\left(\frac{1}{\tau}c_{i,t+1}+\tilde{c}_{i,t+1}\right)}{(\sigma-1)(1+r_t)\gamma e^{(\sigma-1)\chi}}. \end{aligned} \tag{4.11}$$

where $c_{i,t+1}+\frac{1}{\tau}\tilde{c}_{i,t+1}$ is the total production by the home region firm i to be sold domestically and exported in period $t+1$. Using demand functions in Equation 4.5 (modified for demand of exported goods produced in the home region), unskilled wages in Equation 4.4a prices in Equation 4.8 and the simplified price index in Equation 4.8a, market clearing implies:

$$c_{i,t+1} + \frac{1}{\tau}\tilde{c}_{i,t+1} = \frac{\sigma-1}{\sigma\beta}\left(\frac{\mu E_{t+1}}{n_{t+1}+\tilde{n}_{t+1}\tau^{\sigma-1}} + \frac{\mu \tilde{E}_{t+1}\tau^{\sigma-1}}{n_{t+1}\tau^{\sigma-1}+\tilde{n}_{t+1}}\right). \tag{4.12}$$

Isomorphic market clearing applies to foreign firms. Substituting Equations 4.10, 4.11 and 4.2a, Equation 4.12 for home and foreign firms can be rearranged as:

$$\begin{aligned} w_{K,t} &= \frac{\mu\alpha}{\sigma}\left(\frac{E_t}{L_{K,t}+L_{K,t}e^{-(\sigma-1)(1-\chi)}\tau^{\sigma-1}} + \frac{\tilde{E}_t\tau^{\sigma-1}}{L_{K,t}\tau^{\sigma-1}+L_{K,t}e^{-(\sigma-1)(1-\chi)}}\right), \\ \tilde{w}_{K,t} &= \frac{\mu\alpha}{\sigma}\left(\frac{E_t}{L_{K,t}e^{(\sigma-1)(1-\chi)}+L_{K,t}\tau^{\sigma-1}} + \frac{\tilde{E}_t\tau^{\sigma-1}}{L_{K,t}e^{(\sigma-1)(1-\chi)}\tau^{\sigma-1}+L_{K,t}}\right). \end{aligned} \tag{4.12a}$$

4.3 Long-run equilibrium

The long-run steady state is defined by a constant number of manufactured varieties, constant growth in the level of quality of manufactured varieties and its regional division as well as the prices and quantities defined by the short run equilibrium above. Migration of skilled workers due to wage pressure leads to the long-run equilibrium, where real wages of mobile workers are equalised between locations. The real wage for a home region worker is defined as $\omega = \frac{w}{\mathbf{P}}$ where $\mathbf{P} = P_T^{1-\mu} P_M^{\mu}$ is the local perfect price quality index describing the price index of utility, P_M is the local manufacturing price quality index and P_T is the price of traditional goods. Similarly for foreign workers $\tilde{\omega} = \frac{\tilde{w}}{\tilde{\mathbf{P}}}$ where $\tilde{\mathbf{P}} = P_T^{1-\mu} \tilde{P}_M^{\mu}$. This model uses a static migration equation, where skilled workers respond to real wage differences at the start of each period. Migration of skilled workers occurs as follows:

$$L_{K,t} = L_{K,t-1} + m_t, \qquad \tilde{L}_{K,t} = \tilde{L}_{K,t} - m_t, \tag{4.13}$$

where m is the migration of skilled workers shifting from the foreign region to the home region in response to differences in wages:

$$m_t = s_{t-1} \left(1 - s_{t-1}\right) \left(\omega_{K,t} - \tilde{\omega}_{K,t}\right). \tag{4.14}$$

and s_{t-1} is the home region's share of skilled workers during the previous period, $\frac{L_{K,t-1}}{L_{K,t-1} + \tilde{L}_{K,t-1}}$. Migration also assumes ad hoc dynamics from the standard core-periphery model in Fujita et al. (1999). Migration in discrete time periods is used for elegance because of the discrete nature of innovations in the deterministic growth model framework. The same conclusions can be reached with continuous time periods, but such a model relies on stochastic innovation where the length of time between innovations varies in relation to the level of investment. The model set up here is kept simple to aid intuition by using deterministic innovations, discrete time and discrete migration. Unskilled workers cannot migrate between locations.

Consider how expenditure is distributed between regions. Consumers in either region spend the same portion of their income $Y = \frac{1}{\vartheta} E$ and $\tilde{Y} = \frac{1}{\vartheta} \tilde{E}$ where ϑ represents the proportion of income that is spent in order to satisfy Equation 4.2a. Since $L_U = \tilde{L}_U$, the term L_U is used to refer to amount of local unskilled labour in either region. Local expenditure is proportional to local income:

$$\left(L_U + L_{K,t} w_K\right) \vartheta = E, \qquad \left(L_U + \tilde{L}_{K,t} \tilde{w}_K\right) \vartheta = \tilde{E}. \tag{4.15}$$

4.3.1 Long run location

This section considers different spatial distributions of workers and economic activity and considers the forces for manufacturing concentration or dispersion. Starting with a distribution of firms operating in both regions and substituting Equation 4.15 into the wage Equation 4.11, generates a system of two equations in $w_{K,t}$ and $\tilde{w}_{K,t}$ that can be solved for equilibrium skilled wages as functions of the spatial distribution of skilled workers $L_{K,t}$ and $\tilde{L}_{K,t}$. Following the approach of Forslid and Ottaviano (2003), with $s_{t-1} = \frac{L_{K,t-1}}{L_{K,t-1}+\tilde{L}_{K,t-1}}$ describing the share of skilled workers in the home region, wages are found as functions of the skilled labour share and the current distribution of manufacturing firms which was determined by the location of skilled workers in the previous period $t-1$:

$$
\begin{aligned}
w_{K,t} &= \frac{\frac{\mu\alpha\vartheta}{\sigma}}{1-\frac{\mu\alpha\vartheta}{\sigma}} L_U \frac{e^{(\sigma-1)(1-\chi)}}{L_{K,t}+\tilde{L}_{K,t}} \frac{e^{(\sigma-1)(1-\chi)}2\tau^{\sigma-1}s_{t-1}+\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)(1-s_{t-1})}{\tau^{\sigma-1}\left(\left(s_{t-1}e^{(\sigma-1)(1-\chi)}\right)^2+(1-s_{t-1})^2\right)+e^{(\sigma-1)(1-\chi)}s_{t-1}(1-s_{t-1})\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)}, \\
\tilde{w}_{K,t} &= \frac{\frac{\mu\alpha\vartheta}{\sigma}}{1-\frac{\mu\alpha\vartheta}{\sigma}} L_U \frac{1}{L_{K,t}+\tilde{L}_{K,t}} \frac{2\tau^{\sigma-1}(1-s_{t-1})+e^{(\sigma-1)(1-\chi)}\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)s_{t-1}}{\tau^{\sigma-1}\left(\left(s_{t-1}e^{(\sigma-1)(1-\chi)}\right)^2+(1-s)^2\right)+e^{(\sigma-1)(1-\chi)}s_{t-1}(1-s_{t-1})\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)}.
\end{aligned} \tag{4.16}
$$

With migration dictated by Equation 4.14, a portion of skilled workers respond to real wage pressure $\omega_K - \tilde{\omega}_K$, by moving to the region with higher real wages. Real wages are found by dividing Equation 4.16 by the perfect price index. Real wage pressure is given by:

$$
\begin{aligned}
\omega_K - \tilde{\omega}_K &= \frac{w_K}{P} - \frac{\tilde{w}_K}{P} \\
&= \Phi \frac{V(s,\tau)}{\tau^{\sigma-1}\left(\left(se^{(\sigma-1)(1-\chi)}\right)^2+(1-s)^2\right)+e^{(\sigma-1)(1-\chi)}s(1-s)\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)},
\end{aligned} \tag{4.17}
$$

where $\Phi = \frac{\mu\alpha\vartheta L_U(\sigma-1)^\mu A_i^\mu \gamma^{\frac{\mu}{1-\sigma}}}{(\sigma-\mu\alpha\vartheta)(\sigma\beta)^\mu\left(L_{K,t}+\tilde{L}_{K,t}\right)}$ and:

$$
\begin{aligned}
V(s,\tau) = {} & e^{(\sigma-1)\chi} \frac{e^{(\sigma-1)(1-\chi)}2\tau^{\sigma-1}s+\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)(1-s)}{\left(\left(\frac{L_{K,t-1}}{e^{\sigma-1}}+\frac{L_{K,t-1}}{e^{(\sigma-1)(1-\chi)}}\left(\frac{1}{\tau}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\right)^\mu} \\
& - \frac{2\tau^{\sigma-1}(1-s)+e^{(\sigma-1)(1-\chi)}\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)s}{\left(\left(\frac{L_{K,t-1}}{e^{\sigma-1}}\left(\frac{1}{\tau}\right)^{1-\sigma}+\frac{L_{K,t-1}}{e^{(\sigma-1)(1-\chi)}}\right)^{\frac{1}{1-\sigma}}\right)^\mu}
\end{aligned} \tag{4.18}
$$

To be a steady state, real wage pressure must subside ($\omega_K - \tilde{\omega}_K = 0$) so there is no wage pressure to incentivise migration or all workers must be already in the location with the highest wages such that all skilled workers choose to remain in their incumbent locations. In addition, a steady state is stable if a small deviation (by ad hoc dynamics) results in pressure for the distribution to return to the initial state. If this small deviation results in a change in the stable distribution of skilled workers, then the original steady state is considered unstable. An interior equilibrium is found by the solutions to $V(s,\tau) = 0$ and is stable if the slope of $\omega_K - \tilde{\omega}_K$ is not positive for a marginal change in s. One solution is clear by inspection where $s = \frac{1}{2}$ such that there is an even distribution of skilled workers and manufacturing firms. Differentiating $\omega_K - \tilde{\omega}_K$ with respect to s, substituting $s = \frac{1}{2}$ and setting $\frac{\partial}{\partial s}(\omega_K - \tilde{\omega}_K)$ to zero finds that $s = \frac{1}{2}$ is a steady state if transport costs are below the threshold:

$$\tau^{\sigma-1} < \phi_b = \frac{1 - \frac{\mu\alpha\vartheta}{\sigma}}{1 + \frac{\mu\alpha\vartheta}{\sigma}} \frac{1 - \frac{1}{\sigma} - \frac{\mu}{\sigma}}{1 - \frac{1}{\sigma} + \frac{\mu}{\sigma}}, \qquad 0 \leq \tau \leq 1. \tag{4.19}$$

This is the "break point" ϕ_b where if transport costs exceed this threshold, the equal distribution is no longer a stable steady state. The equilibrium will revert to the full agglomeration outcome in one region or the other. It can be ensured that the break point is greater than zero by assuming $\mu < \sigma - 1$ and $\mu\alpha\vartheta < \sigma$. This second condition is required to extend the "no black holes" condition for the rare event that $\mu\alpha\vartheta > 1$. Although α is always less than one and ϑ is also typically less than one, there is a theoretical possibility for ϑ to be greater than one. There are also two other interior solutions for $V(s, \tau) = 0$ which are found to be unstable steady states similar to Forslid and Ottaviano (2003) so, these are not investigated further here.

The distribution of economic activity includes the standard location forces present in all core-periphery models (market access effect, cost of living effect and market crowding effect) and a new location force: the technology clustering effect. This idea is expanded on in Chapter 5 and is also termed the clustering effect for consistency and to differentiate it from an agglomeration effect. The "clustering effect" refers to the attraction to regions where firms have similar knowledge inputs to innovation. The home market effect is the overall attraction for the firm (and workers) to choose the larger region. It is made up of the market access effect and cost of living effect, as discussed previously in sections 2.3.1 and 3.3.3, and technology clustering effect. The market access effect is where the firm prefers to locate in the larger market to minimise transport costs and the cost-of-living effect is where workers prefer the location with the lowest cost of living and attempt to minimise transport costs by locating in the region with the most producers. That is, workers prefer to live in the larger market because it has the lower overall price index. At the same time, the market crowding effect means peripheral regions are attractive to producers because of high prices due to transport costs. When transport costs are high, the market crowding effect dominates the home market effect such that a smaller region offers higher skilled wages dispersing economic activity and resulting in an equal distribution of manufacturing and skilled labour. There is an overlap between the core-periphery and equal distribution steady states because the home market effects depends upon the distribution of firms and only occurs with unequal states. The three forces of market access, cost of living and market crowding are pecuniary externalities that determine steady states in the standard core-periphery model. In addition, the model here adds another force for agglomeration to the home market effect, the "clustering effect", where skilled workers prefer to locate alongside other firms due to knowledge spillovers that increase research and development productivity and reduce the cost of innovating.

It is this technology clustering effect and market contestability by innovation that are the key differences to the standard footloose entrepreneur model (Forslid and Ottaviano, 2003) or other product variety models of regional growth (Baldwin et al., 2003). While assumptions are made to maintain symmetry, if there are unequal shares of labour, firms in the larger region have greater access to knowledge. This means that to enter the market, peripheral firms have to choose a technology level that is competitive with home region firms or else lose their market position, yet foreign firms have a lower index of technological opportunity. As a result, foreign firms have lower productivity in producing innovations and require a greater number of skilled workers to achieve the quality target required for market entry. Peripheral firms have lower wages for skilled workers in order avoid negative profits. This lower wage results in further migration to the core region. With additional workers, there are now a greater number of firms in the larger region and even further migration. This clustering effect is caused because the productivity of skilled workers in research and development benefits from greater access to knowledge.

4.3.2 Agglomeration

The other scenario to consider is where all skilled workers and manufacturing firms are already located in a single region. In the typical footloose entrepreneur model, the solution is found in the same way as for an equal distribution, but the model here is richer, because any shift to the periphery requires a firm to switch its location or replace an existing variety, rather than create an expanding variety of goods. While no skilled workers are initially employed in the periphery, the potential skilled wage is not non-existent, but requires a migrant to start a firm in the peripheral region. Switching entrepreneurs must transfer knowledge from the agglomerated location to the periphery such that setting up the first firm in the periphery requires a greater number of skilled workers to achieve a quality improvement large enough to secure entry. This first firm will have the same quality target required for entry in the home region (to ensure entry) and will have a nominal wage such that the firm makes zero profit in equilibrium, but also has all knowledge inputs to innovation located in the alternate (core) region. Therefore, the establishment of manufacturing capability in the periphery is an additional barrier to skilled worker dispersion. This nominal wage is the foreign wage for skilled workers considering migration from the agglomerated region to the periphery.

Consider an example where the home region is fully agglomerated with all skilled workers and all manufacturing firms. A migrating skilled worker would have to start a firm in the foreign region in the coming period $t+1$ by developing a quality target equal to the target of home region firms in order to ensure market entry: $A_{i,t+1} = \frac{\sigma-1}{\eta}\left(A_{i,t} + \lambda_V \overline{A_{w,t}}\right)$. Any quality level below this would mean that an alternative firm could develop a greater quality improvement for that variety and hold the right to manufacture that variety.

Starting from symmetry in quality levels and production (all firms have the same quality level and output), the fixed input requirement of migrant skilled labour to develop a large enough quality improvement is $\gamma e^{(\sigma-1)\frac{n_{t-1}+\tilde{n}_{t-1}+\lambda_V(n_{t-1}+\lambda_R\tilde{n}_{t-1})}{\lambda_R(n_{t-1}+\tilde{n}_{t-1})+\lambda_V(\lambda_R n_{t-1}+\tilde{n}_{t-1})}}$. The function cannot be simplified with $\chi = \frac{n_{t-1}+\tilde{n}_{t-1}+\lambda_V(n_{t-1}+\lambda_R\tilde{n}_{t-1})}{n_{t-1}+\tilde{n}_{t-1}+\lambda_V(\lambda_R n_{t-1}+\tilde{n}_{t-1})}$ due to the addition of λ_R representing the transfer of knowledge about the quality of the variety that is transferred between regions. With full agglomeration in period t, ($\frac{n_t}{n_t+\tilde{n}_t} = 1$ and $\tilde{n}_t = 0$) the $1+\lambda_V$ cancels out from the fraction so the fixed input of migrant skilled labour to develop a large enough quality improvement in the foreign region simplifies to $\gamma e^{\frac{\sigma-1}{\lambda_R}}$. It is assumed that a migrant does not take account of the effect of her own migration on wages which is reasonable for a large number of skilled workers (i.e. each agent making the migration decision is small enough to have no noticeable effect on wages). With labour market clearing, Equation 4.10 is revised to $\tilde{n}_{t+1} = \frac{\tilde{L}_{K,t}}{\gamma e^{\frac{\sigma-1}{\lambda_R}}}$. The rest of the analysis follows as above such that (potential) skilled wages are defined by the revised number of firms and $s=1$ (agglomeration in home):

$$\begin{aligned} w_{K,t} &= \frac{\frac{\frac{\mu\alpha\vartheta}{\sigma}}{1-\frac{\mu\alpha\vartheta}{\sigma}} L_U \frac{e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}}}{L_{K,t}+\tilde{L}_{K,t}}}{\frac{e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}} 2\tau^{\sigma-1}}{\tau^{\sigma-1}\left(e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}}\right)^2}}, \\ \tilde{w}_{K,t} &= \frac{\frac{\frac{\mu\alpha\vartheta}{\sigma}}{1-\frac{\mu\alpha\vartheta}{\sigma}} L_U \frac{1}{L_{K,t}+\tilde{L}_{K,t}}}{\frac{e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}}\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)}{\tau^{\sigma-1}\left(e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}}\right)^2}}. \end{aligned} \tag{4.20}$$

Revising wage pressure given in Equation 4.17 for wages with agglomeration obtains:

$$\begin{aligned} \omega_K - \tilde{\omega}_K &= \frac{w_K}{P} - \frac{\tilde{w}_K}{P} \\ &= \Phi \frac{V(s=1,\tau)}{\tau^{\sigma-1}\left(e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}}\right)^2}, \end{aligned} \tag{4.21}$$

where Φ is as above and $V(s=1,\tau)$ given by:

$$\begin{aligned} V(s=1,\tau) &= e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}}\left(e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}} 2\tau^{\sigma-1}\right) \\ &\quad - \frac{1}{(\tau^{\sigma-1})^{\frac{\mu}{1-\sigma}}}\left(e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}}\left(1-\frac{\mu\alpha\vartheta}{\sigma}+\left(1+\frac{\mu\alpha\vartheta}{\sigma}\right)\tau^{2(\sigma-1)}\right)\right) \end{aligned} \tag{4.22}$$

with Φ evaluated for agglomeration in home, $\Phi = \frac{\mu\alpha\vartheta L_U(\sigma-1)^\mu A_i^\mu}{(\sigma-\mu\alpha\vartheta)(\sigma\beta)^\mu\left(L_{K,t}+\tilde{L}_{K,t}\right)n_t^{\frac{\mu}{1-\sigma}}}$. $V > 0$, represents a stable steady state because the value of the firm choosing the home location is positive and therefore the preferred location for all home firms. Solving $V(s = 1, \tau) > 0$, finds that full agglomeration in home is a stable steady state whenever trade costs are small enough that $\phi = \tau^{\sigma-1}$ is above the threshold ϕ_s, that is implicitly defined by:

$$1 - \frac{\mu\alpha\vartheta}{\sigma} + \left(1 + \frac{\mu\alpha\vartheta}{\sigma}\right)\phi_s^2 - e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}}\phi_s^{\frac{1+\mu}{1-\sigma}} = 0. \tag{4.23}$$

This is the sustain point. As long as transport costs are greater than the threshold ϕ_s, full agglomeration in home is a steady state. Isomorphic analysis applies to agglomeration in the foreign region, resulting in the same sustain point ϕ_s. Note the break point is unaffected by regional knowledge spillovers, but the sustain point is affected such that agglomeration is a steady state at a lower freeness of trade (higher transport cost) than in the footloose entrepreneur model, because the model includes the technology clustering effect in addition to the standard location forces in the core-periphery model. That is, workers are not as mobile between regions when they have to take existing knowledge with them to develop innovations in a new location. The sustain point increases (decreasing transport costs) with λ_R, because a higher λ_R means technology and entrepreneurs are more mobile.

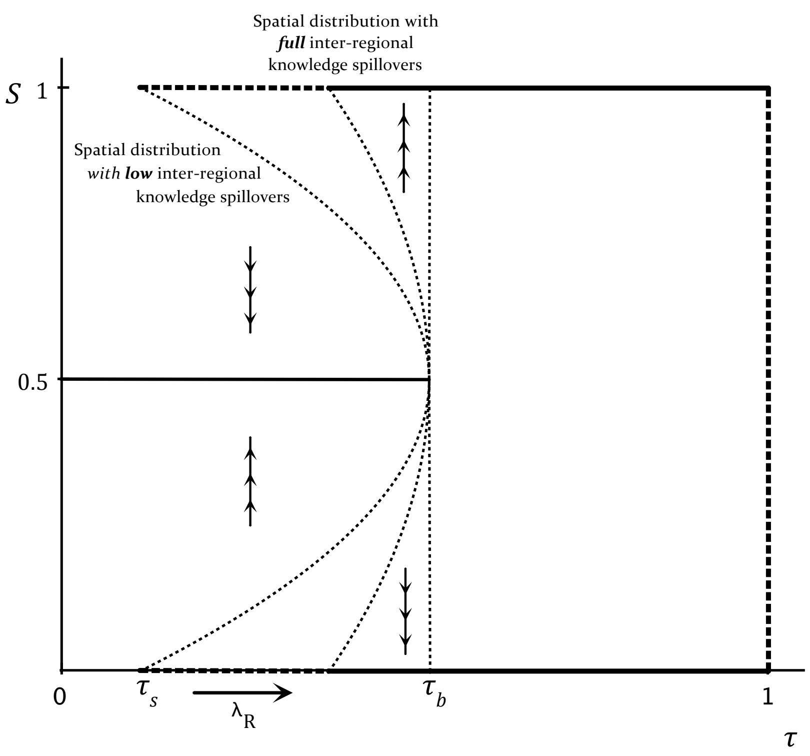

Figure 4.1: Tomahawk diagram with inter-regional knowledge spillovers

The Tomahawk diagram given in Figure 4.1 describes the stable steady state shares of skilled labour ($\frac{L_{K,t-1}}{L_{K,t-1}+\tilde{L}_{K,t-1}}$) for the range of trade freeness and inter-regional knowledge spillovers. The two possible stable steady states, manufacturing concentration or equal distribution, are described by the solid lines and the arrows describe the direction of movement by firms and skilled workers towards a steady state (if starting in a state that is not steady). The dotted lines describe unstable internal steady states which are still steady, but unstable because a small change would trigger a switch to either full agglomeration or equal distribution (according to the direction of the arrows). The dashed lines describe equilibria that are only stable under specific conditions such as low knowledge spillovers between regions that are geographically separate or zero transport costs ($\tau = 1$), respectively.

Starting from an equal distribution of skilled labour with low trade freeness ($\tau = 0$), as freeness of trade increases the break point τ_b, defined by $\tau_b = \phi_b^{1-\sigma}$, is the level of trade freeness where the home market effect exactly equals the market crowding effect. There is no cost of living or clustering effects as these are dependent on an unequal distribution. With freeness of trade below the break point, a marginal change in the labour distribution reverts back to the equal distribution outcome. As freeness of trade increases past the break point, a marginal change in the labour distribution would trigger catastrophic agglomeration in one region. Alternatively, starting with free trade ($\tau = 1$) and the manufacturing concentration outcome, as freeness of trade decreases below the break point the equal distribution outcome is now stable, but only if firms and workers migrate between locations. A marginal change in the distribution of labour will still revert to the agglomerated state, but a large shock could induce a switch to the equal distribution outcome. Both types of equilibria are possible for some range of trade freeness. As freeness of trade declines further, the sustain point is reached, defined by $\tau_s = \phi_s^{1-\sigma}$, where the market crowding effect exactly equals the technology clustering, home market and cost of living effects. If freeness of trade declines further, it would trigger a sudden dispersion of industry between the two regions. The range of the overlap depends upon the level of interregional knowledge spillovers λ_R and other model parameters that determine the size of these forces for manufacturing concentration. An increase in λ_R results in a higher sustain point τ_s.

4.3.3 Stability analysis

Most two-region models of growth and location implicitly or explicitly consider the impact of market integration on economic activity. The model in this chapter provides a unique opportunity to consider another dimension of market integration: the effect of the ability to transfer technology between locations as a prerequisite to vertical innovation and production. Similarly, Baldwin and Forslid (2000) consider the effect of spatial limitations on knowledge spillovers in a product variety model, but their analysis is unable to capture the characteristics of economic activity where the market is contestable through technology improvement. This chapter expands their conclusion regarding the stabilising nature of knowledge spillovers suggesting that increasing the ability for technology to transfer between locations is destabilising for the core-periphery outcome, because it increases the mobility for entrepreneurs to develop quality improvements in other locations. However, changes in the level of knowledge spillovers have no effect on the stability of the equal distribution outcome because the break point is unaffected by regional knowledge spillovers. Even so, it makes it more likely that skilled workers can remain in the equal distribution outcome because it is only destabilising for the core-periphery steady state and makes full agglomeration less likely.

To examine stability, consider the range of transport costs where each outcome is stable or unstable. The boundaries of this range are defined by the break and sustain points. The break point as given by Equation 4.19 is constant for all levels of λ_R. Understanding the relationship between the sustain point and regional knowledge spillovers is more complex. Figure 4.2 shows the break and sustain points evaluated for different levels of regional knowledge spillovers λ_R. As regional knowledge spillovers increase, the sustain point increases in freeness of trade, as agglomerated firms are more mobile and more inclined to shift to the periphery. Inspecting the sustain pont defined implicitly by Equation 4.23 and noting that $0 \leq \lambda_R \leq 1$, it can be seen that the component $e^{(\sigma-1)\frac{1-\lambda_R}{\lambda_R}} \phi_s^{\frac{1+\mu}{1-\sigma}}$ decreases with λ_R to a specific limit when λ_R is one. As a result, the sustain point increases with λ_R. The range of transport costs where both full agglomeration (core-periphery) and equal distribution are stable steady states decreases with regional knowledge spillovers. Therefore, regional knowledge spillovers are destabilising, because the market becomes more contestable as regional knowledge spillovers increase.

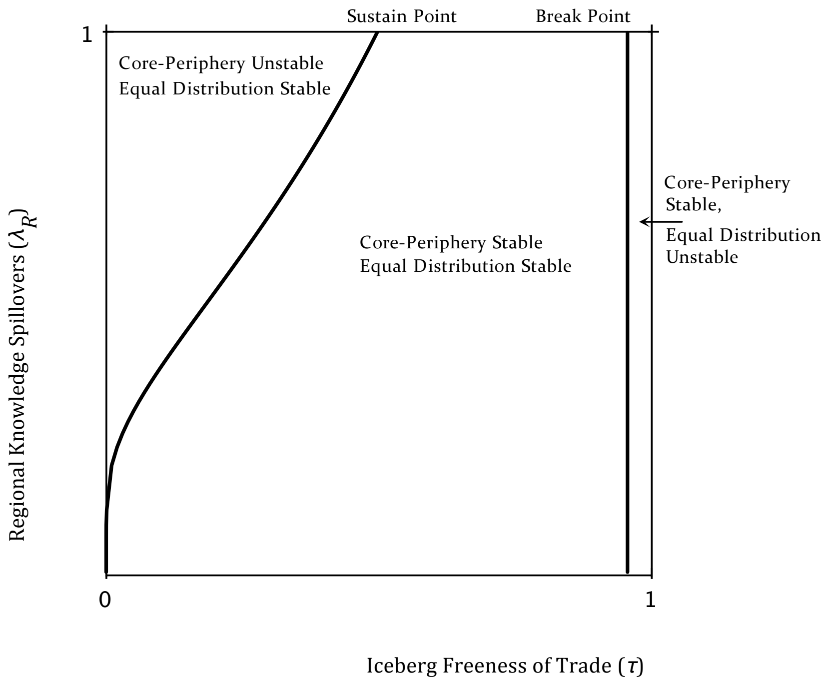

Figure 4.2: Stability chart of core-periphery and equal distribution outcomes

When a large agglomerated economy integrates with a small regional or peripheral economy and thereby increases the ability to transfer knowledge between locations, it is easier for firms in the agglomeration to relocate to the periphery. At some levels of transport costs and knowledge spillovers, this could be catastrophic such that there is a sudden dispersion of industry between the two integrating regions. This destabilising nature of knowledge spillovers for agglomerated regions can be observed in Figure 4.2 by the range of freeness of trade at different levels of λ_R. With increases in λ_R, the range of freeness of trade decreases for which the core-periphery outcome is a stable steady state. Alternatively, regional knowledge spillovers could be considered stabilising for peripheral regions or equally distributed outcomes, because regional knowledge increases the range of transport costs where only the equal distribution outcome is stable allowing the region to retain manufacturing capability. This phenomenon can be observed in the figure. With increases in λ_R, the range of freeness of trade increases for which only the equal distribution outcome is a stable steady state.

This is a slightly different conclusion for agglomerated locations compared to existing two-region models of growth in product variety, although the underlying intuition is similar. Comparing Figure 4.2 with Figure 3.1 reproduced from Baldwin and Forslid (2000) shows the direct comparison. With increases in λ in Baldwin and Forslid's (2000) model, the range of trade freeness increases for which both outcomes are a stable steady state, leading to the conclusion that regional knowledge spillovers are generally stabilising. Their model also finds that with increases in λ, the range of trade-freeness decreases for which the core-periphery outcome is a stable steady state, so λ is also destabilising for the core-periphery outcome. However, the impact is not the same. The result in their model would be converging growth rates in manufacturing between regions as entrepreneurs in the periphery are able to start taking advantage of knowledge transfers from the core region. Alternatively the model in this chapter would suggest a hollowing out of manufacturing from the core region as entrepreneurs disperse into the periphery.

Regional knowledge spillovers are equity-enhancing, because they reduce the localisation of knowledge and encourage the dispersion of manufacturing between locations. If the market is contestable by quality improvements, this dispersion only impacts the agglomerated outcome (by shifting the sustain point). The model here suggests that two integrated regions have quite different incentives for market integration. It is in the interests of peripheral regions to increase knowledge transfers such that currently agglomerated firms are more able to set up in a peripheral location. For an agglomerated location, there is a risk that the dispersion of innovation will result in a manufacturing decline. Instead agglomerated regions prefer a policy approach that reduces transport costs, as this allows both regions to be better served by the agglomeration. Consumers in peripheral regions also benefit from a decline in transport costs as it reduces the cost of imported manufactured products. However, if starting in an equal distribution outcome, it could trigger agglomeration in one region. The implications for integration policy of the model developed in this chapter are that peripheral countries should support initiatives which increase the ability for knowledge to transfer between locations, but they may face opposition from core regions and countries with large agglomerated cities.

4.4 Growth

So far this chapter has considered equilibrium prices, quantity and investment, the steady state distribution of economic activity and forces for agglomeration. It is possible that agglomeration facilitates greater quality improvements that could compensate the periphery for its loss in real incomes due to a switch from the equal distribution outcome to the core-periphery outcome. The desire for policy to promote equitable outcomes can be balanced with the additional benefits of greater growth through agglomeration. This section explores the relationship between growth and agglomeration (whether steady states or not) and between growth and knowledge spillovers.

Equations 4.9a and 4.9b, describe each firm's quality target. The target depends upon the elasticity of substitution, the parameters of the innovation production function and the level of agglomeration. The elasticity of substitution determines a firm's mark up and therefore its willingness to pay for research. The parameters of the innovation production function and the level of agglomeration determine the productivity of skilled labour in research and the threshold quality level for market entry. In particular, λ_V and λ_R describe the ability for knowledge to transfer between firms. Since growth is fuelled by innovation, it is possible to consider the relationship between these different knowledge externalities and economic growth.

The growth rate of quality per period is $g_A = \frac{A_{i,t+1}-A_{i,t}}{A_{i,t}}$, substituting Equation 4.9b (advanced one period) yields the growth rate of quality:

$$g_A = \frac{\sigma - 1}{\eta}\left[1 + \frac{\lambda_V\left(n + \lambda_R \tilde{n}\right)}{n + \tilde{n}}\right] - 1. \tag{4.24}$$

As capital enters the model in the form of the quality level, g_A is the growth rate of the capital stock, even though the user or owner of that knowledge capital may change through innovation and contestability. Per capita growth in real income (or consumption) is the rate at which utility improves between periods: $g_Q = \frac{\Delta Q}{\Delta t} = \frac{Q_{t+1}-Q_t}{Q_t}$. The growth rate of capital is directly linked to real income growth. While wages, expenditure and prices keep the same nominal value, the price index changes over time such that the ability to purchase (utility) increases as higher quality products replace obsolete varieties. The increase in purchasing power or the decline in the perfect price index is the economy's growth rate.

Using the price quality index for manufactured varieties finds that the index is falling at a rate of $g_A^{\sigma-1}\frac{1}{1-\sigma}$. The overall price index is falling at a rate of $\frac{\mu}{1-\sigma}$ times $g_A^{\sigma-1}$. The growth rate of utility is therefore given by:

$$g_Q = \frac{\mu}{\sigma - 1}\left[\frac{\sigma - 1}{\eta}\left[1 + \frac{\lambda_V\left(n + \lambda_R \tilde{n}\right)}{n + \tilde{n}}\right] - 1\right]^{\sigma - 1}. \tag{4.25}$$

The growth rate is the same in both locations in the steady state, whether it is an equal distribution outcome or manufacturing concentration. This is because, while the price index for manufactured varieties falls, the equivalent for traditional goods remains constant, since consumers still spend a constant portion of their earnings on traditional goods. The overall price index for consumption in either region decreases at the same rates, because producers of traditional goods can trade for manufactures.

4.4.1 Knowledge spillovers and growth

Inter-varietal knowledge spillovers in the form described here do not crowd out a firm's own investment in innovation, so they are essentially a free portion of additional growth. Varietal knowledge spillovers (λ_V) describe how knowledge can be applied across many sectors, beyond the variety where the knowledge was originally developed. Therefore, λ_V describes the level of underinvestment in knowledge creation, as entrepreneurs fail to receive the full benefit created by the development of knowledge. This is part of the market failure of innovation. Additional market failure of innovation is due to the inter-temporal nature of knowledge spillovers. Firms invest in innovation creating new knowledge, but only receive the benefits from that knowledge in the coming period. As a result, firms underinvest in innovative activities overall. These market failures of innovation suggest policies that support innovation activity are efficiency and growth-enhancing.

Inter-varietal spillovers have a greater impact on economic growth if not also diminished by spatial effects. Manufacturing concentration, or increasing the ability for knowledge to transfer between locations that are geographically separated (increasing λ_R), has the effect of reducing the size of the market failure of innovation. Figure 4.3 describes how growth is affected by varietal knowledge spillovers. If $\lambda_R < 1$ and there is manufacturing concentration, growth is represented by the line between growth of $\frac{\mu}{\sigma-1}\left(\frac{\sigma-1}{\eta}-1\right)^{\sigma-1}$, if $\lambda_V = 0$ and growth of $\frac{\mu}{\sigma-1}\left(\frac{2(\sigma-1)}{\eta}-1\right)^{\sigma-1}$, if $\lambda_V = 1$. Alternatively, if there is equal manufacturing in both locations, growth ranges between $\frac{\mu}{\sigma-1}\left(\frac{\sigma-1}{\eta}-1\right)^{\sigma-1}$, if $\lambda_V = 0$ and $\frac{\mu}{\sigma-1}\left(\frac{\sigma-1}{\eta}\left(1+\frac{1+\lambda_R}{2}\right)-1\right)^{\sigma-1}$, if $\lambda_V = 1$. Inter-varietal spillovers have a greater impact on agglomerated locations, because firms in these regions do not face spatial externalities in knowledge transfer between firms that are geographically separated. There are implications for prioritising different types of policy for agglomerated and peripheral locations. Small agglomerated regions should focus on both a clustering policy to attract technologically-related industries and improving regional knowledge spillovers while large agglomerations could focus on firm interaction within the region. Both of these increase domestic varietal knowledge spillovers.

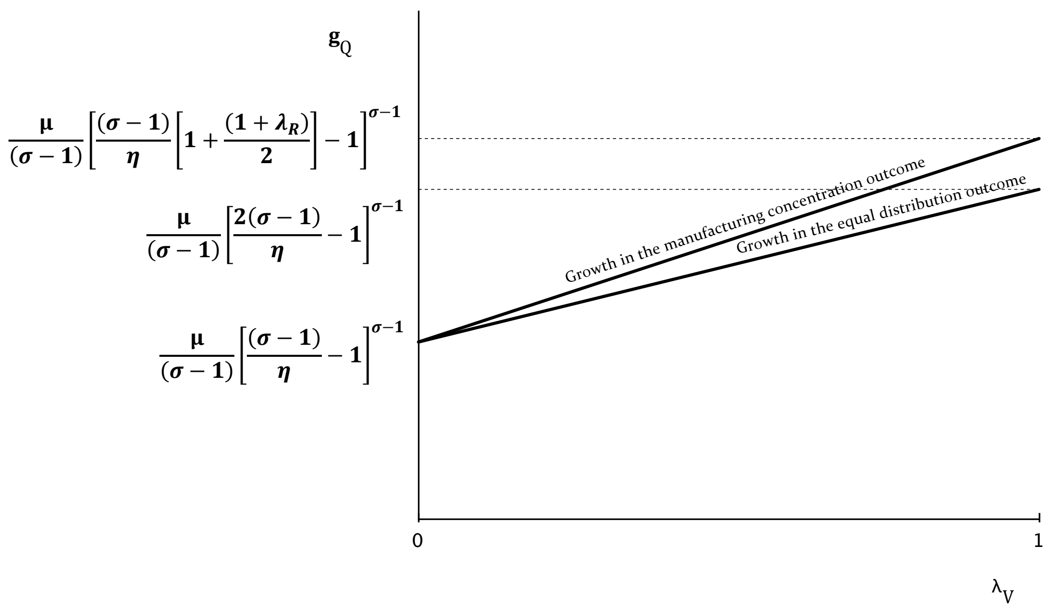

Figure 4.3: Economic growth with varying inter-varietal knowledge spillovers

Greater inter-varietal knowledge spillovers also increase the size of the market failure of innovation, because they increase the benefits of knowledge creation without increasing firms' investment in knowledge generation. While the model here assumed λ_V is constant between all sectors, in the real world knowledge is more complex and is not applied equally across varieties. If a single region is more concentrated with sectors where knowledge is more compatible, this is similar to increasing the size of λ_V. Policies which increase industry concentration are positive for growth because of the effect on λ_V. They have an additional effect of increasing the size of the market failure (by increasing the benefits of knowledge generation without increasing the investment), because a greater λ_V means knowledge is useful to growth in more varieties. Therefore, these types of policies are more effective if complemented by policies that also increase innovation activity because of the increase in the size of the market failure.

Inter-regional spillovers describe how knowledge does not fully transfer between locations that are geographically separated. Under the core-periphery outcome they have no effect on growth, because all manufacturing and research and development are in the same location. Alternatively, under the equal distribution outcome, partial inter-regional knowledge spillovers reduce growth. Figure 4.4 shows how growth varies with different levels of inter-regional knowledge spillovers. Under the equal distribution outcome, increasing inter-regional knowledge spillovers has a positive effect on growth, because it increases the productivity of growth in both locations.

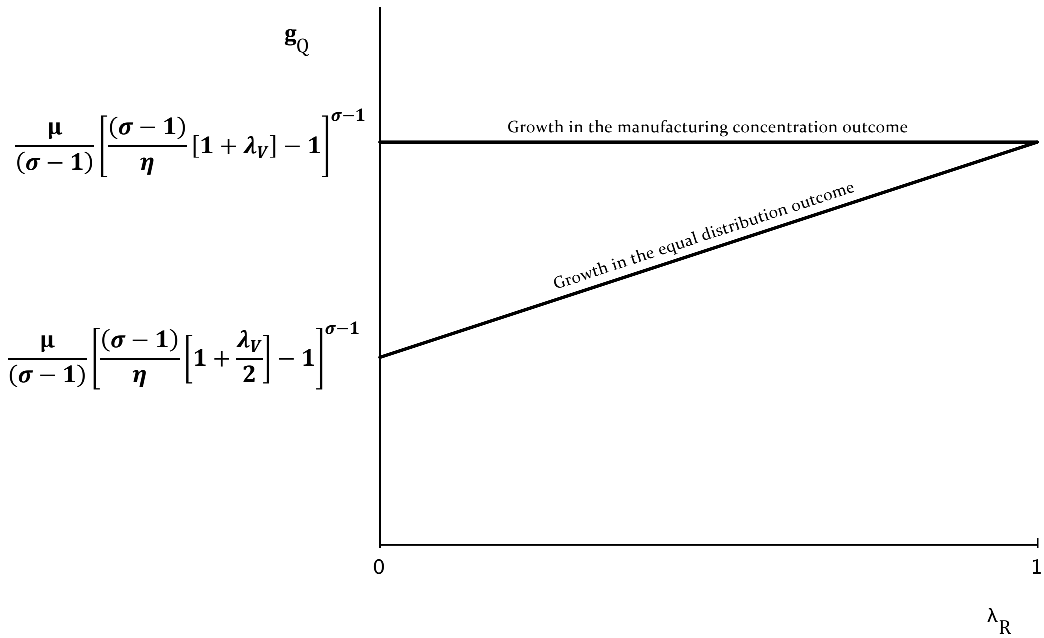

Figure 4.4: Economic growth with varying inter-regional knowledge spillovers

4.4.2 Agglomeration and growth

While the model here was chosen because there are no scale effects, n_t and $\tilde{n}_t$ are affected by the distribution of L_K and the productivity of skilled workers, so growth in quality in each location is affected by migration. Changes in the size of a region lead to transitional effects in growth rates. If the distribution of firms and workers is between stable steady states such that there is migration, innovation activity in the smaller region has lower productivity due to lower knowledge spillovers. As such, the number of products available overall is lower if there is partial agglomeration than in the full agglomeration or equal distribution outcomes, because firms in the smaller region require greater skilled labour input per variety (and lower wages per worker) in order to enter the market. The growth rate of quality of peripheral products is still the same, as this is the quality required for market entry. However, in the following period skilled workers migrate to the agglomerated location, skilled worker productivity increases and the number of varieties expands. Growth has a long-run component of increasing quality, and a medium-run transitional component which is influenced by changes in the distribution of labour. While this chapter does not investigate this transitional component further because it is focused on long run growth, these transitional effects could partially explain the varying growth rates in integrated markets. For example, an explanation for New Zealand's relatively low growth rate could be this transitional component of growth as industry switches to core locations in Sydney or Melbourne to service the wider region. The transitional component of growth comes from new migrants' increased productivity in their new location (and non-migrants' decreased productivity in the periphery) and the transfer of industry between locations toward the endogenously determined stable steady state level.

With an equal distribution of manufacturing between locations, quality improves by a factor of $\frac{\sigma-1}{\eta}\left(1+\frac{\lambda_V(1+\lambda_R)}{2}\right)$. Conversely with the core-periphery outcome, the agglomerated location sees quality improve by a factor of $\frac{\sigma-1}{\eta}(1+\lambda_V)$. With $\lambda_R < 1$, the quality improvement under a core-periphery outcome is always higher than under the equal distribution outcome (or any other interior distribution). Since all firms are local, knowledge spillovers do not diminish due to distance, but due to varietal factors only.

Figure 4.5 shows the levels of growth with different levels of agglomeration. Agglomeration in the home region is shown by its share of skilled workers.

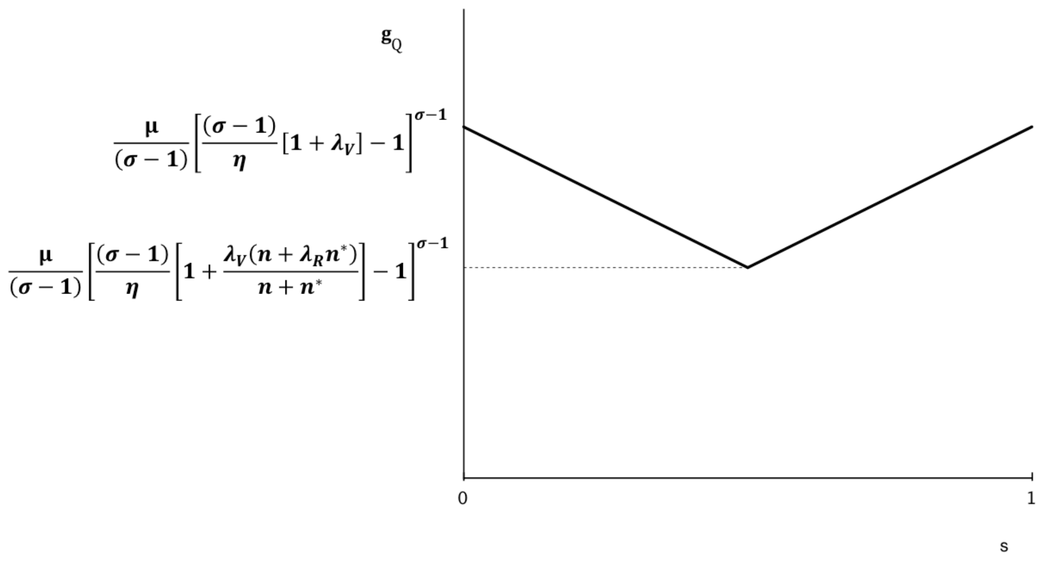

Figure 4.5: Economic growth with varying levels of agglomeration

Growth is highest if all manufacturers are in one region. Growth is at its minimum when each region has half the manufacturers. The reason the periphery experiences growth, even with no manufacturers (and no innovation), is because the periphery can trade traditional goods for manufactured varieties. While the real income level of the periphery is lower than in the agglomeration, because all manufactured varieties must be imported, the growth level in the periphery is the same as in the agglomerated location.

4.4.3 Policy implications for supporting economic growth

By including technology in the firm's location, investment and production decisions, the model here helps explain the emergence of technologically advanced locations. research and development subsidies in a single location could initiate and speed up the agglomeration process by increasing the wage differential. Alternatively, they could slow down the agglomeration process if a subsidy in the periphery decreases the wage differential. Policies which attempt to retain manufacturing in the periphery may have a detrimental effect on long-run growth, but are always equity enhancing.

research and development subsidies were dismissed by Young (1998) as ineffective for improving growth rates if there are no scale effects. Subsidies only change the pool of rents, shared by skilled labour. However, differing subsidies in each location could influence the location of economic activity. Furthermore, if subsidies prevent the agglomeration of industry, both locations miss out on the quality improvements due to knowledge spillovers that could have occurred within an agglomerated location. Subsidies in the periphery may have a short-term effect of retaining some industry and maintaining higher real peripheral wages, but are ultimately detrimental to global economic growth and income. However, peripheral innovation subsidies are equity-enhancing and may achieve other desirable outcomes. Despite this, equity could perhaps be better facilitated by other policy initiatives such as enhancing the ability for knowledge to transfer between firms that are geographically separated. Initiatives such as regional universities, research institutions, incubators for start-ups, national communications and transport infrastructure that all aid the ability for people and knowledge to move between locations, are both equity and growth-enhancing.

The quality ladders or creative destruction approach has some interesting implications, because it particularly highlights the effect of historical "accidents". That is, events which have affected growth for a period in the past have permanent effects on income, technology and the distribution of economic activity. Similarly, sudden changes to market integration of two regions could have catastrophic and irreversible effects on income levels. These effects are known as hysteresis.

By taking a related-variety approach to knowledge spillovers, it also means there are different policy implications for core and peripheral regions typically ignored by the growth or trade literature. Regional knowledge spillovers are a greater proportion of quality improvement in a location with little manufacturing than in a location with many manufacturers. To encourage quality improvements, the low manufacturing location may be better off supporting regional knowledge spillovers than additional support for research and development. An agglomerated location on the other hand may support growth by encouraging relationships between the location's firms which support spillovers between different varieties.

By supporting inter-regional knowledge spillovers a location will have greater quality improvements through knowledge spillovers. This could lead to a dispersion of industry as firms are indifferent between locations to conduct research and development. If regional knowledge spillovers increase, the equal distribution outcome becomes more stable, but the core-periphery outcome is less stable. This conclusion is in contrast to Baldwin and Forslid (2000) who show that integration policies which lower the cost of trading information are always stabilising. The model here is similar, but is able to differentiate between the sources of knowledge spillovers. In addition, there is a hysteresis effect where firms typically prefer to continue to conduct research and manufacturing in the same location, a result which cannot be shown in a product variety model.

4.5 Conclusion

This chapter brings the quality ladders approach of endogenous growth theory to the two-region setting and provides a framework for dealing with different sources of knowledge spillovers. It makes it possible to think about the localisation effects of related-variety spillovers within a quality-improving innovation growth model, alongside inter-temporal and spatial effects on knowledge spillovers. The results highlight different growth and integration policy approaches and preferences for agglomerated and peripheral locations. In particular, peripheral locations should focus on policies which increase the ability for knowledge to transfer between locations that are geographically separated, while agglomerated regions will be cautious of this integration approach, because there is a risk of manufacturing dispersion.

In the core-periphery outcome, growth is the same in manufacturing and traditional goods locations, even though one region is technologically advanced, because traditional goods locations benefit from trading for innovative products that achieve higher quality improvements in the agglomerated region. Real earnings in traditional goods locations may be lower, but they grow at the same rate as in the agglomeration. Varying growth rates between different locations are perhaps a symptom of the transition from separate economies to highly integrated markets. In the longer run, growth rates in both locations would be expected to equalise. Growth rate differences are both a reflection of lower productivity innovation in a less agglomerated location and the transition to long-run income levels. Closer integration for trading goods can cause a shift of industry from one location to another, but in the long run, growth is the same in both locations.

From a policy perspective, a number of interesting implications arise. The results of Young (1998) that research and development subsidies have no effect on innovation effort are extended by noting that the wage effect of unequal subsidies between different locations can speed up or slow (or even prevent) agglomeration and therefore has implications for growth. research and development subsidies in a low manufacturing location may sustain technology growth in the region in the short run but reduce long-run growth rates by slowing the transition to the agglomerated outcome, because it reduces, or even reverses, the real wage differential which drives migration. On the other hand, subsidies in a high manufacturing location speed up the transition to agglomeration. For an internal equilibrium, a low or equal manufacturing location attempting to retain manufacturing capability may be better off focusing on policies which encourage inter-regional knowledge spillovers rather than policies for research and development. Such policies are stabilising and retain manufacturing when markets integrate. A location with manufacturing concentration depends less on inter-regional linkages and more on local spillovers for economic growth. Agglomerated locations will instead be supportive of integration policies which reduce transport costs such that the agglomeration can supply a larger market.

As in Baldwin and Forslid (2000), regional knowledge spillovers are a stabilising force, but only for the equal distribution outcome. The break point is unaffected by these spillovers so the range of transport costs where both equilibria are stable is decreasing with knowledge spillovers. Regional knowledge spillovers are destabilising for the core-periphery outcome. In particular, if the market is contestable through quality improving innovations as in the model here, the destabilising nature of knowledge spillovers could trigger a sudden dispersion of economic activity between regions. This highlights the importance of hysteresis in integration and the effects on technology levels. Sudden and incomplete integration through either transport costs or the ability to transfer knowledge between locations could result in drastic and irreversible changes in the location of economic activity, with catastrophic consequences for incomes and production technology levels in some regions. The conflicting incentives for integrating regions should be an important factor in economic integration approaches.

Chapter 5

A multi-sector model of growth with innovation clustering

This chapter develops a multi-sector model with endogenous growth through quality improving innovations. Each firm's technology sector and the location of other firms play a role in each firm's ability to improve its own technology. As a result, firms prefer to co-locate in technologically compatible clusters. Technology growth is highest at clustered and agglomerated locations. Production of each variety is contestable through vertical innovation and firms choose a sector, variety and location to maximise profit by maximising the productivity of research and development. The direction and magnitude of the clustering effect is specific to each sector depending upon the spatial distribution and knowledge intensity of all firms and sectors. Firms choose an optimal location for research and development by balancing the clustering effect with the manufacturing agglomeration effect. Clusters of firms in the same sector can develop in peripheral locations if the innovation function captures a sufficiently large share of knowledge spillovers from firms in their own sector. The model suggests that governments in peripheral regions or countries could play an important role to sustain peripheral clusters by monitoring innovation performance and assisting the research and development sector.

5.1 Introduction

Endogenous growth theory recognises the role of knowledge spillovers in research and development decisions and implies that firm clustering has an important role in innovation and growth processes. Mechanisms which lead to industrial clustering (and concentration of firms and people) have been considered since Marshall (1890). These include knowledge spillovers, transport costs and factor prices. While the literature has long studied factors that encourage clustering, modern endogenous growth theories fail to incorporate the proximity characteristics of knowledge and innovation that lead to industrial clustering - a factor that is particularly important for peripheral regions. Growth models fail to explain why some industries remain in distant or isolated locations, while others shift to large agglomerations.

Knowledge spillovers between firms vary with the concentration of people, production and related technologies. Clustering of technologically related firms has been established as a mechanism to facilitate technology spillovers. Endogenous growth theory should consider how clustering forces interact with innovation and growth. Agglomeration and knowledge spillovers reinforce each other with circular causality in regional models of growth, yet these models typically ignore the formation or impact of clustering. Clustering allows technologically related firms to remain in peripheral locations to benefit from the knowledge spillovers within the cluster that are required as an input to innovation.

Models of growth and agglomeration describe the formation and continued expansion of large and diverse cities, but the concept of specialised cities is less understood. Clusters maintain technological advantages over other locations and further encourage the co-location of technologically related firms. What policy choices (if any) are available for (relatively) declining or lagging regions to keep up with modern globally connected agglomerations? The model considered has multiple industrial sectors where the cost of innovation is dependent on the location and technology of all firms and the selected level of technological improvement. It can be profitable for firms to cluster alongside technologically related firms in a small region even though innovation is more expensive than if all firms and sectors were agglomerated in the same location. While the entire sector could be better off with agglomeration, no single firm can be made better off by relocating. Moreover, this dispersion of economic activity to specialised peripheral cities is also equity-enhancing.

A range of models combine spatial factors with endogenous growth. These models attempt to explain variations in economic growth between regions and nations through both pecuniary and knowledge spillover externalities. Several models combine increasing product variety (Romer, 1990; Grossman and Helpman, 1991a) with the NEG (Baldwin et al., 2001; Baldwin and Forslid, 2000; Martin and Ottaviano, 1999; Martin, 1999; Fujita and Thisse, 2003). This combination is facilitated by the fundamental common use of monopolistic competition as formalised by Dixit and Stiglitz (1977). The different models have varying assumptions on the mobility of capital and labour and assumptions regarding intermediate versus consumer demand to influence the forward and backward linkages. However, full agglomeration is inevitable at low transport costs in these models with mobile capital and/or labour. Furthermore, agglomeration is often referred to as catastrophic, because it can be triggered by a small decrease in transport costs (or a shock) and it is irreversible. While NEG and growth models typically feature agglomeration as a characteristic of equilibrium, industry-specific clustering is a rare feature. In addition, spatially varying technology levels should be recognised as a key factor in location decisions related to technology spillovers. Since growth in conventional NEG and growth models comes from an increasing variety of products, this range of models is unfit to consider the role of technology levels in research and development spillovers.

Most trade models highlight comparative advantage, but in these models agglomeration is caused by vertical linkages between intermediate goods used in production and the pecuniary externality of transport costs. For example, Venables (1999) creates a model of international specialisation, extending Krugman (1991b) by assuming industries both use and produce intermediate goods. In his model, clustering is due to pecuniary externalities in manufacturers' consumption of intermediate goods rather than technical externalities in knowledge transfer as found in the innovation economics literature. Yamamoto (2003) has partial spatial knowledge transfer externalities in the research and development sector and pecuniary externalities in research and development consumption of intermediate goods, but with only one manufacturing sector this model must assume labour is immobile to avoid catastrophic agglomeration.

Products are today increasingly manufactured in global value chains using a multitude of components produced in many countries. With low transport costs and many suppliers and locations, pecuniary externalities from transport costs are becoming less relevant to firm location choices and trade. Despite the apparent ability to operate anywhere, the world economy is not flat; it is increasingly uneven (Christopherson et al., 2008; McCann, 2008). The innovation input to final goods production is dependent on knowledge inputs and not all knowledge is directly transferable between locations. This increases the premium associated with face-to-face contact (Gaspar and Glaeser, 1998; Storper and Venables, 2004; Rodríguez-Pose and Crescenzi, 2008) because of technical externalities in research and development on the inputs side of the production process in knowledge-intensive industries. Cities are becoming more important to production due to the nature of knowledge and technology. Therefore, it is insightful to consider a trade and growth model in which knowledge transfer and innovation are at the centre of explaining specialised innovation clusters and agglomerations. This chapter develops and analyses a model in which knowledge transfer externalities are the key cause of clustering and agglomeration.

Evolutionary economics describes firm innovations, economic development, institutional change, and region specialisation as a branching process (Frenken and Boschma, 2007; Nelson and Winter, 1982; Freeman and Soete, 1997). This conceptual framework is similar to the results implied by the model developed here. A firm's ability to develop new innovations is related to both its technological and spatial proximity (Boschma, 2005). This type of mechanism gives technologically related firms an incentive to locate in clusters. Surprisingly, as far as the author is aware, this feature has not previously been incorporated into endogenous growth models.

Duranton (2007) describes a model of specialised locations by applying the Grossman and Helpman (1991b) quality ladders model to an urban framework to study industry churn. The model uses strictly local knowledge spillovers from the quality leader within each industry and industry churn is generated by a stochastic process in research and development that ignores any interdependence between research firms. As a result, cities are specialised clusters of industries, but stochastic innovations in other industries lead to the shifting of industries between cities. Alternatively, Brezis and Krugman (1997) use an urban model with learning by doing to explain technology lock-in within clusters such that the stochastic emergence of new technologies at other locations cannot be adopted by locked-in incumbents. This leads to the emergence of new clusters and the constant stochastic churning of industries from agglomerations to new clustered locations. These models are interesting descriptions of the dynamics of stochastic industrial clustering and churning. However, they are inappropriate for studying the relationship between clustering and growth, because there are no endogenous forces for industry location. Industry location and churn are instead determined stochastically and maintained in the medium term by hysteresis and strong assumptions about the inability for knowledge to transfer between locations. In reality, research firms across all industries use some shared knowledge and technology to develop innovations. Rather than being stochastic, the locations of industries are likely to be dependent upon the proximity of related technologies. This has implications for economic growth and development policy. Consequently, these technology spillovers are essential to modelling endogenous growth with clusters.

What is lacking in these models is any relationship between research and development and the presence of technologically related sectors. Research papers that develop growth models often claim that the shifting of industry is by the accidental emergence of new technologies. This chapter develops a model where the location of clusters is not accidental. Some level of technology improvement occurs because of related technology spillovers. In the proposed model, clustering is a characteristic of the steady state and the type of sectors that can remain clustered in the periphery is endogenous due to the characteristics of each sector and the spatial distribution of other firms and sectors. Endogenous growth theory assigns importance to technology spillovers, but existing models fail to consider the role of proximity in facilitating technology transfer and its effects on economic growth policy. The model developed here addresses and overcomes this weakness.

With modern communication technologies, knowledge spillovers are not restricted to the local city area. The creative destruction/quality ladders approach (Grossman and Helpman, 1991b; Aghion and Howitt, 1992; Young, 1998) offers an opportunity to develop a model in which firms make use of knowledge from many sources and locations across both geographic and technological space. In these models, growth comes from increasing product quality instead of product variety. The cost of improving quality depends on the availability of knowledge. This chapter develops a model with interdependencies between (research) firms due to their common knowledge inputs, even across different locations. As a result, an endogenous growth model is developed that includes the possibility of modelling clusters of technologically related firms. Such clusters are able to survive sustainably in the periphery, even with factor mobility - a feature that could not be modelled in previous models of growth and location with mobile capital or labour. Consequently, the model offers interesting implications for growth policy in peripheral countries or regions.

As in Chapter 4, the starting point is Young (1998), which uses a deterministic quality ladders model without scale effects, where intertemporal knowledge spillovers determine the cost of innovation. Chapter 4, extends this model to recognise the spatial and related-technology characteristics of knowledge spillovers and innovation. With only one manufacturing sector, full agglomeration is inevitable in this model at low transport costs. The model developed in this chapter includes multiple manufacturing sectors and develops the knowledge spillovers mechanism further by following a proximity (Boschma, 2005) approach. This facilitates the innovation clustering of firms with related technology. Importantly, the model allows for the possibility of firms within each industry to cluster in a location other than a large agglomeration. For simplicity, transport costs are removed to focus solely on the clustering and agglomeration effects of technical externalities in research and development. The importance of transport costs to firm location is already well-documented in the NEG literature and is not the purpose of this chapter. While there are many factors which firms balance when deciding on a production location, this chapter focuses on the trade-offs for firm location decisions based on the location of other firms and spatial technical externalities in research and development. Even with zero transport costs, agglomeration effects in this set-up may be overcome by clustering effects in peripheral locations as firms prefer to co-locate alongside technologically related firms.

Section 5.2 describes the model and the short-run equilibrium prices, wages and investment in research and development. Section 5.3 discusses the forces for clustering, the allocation of industry in the steady state and illustrates quality improvement in the steady states through simulations. Section 5.4 provides a further discussion of the model, including the potential policy implications. Lastly, Section 5.5 concludes.

5.2 The model

A core feature of the model is that firms invest in research and development in order to produce quality improvements and manufacture products for consumers at multiple locations. Technological knowledge related to all varieties is an input to research and development, but this knowledge does not transfer perfectly between locations or varieties. The usability of knowledge for innovation is associated with the spatial and technological proximity of related varieties. There is a greater ability for firms to use knowledge of varieties produced in the same location or within the same technological sector. Firms which choose to locate alongside other manufacturers with technologically related varieties therefore have greater access to technological knowledge and a lower cost of quality improvement. A firm chooses its location by considering knowledge spillovers, the location of other firms and their technological relatedness. This last factor is the mechanism responsible for industrial clustering.

Consider an economy with two production sectors, traditional and manufactured goods. Each has its own factor of production: unskilled and skilled labour respectively. In the manufacturing sector, each firm uses skilled labour and existing technological knowledge in research and development to develop a quality improvement for a single variety from one of many industrial subsectors. If a firm develops a substantive quality improvement, it is awarded with a patent, giving the firm the monopoly right to produce this variety at that quality level. The patent effectively expires when a successful innovator manages to develop the following quality step. If a firm is able to achieve the greatest quality improvement of variety i, they are able to "take the market", producing that variety in the following period. Hence, "taking the market" means that the firm develops the best quality improvement of a variety and supplies the entire "niche" market. If the firm invests too little in research and development, there is an opportunity for an alternative producer to research the quality improvement for that variety and take the market from the incumbent manufacturer.

Each period, a quality improvement destroys the value of the existing quality level, because a firm cannot maintain its niche market without research and development. This is the process of "creative destruction" as described by Aghion and Howitt's (1992) endogenous growth model. Aghion and Howitt (1992) use a contestable monopoly market for an intermediate product. In the model developed here, the contestable market idea is also used, but with multiple industrial sectors, varieties and monopolistic competition. Each variety competes with other varieties in its own subsector through Dixit-Stiglitz preferences. Dixit-Stiglitz preferences, however, are not integral to the model, as growth does not come from increasing product variety. Instead, this provides a simple way of modelling monopolistic competition - other forms of monopolistic competition would yield a similar outcome.

Existing technological knowledge is an input to innovation. Greater knowledge inputs reduce the cost of developing a quality improvement by innovation. Knowledge from the same variety is directly transferable to developing quality improvements in the following period. In addition, knowledge associated with the innovations in other varieties is partially contributing to knowledge inputs in the research and development activities of any given variety. If there is a quality improvement in one variety, it provides some knowledge input to other varieties when quality improvements are developed in future periods. Furthermore, knowledge from other firms is not equally weighted. That is, knowledge from manufacturers within a firm's related technology sector is weighted higher than knowledge from other (less related) technology subsectors. Lastly, knowledge is only partially transferable between locations. If a variety is produced in one location, the knowledge associated with its innovations does not fully transfer to other locations as an input to developing quality improvements.

The result is that firms choose to locate close to other manufacturers that share the same knowledge inputs. Knowledge spillovers provide additional growth. As in the other models of growth and location, growth is highest when there is full agglomeration in one location, as knowledge spillovers are greater with manufacturing concentration. Agglomerated locations are more reliant on local inter-varietal knowledge spillovers for growth, while less agglomerated locations rely on trade, regional knowledge spillovers and local knowledge spillovers within clusters. Traditional goods production is shared between locations and included in the model to ensure trade, even with full manufacturing agglomeration.

The model here includes migration of skilled workers (or footloose skilled labour), in order to help understand innovation and knowledge spillovers in economies where migration may also be a key mechanism for knowledge transfer (Faggian and McCann, 2009). Allowing worker migration gives a better insight into the impact on growth of closer economic integration (beyond trade) between countries and regions. Migration of skilled workers between locations, in response to differences in real wages, equalises real wages in the long run. With zero transport costs, real wages are equal to nominal wages. This simplifies the analysis. However, wage differences may emerge temporarily as certain locations become more advantageous for innovation in specific industries, thereby creating greater demand for skilled labour. Nominal (and real) wages equalise in the steady state due to migrants responding to wage pressure. Through this mechanism, the model offers a better understanding of the influence of migration on economic growth. Migration, competition, location, trade, knowledge and innovation are all affected by closer economic integration between regions and nations. While footloose labour usually makes catastrophic agglomeration inevitable when there are forces for agglomeration, this is not the case here. Instead, the model has a property of (catastrophic) clustering in which industrial clusters are able to remain in the periphery due to hysteresis and the benefits of clustering. Therefore, the model is able to consider growth outcomes and policy in the case of an unequal distribution of economic activity, despite factor mobility.

The model here is highly stylised, but it shows what is happening in high technology, high knowledge sectors where skilled labour is geographically mobile, firm location is not permanent and transport costs are diminishing in relevance for firm location decisions. The model therefore provides a description of differential growth across locations that is particularly suited to describing the present process of global economic development.

5.2.1 Model specification

There are two types of labour: unskilled and skilled labour. As in Grossman and Helpman (1991b) and Young (1998), there is a manufacturing sector and a competitive research and development sector. These two sectors both employ skilled labour. Unskilled labour is employed in the traditional goods sector. Unlike similar models, manufacturing is extended to many sectors. The upper level of the utility function uses a multi-sector Cobb-Douglas function to describe preferences between traditional goods and manufacturing sectors with Dixit-Stiglitz preferences in each industrial sector. To keep the model tractable and without loss of generality, there are assumed to be two locations where firms and workers (consumers) can choose to locate. Consumer demand is indifferent between each manufacturing sector and between domestic or imported goods but consumers have a preference for traditional goods and for variety in each manufacturing sector. Transport costs are assumed zero, so consumers demand home and foreign varieties equally. The analysis is simplified because home and foreign demand for a single variety can be combined in a single demand function from the perspective of the firm.

In each period, a manufacturing firm employs skilled workers to produce the firm's variety with a given quality level for which the firm has a patent. Firms also employ skilled workers in the research and development sector to use existing knowledge of quality levels for all varieties and develop a quality improvement for a variety that the firm will produce in the following period. As in Young (1998), production of each variety is contestable through these quality improvements produced by a competitive research and development sector. For a large enough quality improvement, the firm is granted a patent to produce this improved variety.

As in all endogenous growth models, knowledge spillovers have a vital role. New innovations are based on existing knowledge, but knowledge spillovers transfer imperfectly between firms. The ease of technology transfer depends on technological proximity. The model maintains the approach of Baldwin et al. (2003) that knowledge transfers imperfectly between firms that are spatially separated. The knowledge input from other varieties is weighted according to a related variety approach (Boschma and Frenken, 2009) where the technological relatedness of products determines how useful the knowledge is to innovation in a firm's own variety. This is modelled by assuming that the knowledge input to innovation from within the firm's own sector is weighted higher than knowledge from other sectors. For simplicity, it is assumed that varieties in the same sector are weighted equally and varieties in other sectors are also weighted equally. A modified version of the model could include individual weights for every pair of related varieties but for simplicity in this model, only two possible levels of relatedness are assumed. This technical externality in research and development triggers a "clustering effect" as it induces firms to cluster in locations alongside other firms in their own sector. Firms must also consider this effect alongside incentives to locate in a larger agglomeration where there are more sources of knowledge spillovers from other sectors. This incentive to locate with the larger share of manufacturing, is described as the "agglomeration effect".

Preferences

The representative consumer has a taste for traditional goods and a range of manufactured varieties from several industrial sectors. Consumers also have a preference for higher quality manufactured varieties. The two regions are referred to as home and foreign. The model is described for the home region only and analogous equations apply to the foreign region. Where it is necessary to specify foreign variables, these are denoted by a tilde ($\tilde{}$) above the variable. With $\rho > 0$ as the discount rate, the representative consumer has inter-temporal preferences given by:

$$U = \sum_{t=0}^{\infty} \alpha^t \ln Q_t, \quad \alpha = 1/(1+\rho), \tag{5.1}$$

where Q_t is consumption of traditional goods produced in either region (C_T) and manufactured goods from M sectors in period t with Cobb-Douglas preferences:

$$Q_t = C_{T,t}^{1-\mu} \prod_{i=1}^{M} C_{i,t}^{\frac{\mu}{M}}, \quad 0 < \mu < 1. \tag{5.1a}$$

For simplicity, the time subscript t will be suppressed hereafter where the time dimension is clear. Each of the M industrial sectors is modelled as monopolistically competitive via Dixit and Stiglitz (1977) preferences. The additional factor $A_{i,j}$ represents a further factor of differentiation within each sector: the quality of variety j in industrial sector i,

$$C_i = \left[\sum_{j \in n_i, \tilde{n}_i} (A_{i,j} c_{i,j})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1, \tag{5.1b}$$

where c is the quantity consumed, subscript i indicates each sector and subscript j refers to each of the individual varieties in sector i. Varieties can be produced by domestic manufacturers or imported from foreign manufacturers, where n_i and $\tilde{n}_i$ are the number of home and foreign manufacturers in sector i. A tilde on the variable refers to the location of a foreign manufactured variety (or quality associated with the foreign variety). Symmetry in prices, output and quality levels is assumed for all varieties in each sector i, meaning the variety subscript j can be suppressed and Equation 5.1b can be simplified to:

$$C_i = \left[n_i \left(A_i c_i \right)^{\frac{\sigma-1}{\sigma}} + \tilde{n}_i \left(\tilde{A}_i \tilde{c}_i \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} . \tag{5.1c}$$

where A_i and c_i refer to the quality and quantity of a symmetric variety in sector i. For the remainder of the chapter, variety i, j may be referred to as a single variety in sector i. As will be shown is Section 5.3, the steady state has each sector's output produced in only one location, i.e. either n_i or $\tilde{n}_i$ is equal to zero. While each sector is assumed symmetric, symmetry is not assumed across mulitple sectors and some sectors may have higher or lower measures of quality (and knowledge).

This could also be thought of as a continuum of manufacturing varieties, sectors and continuous time (replacing $\sum$ with integral signs $\int$ and $\prod_{i=1}^{M}$ with the product integral $\prod_{0}^{M}{}^{di}$) but it is easiest to set the model up with a discrete, finite number of varieties, sectors and discrete time periods and assume each n_i is sufficiently large. Continuous sectors would require a Dixit-Stiglitz function (or similar) at the upper level (Equation 5.1a) severely complicating the analysis. Continuous varieties in each sector would yield identical results. Continuous time is not suitable to a deterministic model, because each innovation would be a minuscule incremental increase in quality. Instead the deterministic model in discrete time can be thought of as equivalent to the expected outcome of a stochastic model in continuous time. Discrete variables are used to maintain simplicity, elegance and intuition.

Intertemporal utility optimisation implies the transversality condition and Euler equation:

$$\frac{E_t + \tilde{E}_t}{E_{t-1} + \tilde{E}_{t-1}} = \frac{1+r}{1+\rho}, \tag{5.2}$$

where $E_t + \tilde{E}_t$ is world consumer expenditure in period t, ρ is the rate of time preference and r is the rate of return on savings between periods $t-1$ and t. World expenditure is normalised in each period t such that, $E + \tilde{E} = 1, \forall t$.

Each variety competes only with its own sector via CES preferences. Monopolistic competition in each sector is modelled via CES preferences for simplicity. Ottaviano et al. (2002) point out the limitations of CES preferences in a trade model, but the focus here is the inputs to innovation and the outcomes for growth rather than retail price competition. Other forms of competition may yield a more realistic picture of price competition, but would have similar implications for innovation and firm location and would involve a more complicated derivation.

For simplicity, zero transport costs are assumed such that the price paid and the amount of each variety that arrives at its destination is the price received by the manufacturer and the amount produced. This allows the model to focus exclusively on the location and growth effects of technical externalities in research and development. In an extension of the basic model, it is possible to include trade costs to demonstrate how firms balance many factors in making location and investment decisions, but here the chapter focuses only on technical externalities in research and development. The effect of trade costs are not the focus of this chapter.

Labour

The world is endowed with one unit of labour, consisting of skilled and unskilled workers. In specifying the labour market, the model follows Krugman (1991b) by setting the worldwide stock of skilled workers to μ and the stock of unskilled workers to $(1 - \mu)$, shared equally between regions. The choice of units (μ skilled workers and $1 - \mu$ unskilled workers) follows Krugman (1991b) to ensure that prices and wages in the traditional goods sector are the numéraire, and that the wage rate of unskilled workers equals that of skilled workers. If the number of skilled workers were specified differently, the wages of skilled workers are a constant multiple of the wage rate of unskilled workers. Similarly, a portion of skilled labour could be specified as highly skilled to work in research and development only (as in Chapter 4) and the wages of highly skilled workers would be a constant multiple of the skilled and unskilled workers' wages or using a normalisation to equalise wages. Heterogenous workers, education and skill levels are not the focus of the model and are taken as given. Simplicity is maintained by avoiding these additional multiples. A scaling factor could also be used to calibrate the model to any arbitrary growth or wage rate.

With L_T and $\tilde{L}_T$ the supply of unskilled workers in the home and foreign regions respectively is:

$$L_T = \frac{1-\mu}{2}, \qquad \tilde{L}_T = \frac{1-\mu}{2}, \tag{5.3}$$

so that total production of traditional goods is also shared equally between regions. Unskilled workers cannot migrate between locations. L_K describes the number of skilled workers in the home region (subscript K describes the "knowledge sectors" of manufacturing and research and development) such that:

$$L_K + \tilde{L}_K = \mu. \tag{5.4}$$

Manufacturing and research and development are subject to footloose labour where the migration of skilled workers responds to wage pressure. If there are wage differences between regions at the beginning of a period, there will be migration of skilled workers. The change in skilled workers per period in the home region is given by the migration equation:

$$L_{K,t} = L_{K,t-1} + m_t, \qquad \tilde{L}_{K,t} = \tilde{L}_{K,t} - m_t \tag{5.4a}$$

where M is the migration of skilled workers shifting from the foreign region to the home region in response to differences in wages (or from the home region to the foreign region if M is negative). M is given by:

$$m_t = s_{t-1}\left(1 - s_{t-1}\right)\left(w_{K,t} - \tilde{w}_{K,t}\right), \tag{5.4b}$$

where s is the share of skilled workers in the home region and w_K is the nominal wage of skilled workers in the home region. The ad hoc dynamics from the standard core-periphery model in Fujita et al. (1999) are assumed. With zero transport costs there is no need to define real wages separately from nominal wages for the purpose of migration, because there are no price differences between regions. Workers migrate to the region that provides the highest wage rate. In the model here, manufacturing workers are also skilled and footloose. In all other respects the migration equation used here is identical to Equation 4.14 from Chapter 4.

Firms may face different costs of innovation in different locations. As such, a region with lower innovation costs (i.e. greater knowledge spillovers available) will be more attractive for investment and create a greater demand for skilled labour. This will lead to short-term wage differences that will be overcome by migration. In the steady state there are no real wage differences for skilled workers between locations. While migration has no effect in the steady state, he migration leads to the steady state and provides the requirement that real skilled wages equalise between locations in the long run. Unsteady states are not investigated, where there are uneven real skilled wage states with high rates of migration. Instead the focus here is on the steady state and the effect of industrial clustering on long-run growth. The migration equation above leads to the steady state.

Technology

The model of endogenous growth in this chapter follows Young (1998) with a multi-sector version of the knowledge spillover mechanism in Chapter 4. The innovation cost function allows endogenous growth by vertical innovations without scale effects. A model of growth without scale effects is used intentionally so the clustering effect is solely due to the knowledge spillover externality in research and development and firms do not cluster due to the scale effects present in most endogenous growth models.

In each variety, there is a quality leading firm that developed a quality improvement for this variety in the previous period. In each period, the quality leader produces variety i, j as its niche monopoly and conducts research and development to try to gain a quality improvement large enough to maintain its niche monopoly position for the following period. A firm must conduct research and development to ensure a quality improvement great enough to remain the producer of variety j (or great enough to take the market for variety j from the incumbent firm). Production of an individual variety involves a fixed (labour) investment in innovation (in the previous period) and a constant marginal cost.

Alternatively, Duranton (2007) used a stochastic quality ladders model, where the firm chooses a probability of achieving a fixed level of quality improvement. An incumbent firm never innovates in its own variety because of creative destruction; the firm does not want to destroy the value of its existing patent by developing a quality improvement. In the deterministic model here, the firm knows a rival will develop a quality improvement at the end of each period, so an incumbent can be the innovator in order to continue as the quality leader in the following period. The results of the model here are the same as the expected results of the stochastic model with the exception that an entrant in the next period can be either an incumbent or new entrant. However, the difference is immaterial, because the market outcome after entry from either type of entrant is exactly the same.

The fixed cost of manufacturing in the subsequent period t is the skilled labour requirement in the previous period, $t-1$, to achieve the targeted quality level $A_{i,j,t}$:

$$F_{i,j,t-1}(A_{i,j,t}, \bar{A}_{i,j,t-1}) = \begin{cases} \gamma e^{\eta A_{i,j,t}/\overline{A}_{i,j,t-1}} & \text{if } A_{i,j,t} \geqslant \overline{A}_{i,j,t-1} \\ \gamma e^{\eta} & \text{otherwise,} \end{cases} \tag{5.5}$$

where γ and η are constants used for calibration and $\overline{A}_{i,j,t-1}$ is an index of technological opportunity for variety i, j, representing the intertemporal spillover of knowledge available to variety i, j researchers. The fixed cost can be thought of as two components: a standard fixed cost of γe^{η} irrespective of quality improvement and a cost of $\gamma e^{\eta A_{i,t}/\bar{A}_{i,j,t-1}} - \gamma e^{\eta}$ for achieving a quality improvement.

To simplify the model, only the quality leader in each variety can produce each period. To facilitate this, free entry is assumed for previously produced quality levels below the current quality leader but quality improvements are assumed sufficiently large that only the quality-leading manufacturer of each variety is able to produce each period. All manufacturing firms produce one unit of output per β units of labour. For any firm trying to produce a variety at the non-leading quality level, free entry and unit labour costs lead prices to equal β divided by the wage rate. With an inifinite elasticity of substitution between multiple providers of a single variety, consumers will purchase each variety that has the highest quality to price ratio. The quality leader is able charge a higher limit price and supply the entire market, because consumers prefer the version of variety i with the highest quality to price ratio. Parameters are assumed such that quality improvements are large enough that this limit price is always greater than the profit-maximising monopolist's price. As a result, technology-leading firms will always have the highest quality to price ratio at the profit-maximising price, there are no producers at the non-leading quality level, only the highest quality version of each variety is produced in any period and any former lower quality versions of this variety are no longer produced. This is the contestable market for each variety, because only the quality leader in each variety can produce each period. Therefore, all firms producing in the current period, were innovators in the previous period and were granted a patent to produce in the current period. Moreover, there is no economies of scope and each variety is produced by a different manufacturer.

The model described differs from existing growth models in that the addition of multiple locations and multiple sectors allows a more realistic description of the spillover of knowledge between firms that are technologically and/or spatially separated. The intertemporal spillover of knowledge between firms that are spatially separated is imperfect. Knowledge associated with the quality level of firms in a different location is weighted by the spatial parameter $\lambda_R < 1$. λ_R is a scalar that describes the proportion of knowledge that is available to a firm that is spatially separated from the location of this knowledge. The intertemporal spillover of knowledge also takes place imperfectly between sectors and producers of different varieties. The same logic of firms being separated by geographic space, can also be applied to manufacturers of different varieties being separated by technological space. It is assumed that the knowledge of a firm's own variety is more useful (closer in technological space) than other varieties and knowledge of varieties in the firm's own sector is more useful than knowledge of other sectors. In this way, other firms in the firm's own sector are separated by some technological space from the firm's own variety and firms in other sectors are separated by an even greater technological space.

The knowledge input to innovation is therefore made up of three components: knowledge of the variety's own quality level, knowledge from within the firm's own sector and knowledge from other varieties in other sectors. It is assumed that knowledge from all sources is additive. For developing a quality improvement to produce in period t, the knowledge spillover that is an input to innovation has three weighted components:

1. the knowledge at time $t-1$ from the firm's own variety i, j, represented the by quality level
$$A_{i,j,t-1} \text{ or } \widetilde{A}_{i,j,t-1}, \tag{5.5a}$$

2. the weighted average knowledge of quality from varieties within the firm's own sector i weighted by location
$$\overline{A}_{i\forall(n_i+\tilde{n}_i),t-1} = \frac{\sum_{j\in n_i} A_{i,j,t-1} + \lambda_R \sum_{k\in\tilde{n}_i} \widetilde{A}_{i,k,t-1}}{n_i+\tilde{n}_i} \qquad \text{and} \tag{5.5b}$$

3. a weighted average knowledge of quality improvement from other manufacturing sectors weighted by location

$$\overline{A}_{i\forall M,t-1} = \frac{\sum_{m\in M,\, j\in n_m} A_{m,j,t-1} + \lambda_R \sum_{m\in M,\, k\in\tilde{n}_m} \widetilde{A}_{m,k,t-1}}{\sum_{m=1}^{M} n_m + \sum_{m=1}^{M} \tilde{n}_m}, \tag{5.5c}$$

where A describes the quality improvement in each period and λ_R represents the weighting for knowledge that is sourced from firms in a different location than the firm producing variety i, j. Note that in the steady state, the firms in each sector are clustered in either the home region or the foreign region, but not both. Therefore, each of these components will include only the home region variables or the foreign region variables. In an unsteady state, or between steady states, both types of variables could be included.

Each component of knowledge is assumed to be weighted according to a related variety approach (Boschma and Frenken, 2009), where the relatedness of technology describes how useful the knowledge is to innovation in a firm's own variety. Knowledge of a firm's own variety is given a weight of one, knowledge from innovations within the firm's own sector a weight of $\lambda_V < 1$ and knowledge of other sectors a weight of $\lambda_M < 1$, such that $1 > \lambda_V > \lambda_M \geq 0$. If the firm's selected variety was previously produced in the foreign region, it has the same spatial weight as any foreign knowledge $\lambda_R \leq 1$. Evidently, the relatedness of different varieties in the real world is not as simple. To reflect this, it is possible to weight knowledge from every individual pair of varieties by some kind of proximity measure from which firms choose an optimal location (Boschma, 2005) but this would not demonstrate additional insight in a theoretical model of growth. Therefore, for simplification, it is assumed that weights are constant for all varieties within a firm's sector and constant for all varieties not in the firm's sector. For a home region firm producing variety i, j, the overall index of technological opportunity is given by:

$$\overline{A}_{i,j,t-1} = \max\left(A_{i,j,t-1}, \lambda_R \widetilde{A}_{i,j,t-1}\right) + \lambda_V \overline{A}_{i\forall(n_i+\tilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1}. \tag{5.5d}$$

That is, the index of technological opportunity is the maximum knowledge associated with innovations in the firm's own variety weighted by location plus a weighted average of the knowledge associated with innovations of all other varieties weighted by location and technological relatedness. As a result, firms may face a trade-off between the costs of innovation by locating in a cluster of technologically related firms or locating in an agglomeration of relatively unrelated firms. It is this trade-off which leads to the possibility of including clusters in an endogenous growth model. This has interesting implications for growth policy.

It is assumed that the number of sectors is fixed such that there are always M sectors. This assumption maintains tractability and allows each sector's price index to include only symmetric varieties within its own sector. The number of varieties in each sector is determined by the parameters of the model and new varieties can emerge and replace existing varieties. If the variety has never been produced before, the knowledge of a firm's own innovations is replaced by a weighted average of innovations for its selected sector i. This maintains symmetry in each sector even when a new variety is introduced. The index of technological opportunity for new varieties is given by:

$$\begin{aligned} \overline{A}_{i,j,t-1} &= \frac{\sum_{j\in n_i} A_{i,j,t-1}+\lambda_R \sum_{k\in\tilde{n}_i} \widetilde{A}_{i,k,t-1}}{n_i+\tilde{n}_i} + \lambda_V \overline{A}_{i\forall(n_i+\tilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1} \\ &= (1+\lambda_V)\,\overline{A}_{i\forall(n_i+\tilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1}. \end{aligned} \tag{5.5e}$$

Analogous equations exist for foreign firms.

Following a fixed investment in period $t-1$, firms may produce any quantity in period t at a constant marginal (labour) cost of $w_K\beta$. All workers provide one unit of labour per period.

5.2.2 Short-run equilibrium

The short-run equilibrium describes the prices, wages, production and investment in innovation in each period. Equilibrium prices, wages and production follow from optimisation of Cobb-Douglas utility and within each manufacturing sector from Dixit-Stiglitz utility. The overall optimisation requires firms to choose a location, sector, variety and investment in quality-improving innovations based on the consumers' location, taste for variety and quality and the firms' access to knowledge in alternative locations.

Prices, wages and production.

Consumers allocate expenditure across all sectors to maximise utility subject to the budget constraint

$$P_T C_T + \sum_{i\epsilon M} P_i C_i \leq E + \tilde{E}, \tag{5.6}$$

where $E+\tilde{E}$ describes world expenditure (which is normalised to one), P_T is the price of traditional goods, P_i is the price index of each manufacturing sector i and C_i is the total consumption index of each sector i given by Equation 5.1b. By optimisation and the nature of Cobb-Douglas preferences, $1-\mu$ is the share of expenditure spent on traditional goods and $\frac{\mu}{M}$ is the share spent on all varieties in each industrial sector i:

$$P_T C_T = (1-\mu)\left(E+\tilde{E}\right), \qquad P_i C_i = \frac{\mu}{M}\left(E+\tilde{E}\right) \quad \forall i\epsilon M. \tag{5.6a}$$

Traditional goods

While the focus of this chapter is the knowledge sectors characterised by manufactured goods with increasing quality levels, the traditional goods sector is included for several reasons. In the real world, workers in the traditional goods sector are not necessarily unskilled or immobile, but these are common properties of the traditional sector in trade models. A portion of the workforce remains in each location, even if there is full agglomeration and regions continue to trade. Some factors of production in the real world, such as land, are immobile. While land is not included as a factor of production in the model here, the immobility of unskilled workers is a proxy for immobile production factors. The important feature of the traditional goods sector is that the factor of production is immobile and some portion of all consumption is immobile. "Unskilled labour" (or "peasants") is the commonly used term for this factor in trade models. That is not to say that the types of clustering effects described in high-tech sectors are not also occurring in lower-technology or even traditional sectors, but the overall spatial pattern of economic activity will vary on the basis of the mobility of factors of production and the mobility of knowledge. The model here includes these assumptions in order to elegantly model the general equilibrium and steady state spatial pattern. The implications of the model can then be applied to a range of different sectors and adjusted according to the characteristics of each sector. The traditional goods sector is addressed briefly and otherwise the remainder of the analysis focuses on the knowledge sectors of innovation and manufacturing.

The traditional goods sector is perfectly competitive and has constant returns to scale. Unskilled workers provide one unit of production per period and with 1:1 technology, total production of traditional goods is $C_T = 1-\mu$ produced by both regions. Free trade of goods ensures equal nominal prices and wages in the two regions. With full employment of $1-\mu$ unskilled workers, 1:1 technology and normalised world expenditure $E + \tilde{E} = 1$, the first part of Equation 5.6a is solved such that the traditional goods sector is the numéraire:

$$\begin{aligned} w_T C_T = w_T(1-\mu) = P_T C_T = P_T(1-\mu) = (1-\mu)\left(E+\tilde{E}\right) \qquad \text{and} \\ \tilde{w}_T C_T = \tilde{w}_T(1-\mu) = \tilde{P}_T C_T = \tilde{P}_T(1-\mu) = (1-\mu)\left(E+\tilde{E}\right) \end{aligned} \tag{5.6b}$$

Hence:

$$w_T = \tilde{w}_T = P_T = \tilde{P}_T = E + \tilde{E} = 1. \tag{5.6c}$$

Manufacturing

Cobb-Douglas preferences between manufacturing sectors mean consumers allocate a specific constant proportion of expenditure to each industrial sector. Therefore, each variety competes only with its own sector via CES preferences. In the manufacturing sector, the demand function from home region consumers for variety i, j is:

$$c_{i,j} = \frac{\mu\left(E+\tilde{E}\right)}{M} A_{i,j}^{\sigma-1} p_{i,j}^{-\sigma} P_i^{\sigma-1}, \tag{5.7}$$

where $p_{i,j}$ and $A_{i,j}$ represent the price and quality of variety i, j respectively and P_i is an index for sector i of the price and quality of all varieties in that sector given by:

$$\begin{aligned} P_i &= \left[\sum_{k=1}^{n_i+\tilde{n}_i} A_{i,k}^{\sigma-1} p_{i,k}^{1-\sigma}\right]^{1-\sigma} \qquad (5.7a) \\ &= \left[n_i A_i^{\sigma-1} p_i^{1-\sigma} + \tilde{n}_i \tilde{A}_i^{\sigma-1} \tilde{p}_i^{1-\sigma}\right]^{1-\sigma}. \end{aligned}$$

The sector price index is defined such that the total amount spent on sector i is $P_i C_i = \frac{\mu}{M}\left(E+\tilde{E}\right)$ as in Equation 5.6a. The derivation of this demand function and sector price index is a simple extension of the method described in Appendix A, dropping transport costs and making each manufacturing sector represent a $\frac{\mu}{M}$ share of expenditure. While expenditure was normalised to one in each period, $E+\tilde{E}$, is retained in the notation because expenditure is not normalised across multiple periods. That is, when dynamics are reintroduced with regards to innovation, $E_t+\tilde{E}_t$ does not equal $E_{t-1}+\tilde{E}_{t-1}$ and only one of these can be normalised to one.

In the period prior to production, potential investors/firms choose whether to enter, and if they enter, they select a sector, variety and a level of quality improvement. Since each niche monopoly is contestable by quality improvement, no firm can appropriate the intertemporal knowledge spillover and a fixed cost investment must be made each period in order to produce in the following period. There is no strategic interaction, so production takes place under symmetric monopolistic competition with other varieties in the same sector. Due to Cobb-Douglas preferences between sectors (Equation 5.1a), firms do not compete with firms in other sectors as varieties are only substitutes within their own sector and not substitutes for varieties in other sectors. Since all varieties in the sector are symmetric, quality will be the same for each variety in each specific sector, but quality levels do not have to be the same across sectors.

Relative quality magnitudes between varieties within each sector describe the relative effect on consumer utility. Therefore, all varieties in the same sector have the same impact on consumer utility because symmetry is assumed. However, relative quality magnitudes between sectors do not describe any relative impact on utility because firms compete only within their own sector. By the Cobb-Douglas preferences specified in Equation 5.1a, each sector in aggregate contributes equally to utility. Instead, relative quality magnitudes between sectors describe the knowledge intensity of each sector as an input to innovation relative to the knowledge inputs from other sectors. This comparative technology measure is described as "own sector knowledge intensity" and it is expressed by a relatively higher A_i for sector i. That is, if sector i has a higher own sector knowledge intensity, it means firms in sector i source a higher share of knowledge from within their own sector compared to firms in other sectors who source a lower share from their own sectors.

In period $t-1$, each firm selects a quality improvement for production in period t with its associated cost of innovation (Equation 5.5) and its period t price to maximise the monopolistically competitive profits discounted for time:

$$\max_{p_{i,j,t}, A_{i,j,t}} \frac{(p_{i,j,t} - \beta w_{K,t})\, c_{i,j,t}}{1+r} - w_{K,t} F_{i,j,t-1}\left(A_{i,j,t}, \overline{A}_{i,j,t-1}\right), \tag{5.8}$$

where c_i is demand from domestic and foreign consumers for variety i produced in the home region, β is the marginal and per unit skilled labour requirement of producing one more unit of variety i, j. $F_{i,j,t-1}$ is the number of skilled workers required by the firm in the research and development sector to achieve a target quality level of $A_{i,j,t}$. Each firm maximises profit subject to the demand function for the domestic and exported variety the firm produces. A home region manufacturer is subject to the demand function given by Equation 5.7. As innovation occurs in the period prior to production, firms discount future profit for time. It is assumed that each firm takes expected price and quality setting behaviour of other firms as given and ignores the effects of its own pricing decisions on the price quality index. That is, P_i (or $\tilde{P}_i$) is treated as fixed when differentiating by p_i. These assumptions are plausible with a sufficiently large number of firms in each sector. Firms compete using Bertrand competition, although, with a sufficiently large number of firms, the mode of competition (Bertrand or Cournot) has no effect. By differentiation, the first order conditions are given by:

$$c_{i,j,t} + (p_{i,j,t} - \beta w_{K,t}) \frac{\partial c_{i,j,t}}{\partial p_{i,j,t}} = 0, \tag{5.8a}$$

$$\frac{(p_{i,j,t} - \beta w_{K,t})}{1+r} \frac{\partial c_{i,j,t}}{\partial A_{i,j,t}} + \frac{(\tilde{p}_{i,j,t} - \beta w_{K,t})}{1+r} \frac{\partial \tilde{c}_{i,j,t}}{\partial A_{i,j,t}} - w_{K,t} \frac{\partial F_{i,j,t-1}}{\partial A_{i,j,t}} = 0. \tag{5.8b}$$

As is usual in Dixit-Stiglitz models, free entry is assumed and the number of varieties in each sector is determined endogenously by the elasticity of substitution and the size of each region. With free entry into the research and development sector, firms make zero profits as its entire mark up is invested in innovation. Profits are zero in equilibrium, because if profits were positive, the marginal skilled worker could shift to the R&D sector and produce greater quality improvements or an additional variety. The free entry condition is given by:

$$\frac{(p_{i,j,t} - \beta w_{K,t})\, c_{i,j,t}}{1+r} = w_{K,t} F_{i,j,t-1}\left(A_{i,j,t}, \overline{A}_{i,j,t-1}\right). \tag{5.8c}$$

From the first-order conditions, rearranging Eqution5.8a finds that set prices to a constant mark-up over marginal cost based on the elasticity of substitution:

$$p_{i,j,t} = \frac{\sigma}{\sigma - 1}\beta w_{K,t}. \tag{5.9}$$

With zero transport costs, exported varieties have the same prices as domestic varieties.

In the steady state, (real) wages equalise between locations. Since there are zero transport costs, nominal wages are real. By symmetry and the allocation of labour between skilled and unskilled workers using Krugman's (1991) modelling trick, in the steady state skilled wages are also the numéraire. It is also possible to normalise manufacturing prices to one by setting $\beta = \frac{\sigma-1}{\sigma}$, but in the model here, β is left as it is in order to follow how the productivity of labour affects production.

R&D investment and quality improvement

Dividing the second first order condition (Equation 5.8b) by the free entry condition (Equation 5.8c) and rearranging, firms select a quality improvement where the elasticity of research cost with respect to quality is equal to the elasticity of demand with respect to quality.

$$\begin{aligned} \varepsilon_{A_{i,j,t}}^{c_{i,j,t}} &= \varepsilon_{A_{i,j,t}}^{F_{i,j,t}} \\ \sigma - 1 &= \eta \frac{A_{i,j,t}}{\overline{A}_{i,j,t-1}}. \end{aligned} \tag{5.10}$$

Rearranging Equation 5.10 obtains:

$$\frac{\sigma - 1}{\eta} = \frac{A_{i,j,t}}{\overline{A}_{i,j,t-1}}, \tag{5.10a}$$

which describes the preference of firms to invest in quality improvement. By substitution into Equation 5.5, the cost of innovation or preference to invest by each firm per period is:

$$F_{i,j,t} = \gamma e^{\frac{\eta A_{i,j,t}}{\overline{A}_{i,j,t-1}}} = \gamma e^{\sigma - 1}. \tag{5.10b}$$

Therefore, the skilled labour requirement for innovation by each firm is the same for every variety. Rearranging further, firms select a quality target of:

$$A_{i,j,t} = \frac{\sigma - 1}{\eta}\left[\max\left(A_{i,j,t-1}, \lambda_R \widetilde{A}_{i,j,t-1}\right) + \lambda_V \overline{A}_{i\forall(n_i+\widetilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1}\right]. \tag{5.10c}$$

This is a quality improvement multiplier of:

$$\frac{A_{i,j,t}}{\max\left(A_{i,j,t-1}, \lambda_R \widetilde{A}_{i,j,t-1}\right)} = \frac{\sigma - 1}{\eta}\left[1 + \frac{\lambda_V \overline{A}_{i\forall(n_i+\widetilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1}}{\max\left(A_{i,j,t-1}, \lambda_R \widetilde{A}_{i,j,t-1}\right)}\right]. \tag{5.10d}$$

Assuming this multiplier is always greater than one, there are always quality improvements in equilibrium. Since the firms in each sector are symmetric, they are the same size and have equal quality improvements and levels in the steady state.

Quality improvement per period is given by:

$$I_{i,j,t} = A_{i,j,t} - \max\left(A_{i,j,t-1}, \lambda_R \widetilde{A}_{i,j,t-1}\right) = \\ \left(\frac{\sigma-1}{\eta} - 1\right)\max\left(A_{i,j,t-1}, \lambda_R \widetilde{A}_{i,j,t-1}\right) + \frac{\sigma-1}{\eta}\left[\lambda_V \overline{A}_{i\forall(n_i+\widetilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1}\right]. \tag{5.11}$$

Intuitively Equation 5.11 (also Equations 5.10c and 5.10d) has two components. Quality improvement is made up of the innovation from direct investment in R&D (or R&D-based innovation):

$$\left(\frac{\sigma-1}{\eta} - 1\right)\max\left(A_{i,j,t-1}, \lambda_R \widetilde{A}_{i,j,t-1}\right) \tag{5.11a}$$

plus the quality improvement due to the variety specific knowledge spillover (or knowledge spillover/diffusion based innovation):

$$\frac{\sigma-1}{\eta}\left[\lambda_V \overline{A}_{i\forall(n_i+\widetilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1}\right] \tag{5.11b}$$

giving the total quality improvement as given in Equation 5.11.

Labour market clearing and endogenous variety

Labour market clearing requires that the total labour used in home region manufacturing (L_M) and R&D (L_R) are equal to the total supply of home region skilled workers (L_K). The skilled labour used in manufacturing in the home region is the worldwide expenditure on manufactured goods produced in the home region divided by the price per unit and multiplied by its marginal cost. Since the model has homogenous skilled workers and the same marginal cost in each variety and sector, there is no need to distinguish between varieties i or to distinguish between industries when considering labour market clearing. In the perfect symmetric equilibrium, the number of skilled workers used in goods production by all manufacturers in the home region is:

$$L_M = \frac{\mu S\left(E + \tilde{E}\right)}{p}\beta = \frac{\sigma-1}{\sigma}\mu S\left(E + \tilde{E}\right), \tag{5.12}$$

where $S = \frac{\sum_{i \epsilon M} n_i p_i c_i,}{\sum_{i=1}^{M} n_i p_i c_i, + \sum_{\tilde{i}=1}^{M} n_i \tilde{p}_i \tilde{c}_i} = \frac{\sum_{i \epsilon M} n_i p_i c_i,}{\mu\left(E+\tilde{E}\right)}$ is the total market share of manufacturing expenditure by home region firms, p_i and c_i are the symmetric price and demand of each variety j in sector i and expenditure on each sector i as found in Equation 5.6a provides the simplification. The labour used in research is equal to the number of firms in the next period multiplied by the investment in research labour by each individual firm:

$$L_{R,t} = \gamma e^{\sigma-1} \sum_{i=1}^{M} n_{i,t+1}. \tag{5.13}$$

Home region skilled labour market clearing in period t therefore requires:

$$L_{K,t} = \frac{\sigma-1}{\sigma} \mu S \left(E + \tilde{E} \right) + \gamma e^{\sigma-1} \sum_{i=1}^{M} n_{i,t+1}. \tag{5.14}$$

Analogous equations exist for foreign region manufacturers.

The reward for investing in R&D to develop a quality improvement is the operating profit in the following period. With constant mark-up over marginal cost, the operating profit π is the value of sales shared equally between firms, divided by σ: $\pi = \frac{1}{\sum_{i=1}^{M} n_i} \frac{\mu S \left(E+\tilde{E}\right)}{\sigma}$. By the free entry condition (Equation 5.8c), the amount each firm spends on innovation is its present value share of profit. The free entry relation can be rearranged to:

$$\pi_t = \frac{\frac{\sigma-1}{\sigma} \mu S \left(E + \tilde{E} \right)}{\sigma \sum_{i=1}^{M} n_{i,t}} \frac{1}{(1+r)} = \gamma e^{\sigma-1} = F_{t-1}. \tag{5.15}$$

Substituting this expression (advanced one period) into Equation 5.14, obtains:

$$L_{K,t} = \frac{(\sigma-1)}{\sigma} \mu S_t \left(E_t + \tilde{E}_t \right) + \frac{\mu S_{t+1} \left(E_{t+1} + \tilde{E}_{t+1} \right)}{\sigma (1+r)}. \tag{5.15a}$$

Rearranging the Euler equation given in Equation 5.2 and solving for the value of consumer expenditure on manufactured goods as a function of the parameters gives:

$$\mu S_t \left(E_t + \tilde{E}_t \right) = \frac{\sigma L_{K,t}}{(\sigma-1)+\alpha}. \tag{5.15b}$$

Equilibrium requires the economy to move to a steady state level of consumer expenditure with a constant interest rate $1 + r = 1 + \rho = \frac{1}{\alpha}$. Substituting (Equation 5.15b) into the modified free entry relation (Equation 5.15) and rearranging determines the total number of firms per region across all industries as a function of the parameters:

$$\sum_{i=1}^{M} n_i = \frac{L_K \alpha}{[(\sigma-1)+\alpha] \gamma e^{\sigma-1}}. \tag{5.15c}$$

An analogous function exists for foreign firms. Alternatively, similar rearrangements can be used to find the total number of firms in each industry:

$$n_i + n_{\tilde{i}} = \frac{\left(L_K + \tilde{L}_K \right) \alpha}{M [(\sigma-1)+\alpha] \gamma e^{\sigma-1}} = \frac{\mu \alpha}{M [(\sigma-1)+\alpha] \gamma e^{\sigma-1}}. \tag{5.15d}$$

The last remaining step is to determine how firms within each sector are allocated between regions and whether sectors produce in one or both regions.

5.3 Steady state allocation of manufacturing

This section considers the division of firms and sectors between regions. Starting from any distribution of firms in sector i, even unstable distributions, this chapter studies the incentives for a firm to switch location from the home to the foreign region. By the value of sales alone, firm location would be determined as in Krugman (1991b) such that, with zero transport costs, firms are indifferent between locations. Firms would be located in either region and any distribution of firms would be possible. However, in the model here, firms also consider cost differences in innovation which includes the location and proximity of firms in their own sector and other sectors. Firms must therefore consider the costs of innovation in alternate regions, as the fixed cost of entry in each period.

5.3.1 Contestability, quality improvement and market entry criteria

For a firm to enter its market in period t, the firm must achieve a level of quality improvement that is at least equal to that of other firms in its own sector. If the firm achieves a lower quality improvement, contestability and the free entry condition would mean an entrant could innovate by a greater amount and replace the firm within its niche (variety). The firm would not be able to produce at all, because it could not achieve a patent when it does not have the greatest quality improvement for its niche (or is forced out of the market by having negative profit). Therefore, the firm is required to achieve the highest possible preferred quality from either region in order to enter the market, even if the productivity of innovation in its present location would require additional research workers. It is treated as possible that a firm might still choose to locate in the region that requires the firm to hire more research workers ($\tilde{F} > F = \gamma e^{\sigma-1}$ or $F > \tilde{F} = \gamma e^{\sigma-1}$), in order to understand the trade off faced by the firm in choosing a location. Where a firm considers locating in a less optimal location for R&D with lower knowledge spillovers, the firm must make additional investment by hiring more workers in R&D in order to achieve entry to the market.

The firm's preferred investment in innovation is given by Equation 5.10b and its targeted quality level is given by Equation 5.10c. For a home region firm considering a switch to the foreign location, the function for the preferred investment in innovation in the new location is the same as in Equation 5.10b (with notation $\tilde{F_H}$), but with the knowledge input to innovation adjusted by the new location of the firm:

$$\tilde{\overline{A}}_{i,j,t-1} = \lambda_R A_{i,j,t-1} + \lambda_V \tilde{\overline{A}}_{i\forall(n_i+\tilde{n}_i),t-1} + \lambda_M \tilde{\overline{A}}_{i\forall M,t-1}. \tag{5.16}$$

A switching firm prefers a quality target for production in period t with its associated cost of innovation and its period t price to maximise the monopolistically competitive profits discounted for time. When located in the foreign region the firm ideally also prefers to select a quality improvement where the elasticity of research cost with respect to quality is equal to the elasticity of demand with respect to quality. Elasticities are the same in either region, so Equation 5.10a for the foreign firm is:

$$\frac{\sigma - 1}{\eta} = \frac{\tilde{A}_{i,j,t}}{\tilde{\overline{A}}_{i,j,t-1}} \tag{5.16a}$$

The foreign firm has the same preference in terms of willingness to invest in innovation $\tilde{F}_{i,j,t} = \gamma e^{\sigma-1}$, as the home region firm. Rearranging Equation 5.16a, the preferred quality target for a foreign firm is given by $\tilde{A}_{i,j,t} = \frac{\sigma-1}{\eta}\tilde{\overline{A}}_{i,j,t-1}$.

In assessing the costs and benefits of each alternative location, contestability and the free entry criteria requires the firm achieves the greatest quality improvement available from the alternative location choices in order to participate in the market in period t. That is, $A_{i,j,t} = \max\left(\frac{\sigma-1}{\eta}\overline{A}_{i,j,t-1}, \frac{\sigma-1}{\eta}\tilde{\overline{A}}_{i,j,t-1}\right)$, even if $F_{i,j,t} > \gamma e^{\sigma-1}$. Anything less than this will mean another firm can create a greater innovation in the alternative region and take the market from the incumbent. There are many more factors to consider for firm location such that a firm may choose a less optimal location for other reasons. In the real world, firms would also consider factor prices, trade costs and the value of sales in each location in addition to these innovation cost factors considered here. These additional factors complicate the model and are therefore left aside. The focus here is on the steady state. Greater value of sales or lower factor prices could justify a firm choosing a location that is suboptimal for R&D (or a firm choosing a more optimal location for R&D despite higher factor prices or a lower value of sales). The firm may be willing to do this if there is additional value available by choosing this location.

The purpose of treating a suboptimal location for R&D as possible is to assess the requirements that would encourage a firm to switch regions, leading to the steady state. Furthermore, the approach maintains comparability with other two-region models and allows additional factors of firm location to be added (such as transport costs or rent) as an extension of the model presented here.

5.3.2 The requirements for a steady state

The steady state is defined as an unchanging regional division of manufactured varieties, innovation expenditure, prices and quantities defined by the short-run equilibrium above. This section considers possible distributions of manufacturing (both steady and unsteady states) where migration of skilled workers due to a spatial inequality of real wages and the switching of the location of firms due to differences in knowledge spillovers lead to the steady state. Some models include in the definition of a steady state the requirement of a constant rate of growth in both quality and output. However, with growth in quality defined as $g_A = \frac{I_{i,t}}{A_{i,t-1}}$, the rate of growth in quality changes over time as knowledge spillovers change between periods, even in the steady state. Therefore, the definition of a steady state is modified to a constant investment in R&D, but the productivity of R&D effort can change over time. There is a constant quality improvement from R&D based innovation, as defined by Equation 5.11a, even though there may be declining or increasing diffusion-based innovation (Equation 5.11b). This is a steady state, because firms and workers have no incentive to switch region between periods and therefore the distribution of economic activity is "steady". This definition of a steady state is required, because sectors with a relatively higher quality level A achieve lower rates of quality improvement from diffusion than sectors with a low quality level A and therefore, the spillovers from other sectors change over time. This is a similar relationship to that discussed in the distance to frontier literature (Acemoglu et al., 2006), but focused on the relatedness between different varieties. Distributions of economic activity that are steady in each region are possible, even if growth rates are not steady due to changes in technology varying by sector.

Consider an unsteady state with a cluster of firms in the home region defined by a relatively greater number of firms in the same industrial sector i locating in the home region ($n_i > \tilde{n}_i$). In an unsteady state wages may vary between regions and with $n_i > \tilde{n}_i$, Section 5.6a found that $\tilde{F}_{i,j,t} > \gamma e^{\sigma-1}$. A home region firm in that cluster will only switch if the firm can achieve greater return on investment in the new location. Considering return on investment between each location:

$$\frac{\tilde{V}}{\tilde{w}\tilde{F}} > \frac{V}{wF}, \qquad \tilde{F} < F, \tag{5.17}$$

where V is defined as price less marginal cost multiplied by total production in one period. In the model with zero transport costs and mobile skilled workers: wages, prices and value are equal between locations in the steady state. Therefore, a firm will choose the location where the cost of innovation is the lowest (knowledge spillovers are greatest) The requirement for a firm to remain located in the home region is given by $\tilde{F}\tilde{w} \geq Fw$. When a sector (or large portion of a sector) does switch locations, its demand for labour may temporarily raise the local wage rate in the new location. However, with migration given in Equation 5.4b, unequal wage rates are not sustained because migration equalises wages and reinforces the location of the switching sector. The requirement for a firm to remain located in the home region in the steady state simplifies to:

$$\tilde{F} \geq F, \qquad \forall j, i, \tag{5.17a}$$

although the switch between locations could take multiple periods because migration does not happen instantly.

All firms in the same sector and region have the same cost of innovation. If a single firm switches, the number of workers employed in innovation in the foreign region for other firms in the same sector will decrease. Each individual firm is small relative to the size of the entire market, so it is assumed that individual firms to not account for any effect on wages from switching location. Therefore, if Equation 5.17a does not hold true for one firm in sector i such that the firm switches location, it will also not hold true for other firms in sector i (even more so after the first firm switches) such that all home region firms in the sector will also switch location. In addition, if a sector were shared equally across two regions, a single firm switching means one region would now have the larger share of industry and Equation 5.17a would no longer hold true for firms remaining in the original location. As a result, each sector will remain clustered in one location in the steady state, determined by hysteresis, until Equation 5.17a no longer holds.

Knowledge spillovers are greater with concentration of firms in the same sector ($n_i > \tilde{n}_i$) i.e., clustering, and with manufacturing concentration ($\sum_{i\epsilon M} n_i > \sum_{i\epsilon M} \tilde{n}_i$) i.e., agglomeration. These two factors determine firm loation. Firstly, firms prefer locations with a greater share of their own sector such that firms in each sector cluster in a single location - the so-called "clustering effect." But firms must balance this attraction with a preference to locate where there are more firms overall, because greater concentration of all manufacturers also increases knowledge spillovers. This alternative force for firm concentration with all manufacturing firms is described as the "agglomeration effect". Depending on the distribution of each sector, these forces may be in the same direction but could potentially be in opposite directions. Sectors that cluster in the smaller region may still sustainably produce in that location if the clustering effect is greater than the agglomeration effect and in the opposite direction. This scenario is described as a "peripheral cluster".

As described above, the quality improvement required for entry is set by the highest level of quality from either region that is available for the fixed cost of $\gamma e^{\sigma-1}$. Assume this quality level is obtained in the home region such that $\tilde{A}_{i,j,t} < A_{i,j,t}$ and $F = \gamma e^{\sigma-1}$. The cost of achieving the quality level $A_{i,j,t}$ for a firm that is switching to the foreign location is:

$$\tilde{F} = \gamma e^{\frac{\eta A_{i,j,t}}{\widetilde{\overline{A}}_{i,j,t-1}}} \tag{5.18}$$

Firms select a quality target that is assumed to be determined in the home location, given by Equation 5.10c. Analogous equations exist if the foreign region is the technology leading region for variety j. The intertemporal spillover of the firm's own knowledge diminishes by $1 - \lambda_R$ when the firm switches. Substituting the knowledge input, modified for the foreign region (5.16), and the targeted quality level (Equation 5.10c) into Equation 5.18 gives:

$$\tilde{F} = \gamma e^{(\sigma-1)\frac{A_{i,j,t-1}+\lambda_V \overline{A}_{i\forall(n_i+\tilde{n}_i),t-1}+\lambda_M \overline{A}_{i\forall M,t-1}}{\lambda_R A_{i,j,t-1}+\lambda_V \widetilde{\overline{A}}_{i\forall(n_i+\tilde{n}_i),t-1}+\lambda_M \widetilde{\overline{A}}_{i\forall M,t-1}}}. \tag{5.18a}$$

The difference between entry costs in the home region F and entry costs in the foreign region $\tilde{F}$ is the exponent in $\tilde{F}$ which is multiplied by the ratio of knowledge from all firms in either location, where the weightings depend on the locations of firms relative to each alternative. The relative magnitudes of F and $\tilde{F}$ depend upon the distribution of firms within sector i and on the distribution of all firms. If foreign knowledge spillovers are lower, there will be a greater cost of innovation in the foreign region as given by Equation 5.18a. Substituting (5.18a) and $F = \gamma e^{\sigma-1}$ into Equation 5.17a and rearranging shows that in the steady state, the firm chooses the location where knowledge spillovers are greater. In this case firms choose the home location, because

$$A_{i,j,t-1} + \lambda_V \overline{A}_{i\forall(n_i+\tilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1} \geq \lambda_R A_{i,j,t-1} + \lambda_V \widetilde{\overline{A}}_{i\forall(n_i+\tilde{n}_i),t-1} + \lambda_M \widetilde{\overline{A}}_{i\forall M,t-1}. \tag{5.19}$$

Intuitively, it can be seen that the inequality holds for two types of sectoral steady states. In the first steady state, all industry is agglomerated in a single location. The clustering effect from locating alongside producers of technologically related varieties (i.e. the same sector) and the agglomeration effect from locating alongside other manufacturers, are both in the same direction towards the agglomerated location. An alternative scenario where each sector is shared equally between the two locations is not a steady state, but a knife-edge. If a single firm were to switch locations, it would be to the location that would become the marginally larger region and therefore all firms would also want to switch to the larger region. Therefore, the second alternative steady state is where each sector is clustered in a single location and sectors are shared between locations.

By symmetry and clustering of each sector, the inequality becomes much simpler such that $A_{i,j,t-1}$ for all varieties in sector i and $\overline{A}_{i\forall(n_i+\tilde{n}_i),t-1}$ can be denoted as the quality parameter for any variety in sector i, $A_{i,t-1}$. Furthermore, by substituting 5.5c the inequality can be rearranged to describe a knowledge intensity threshold for sector i to produce sustainably in the home region:

$$A_{i,t-1} \geq \frac{\lambda_M}{1+\lambda_V} \frac{\sum_{m\epsilon M} \widetilde{n}_m \widetilde{A}_{m,t-1} - \sum_{m\epsilon M} n_m A_{m,t-1}}{\sum_{m\epsilon M} n_m + \sum_{m\epsilon M} \widetilde{n}_m}. \tag{5.19a}$$

If sector i has a quality target greater than this threshold level, it is possible for sector i to be clustered in the home region in the current period, even if the home region is not the location of other sectors. If all other sectors are agglomerated in the foreign region, this increases the threshold for the quality parameter in sector i. If this threshold is satisfied for all sectors, this is a steady state for the current period, because no single firm will switch location in the coming period, wages are equal across locations so there is no change in labour endowments in each region, all firms will grow at the same rate in each industry and will continue to grow in future periods. In the steady state, this threshold property is easy to test for each variety, because, to be met in all sectors, it only needs to be tested for the variety (or sector) with the lowest technology level in each location. If technology in any single sector is below this threshold, this sector will switch region and the relevant threshold will be redetermined.

The steady state was defined such that a distribution of economic activity is sustainable indefinitely. Therefore, this technology parameter threshold must be met indefinitely for the distribution to be a steady state. The last case to consider is whether greater innovations in the agglomerated sectors in the foreign region lead the threshold to grow faster than quality in the peripheral cluster. That is, equation 5.19a must be met for all time periods. If it is not met in the current period (for the innovations that occurred in $t-1$), the firm will switch. There are two clear long-run steady states. The core-periphery outcome is where all sectors cluster in a single region and is a long-run steady state where all firms benefit from co-locating. Alternatively, the equal distribution outcome where half the sectors are clustered in each region is a steady state if there is also an equal distribution of technology intensities. Each location will have equal growth in the quality levels of comparable technology-intensive industries, so there will be no incentive for firms or workers to switch location during any time period.

A third type of steady state, the peripheral cluster equilibrium, where clusters of firms in the same sector(s) are located in the region with a smaller share of all industrial sectors, is also possible to be a long-run steady state if the increases in quality levels in the peripheral cluster are greater than or equal to the change in the threshold. This allows the threshold to hold in subsequent periods. Consider how the technology threshold changes over time. Taking the discrete derivative of the threshold (Equation 5.19a), there is an additional threshold that determines whether the distribution is a long-run steady state:

$$\Delta_t A_{i,t-1} \geq \Delta_t \left(\frac{\lambda_M}{1+\lambda_V} \frac{\sum_{m \epsilon M} \widetilde{n}_m \widetilde{A}_{m,t-1} - \sum_{m \epsilon M} n_m A_{m,t-1}}{\sum_{m \epsilon M} n_m + \sum_{m \epsilon M} \widetilde{n}_m} \right). \tag{5.19b}$$

Since each industry is clustered only in one location, either $A_{m,t}$ or $\widetilde{A}_{m,t}$ applies to the spillovers from each of the other industries but not both, because each industry does not exist in both locations. The discrete derivative of the quality target function (5.10c) with respect to time yields:

$$\Delta_t A_{i,t-1} = I_{i,t} = \left(\frac{\sigma - 1}{\eta} - 1 \right) A_{i,t-1} + \lambda_V \overline{A}_{i \forall (n_i + \widetilde{n}_i),t-1} + \lambda_M \overline{A}_{i \forall M,t-1}. \tag{5.19c}$$

Substituting this into the differentiated inequality (Equation 5.19b) and rearranging gives:

$$I_{i,t} \geq \frac{\lambda_M}{1+\lambda_V} \left(\frac{\sum_{m \epsilon M} \widetilde{n}_m \widetilde{I}_{m,t} - \sum_{m \epsilon M} n_m I_{m,t}}{\sum_{m \epsilon M} n_m + \sum_{m \epsilon M} \widetilde{n}_m} \right). \tag{5.19d}$$

As long as the size of innovation is greater than the difference between the aggregate innovations in either region, divided by the total number of firms and multiplied by $\frac{\lambda_M}{1+\lambda_V}$, sector i can last indefinitely in a peripheral cluster.

Since $I_{i,t} = A_{i,t} - A_{i,t-1}$, the thresholds can be combined (Equations 5.19a and 5.19d):

$$\begin{gathered} \frac{\lambda_M}{1+\lambda_V} \frac{\sum_{m \epsilon M} \widetilde{n}_m \widetilde{A}_{m,t-1} - \sum_{m \epsilon M} n_m A_{m,t-1}}{\sum_{m \epsilon M} n_m + \sum_{m \epsilon M} \widetilde{n}_m} \leq A_{i,t-1} \leq \\ A_{i,t} - \frac{\lambda_M}{1+\lambda_V} \left(\frac{\sum_{m \epsilon M} \widetilde{n}_m \left(\widetilde{A}_{m,t} - \widetilde{A}_{m,t-1} \right) - \sum_{m \epsilon M} n_m (A_{m,t} - A_{m,t-1})}{\sum_{m \epsilon M} n_m + \sum_{m \epsilon M} \widetilde{n}_m} \right). \end{gathered} \tag{5.19e}$$

Since Equation 5.19a is already satisfied, Equation 5.19e can be rearranged to:

$$A_{i,t} \geq \frac{\lambda_M}{1+\lambda_V} \frac{\sum_{m \epsilon M} \widetilde{n}_m \widetilde{A}_{m,t} - \sum_{m \epsilon M} n_m A_{m,t}}{\sum_{m \epsilon M} n_m + \sum_{m \epsilon M} \widetilde{n}_m}. \tag{5.19f}$$

This is the same as the earlier threshold advanced one period. Therefore, if the threshold is met for technology levels in the current period for the marginal industries (one in each region), it will be met for all industries in all future periods. As a result, whenever the threshold is met for all sectors, the distribution of technology and economic activity is a steady state.

Summary of steady states

Three three possible steady states have been derived:

1. Equal Distribution: even distribution of technology and number of sectors per region.

2. Core-Periphery: all industry agglomerates in a single region.

3. Peripheral Cluster: an industry that is own industry technology intensive produces sustainably in the periphery. A home region peripheral cluster in sector i must have a knowledge input to innovation that satisfies:

$$A_{i,t-1} \geq \frac{\lambda_M}{1+\lambda_V} \frac{\sum_{m \epsilon M} \widetilde{n}_m \widetilde{A}_{m,t-1} - \sum_{m \epsilon M} n_m A_{m,t-1}}{\sum_{m \epsilon M} n_m + \sum_{m \epsilon M} \widetilde{n}_m}.$$

5.3.3 Simulations

In this section, quality improvement is simulated under various examples of each of the steady states. The simulations here do not reflect the characteristics of a specific economy, but they are a numerical exercise to demonstrate the effects on steady state growth rates from the trade off between the clustering and agglomeration effects. In particular, these simulations emphasise the steady state that is unusual in two region growth models with mobile labour, where clusters with sufficiently high own industry knowledge intensity can grow sustainably in a peripheral location.

Parameters

The model is calibrated based on parameters used in other simulations. Krugman (1991b) uses the parameters $\sigma = 4$ and $\mu = 0.3$. In a product variety growth model, Baldwin and Forslid (2000) use the parameters $\sigma = 5$, $\mu = \frac{1}{4}$ and $\alpha = \frac{1}{2}$ (this implies an annual discount rate of approximately 7% when periods are 10 years).

Since there are multiple sectors where firms only compete with their own sector by Cobb-Douglas preferences, then the elasticity of substitution within each sector in the model should be considerably greater than one[1], while the elasticity of substitution between varieties in different sectors is exactly one. To determine appropriate parameters for the multi-industry model, the simulations use average elasticity of substitution of six between any two varieties. With M industries and $n_i + \tilde{n}_i = n$ firms in each industry there are $\frac{Mn(Mn-1)}{2}$ pairs of products. Only varieties within the same industry are substitutes such that, there are $\frac{Mn(n-1)}{2}$ pairs with an elasticity of substitution equal to $\sigma > 1$, and all other pairs have an elasticity of one. The average elasticity is set to $\bar{\sigma} = \frac{Mn(n-1)\sigma + Mn(Mn-1) - Mn(n-1)}{Mn(Mn-1)} = 6$ such that there is an elasticity of $\sigma = (\bar{\sigma} - 1)\frac{Mn-1}{n-1} + 1$, within each industrial sector. Setting the number of firms per industry to 10 and the number of industries to 10, suggests an elasticity of $\sigma = 56$. 100 varieties across 10 sectors is sufficient to demonstrate growth and location in a peripheral cluster, although more or less sectors and firms would also demonstrate the spatial, clustering and technical externality impacts on economic growth.

[1]The exact choice of elasticity used in the simulations here is arbitrary. All other parameters are calibrated accordingly and the exercise is purely a numerical example of how the model works rather than a representation of any true economy. Using a different value for the elasticity does not affect the qualitative conclusions in this model.

The value of γ is calibrated to other selected parameters and determines the amount of labour required by each firm as its fixed cost of market entry in the following period. Substituting into Equation 5.15c for both regions along with the parameters $\alpha = \frac{1}{2}$, $M = 10$, $n_i + \tilde{n}_i = 10$ and $\left(L_K + \tilde{L}_K\right) = \mu = \frac{1}{4}$, finds that $\gamma = \frac{\left(L_K+\tilde{L}_K\right)\alpha}{M(n_i+n_{\tilde{i}})[(\sigma-1)+\alpha]e^{\sigma-1}} = \frac{0.125}{5550e^{44}}$. The exact choice of parameters does not affect the conclusions drawn. This calibration produces small values of production, as it allows for 100 different varieties to be produced by a labour force of $\mu = 0.25$ and that labour force is shared between R&D and manufacturing. The focus here is not on the exact values of production but on the growth rates of quality levels under different steady states.

The last parameter to calculate is η which is used to calculate the fixed costs a firm faces in each period regardless of its quality target γe^{η}. Calibrating η such that these fixed costs are half of the overall fixed costs of quality improvement gives, $\eta = \ln\left(\frac{1}{2}\right) + \sigma - 1$. The initial values for A_i across the 10 sectors are set in order to demonstrate each of the steady states. Knowledge spillover parameters are set at $\lambda_R = 0.7$, $\lambda_V = 0.5$ and $\lambda_M = 0.2$. These values represent the assumed transmission cost $(1-\lambda)$ of knowledge across geographic space, the technology space between varieties in the same sector and the technology space between varieties in different sectors.

Each of the steady states are simulated. Simulations start with an initial endowment of firms, workers and technology levels in each location that belongs to a long-run steady state. Simulations run for 15 periods, although simulations can be run indefinitely, but do not show any further detail over the long run. The simulations presented are not exhaustive. They represent a numerical exercise which are suggestive of the insights that could be gained from a simulation calibrated to a specific set of regions. Since this would require a range of explained insights to inform calibration, this is clearly beyond the scope of the chapter or thesis. Instead the focus is on calibrating the model as a numerical exercise that draws relevant conclusions about the nature of innovation, growth and policy implications.

Equal distribution

In the equal distribution outcome, quality improvement is simulated with two identical regions where each sector is clustered in one region or the other but both regions have an equal number of sectors. Furthermore, for each sector, there is an equivalent sector in the other region with the same level of own industry knowledge intensity. Starting values of technology levels for sectors 1 to 5 in the home region are set at 1 through 5 and for sectors 6 to 10 in the foreign region are also set at 1 through 5. Since the regions are identical, a firm in either region receives the same level of knowledge spillovers and grows at the same rate as a firm in the other region with the same starting technology level. Figure 5.1 shows quality levels for the 5 sectors in either region.

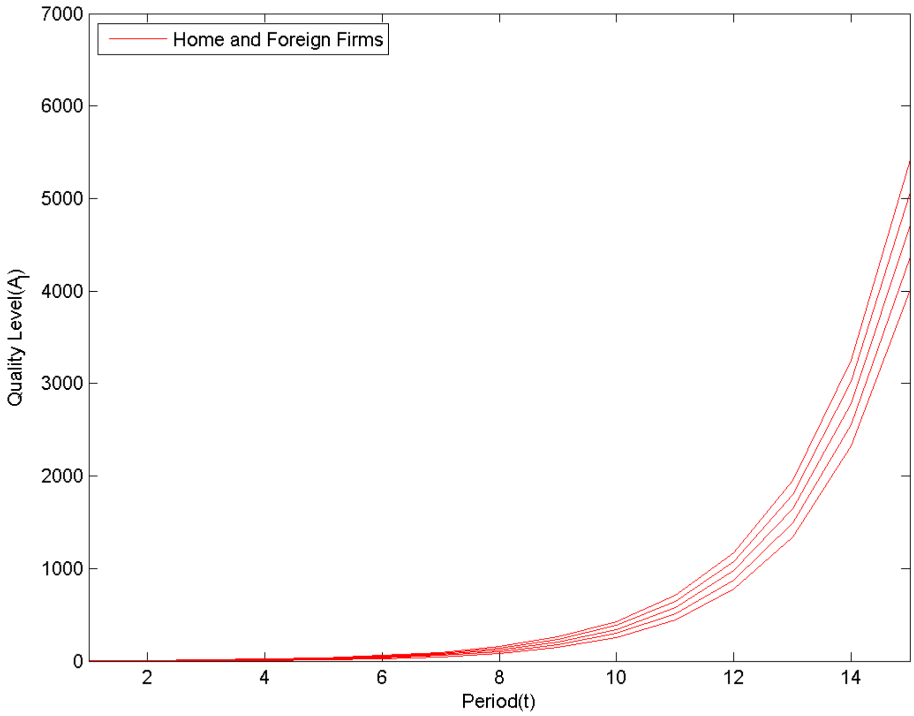

Figure 5.1: Quality levels in the equal distribution steady state

Figure 5.2 shows the same simulation on a log axis of quality levels. It can be seen that lower knowledge intensity sectors have a greater growth rate than higher intensity firms, as these firms benefit from greater knowledge spillovers.

Figure 5.3 shows the growth rate in quality levels per sector and period. Note a period is 10 years, so a growth rate of 1 equates to an average of approximately a 7% quality improvement per year. Over time, growth rates converge, as faster growth in low knowledge intensity sectors results in a convergence of growth rates across sectors. Growth rates of higher knowledge intensity sectors increase over time, while growth rates of lower knowledge intensity sectors decrease over time. This convergence is not due to diminishing returns to scale, as the model is chosen specifically because of constant returns to scale. Instead convergence is caused by diminishing returns to relative scale becuase knowledge spillovers from other sectors vary between sectors on the basis relative "own industry knowledge intensity". Sectors with a knowledge intensity greater than the average, weighted by relative location, start with relatively lower growth rates because they receive relatively lower spillovers from other sectors. But their growth rate increases over time because the effect of diminishing returns to relative scale declines with their relatively lower growth rate. Alternatively sectors with a knowledge intensity less than average, weighted by relative location, start with relatively higher growth rates because they receive relatively greater

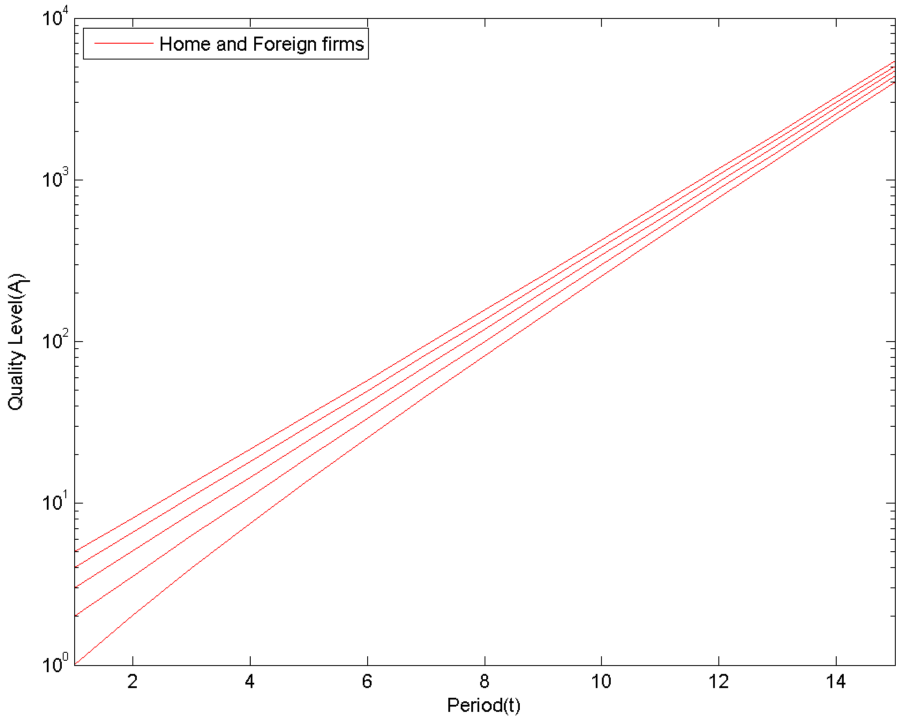

Figure 5.2: Log plot of quality levels in the equal distribution steady state

spillovers from other sectors. Growth rates decline because of the effect of diminishing returns to relative scale increases with their relatively higher growth rate. As a result, own industry knowledge intensity and growth rates converge.

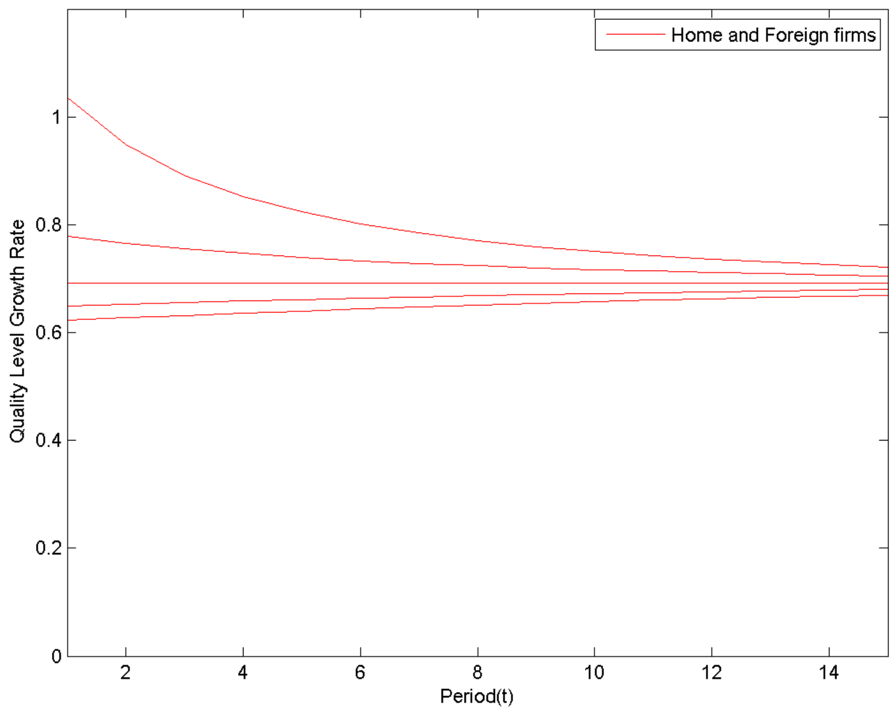

Figure 5.3: Growth rate of quality levels in the equal distribution steady state

Core-periphery

In the core-periphery outcome, quality improvement is simulated where all manufacturing sectors are located in a single agglomerated location. There are 10 sectors with the same technology levels as the equal distribution simulation above, but all firms are agglomerated in a single location. A direct comparison shows the additional growth effects of agglomeration as firms benefit from greater knowledge spillovers with closer spatial proximity. Figures 5.4 and 5.5 show the quality levels for the core-periphery steady state and can be directly compared to Figures 5.1 and 5.2.

Figure 5.6 shows the growth rate in quality levels per sector and period. Growth rates are higher with agglomeration in a single region than in the equal distribution outcome. Once again, growth rates between sectors converge as lower knowledge intensity sectors benefit relatively more from inter-industry spillovers. This reflects the diminishing returns to relative scale identified previously although sectors converge to a higher long run growth rate because there are no spatial effects on the weighted averages of knowledge intensity due to full agglomeration in the foreign region.

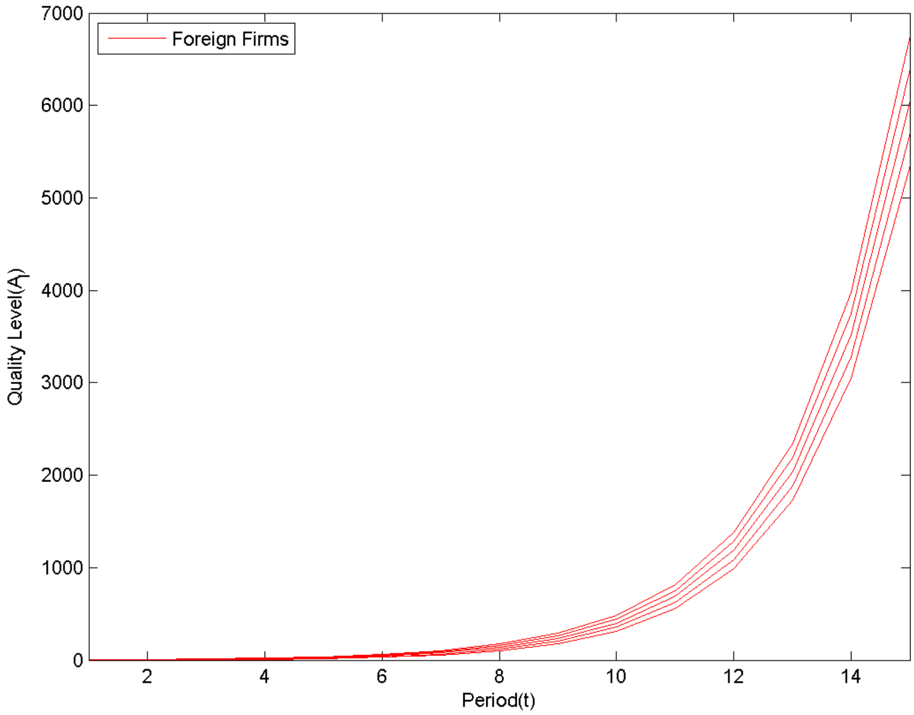

Figure 5.4: Quality levels of foreign firms in the core-periphery steady state

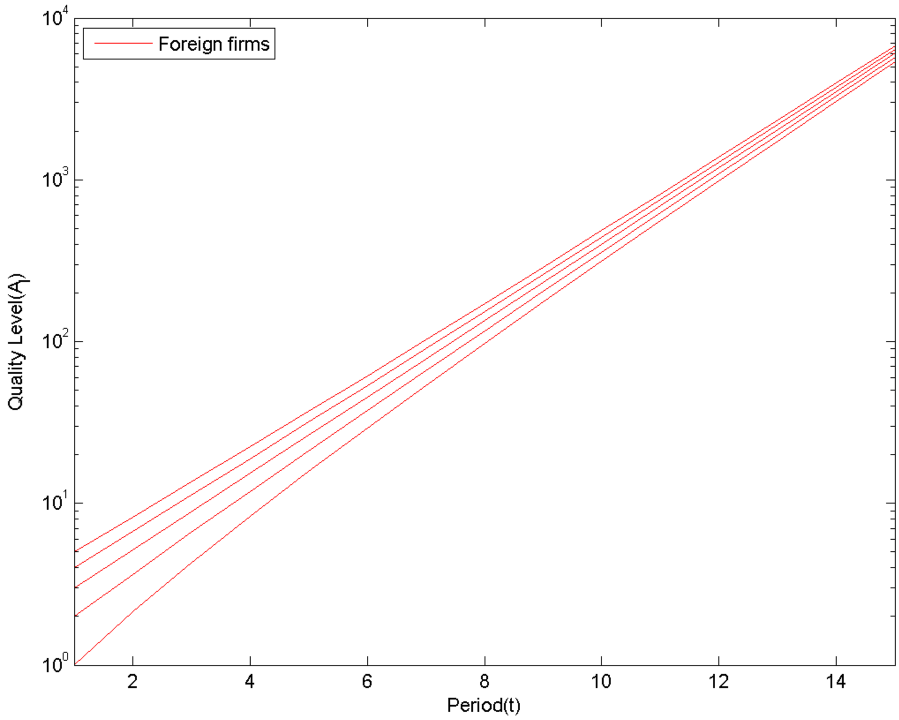

Figure 5.5: Log plot of quality levels in the core-peripheral steady state

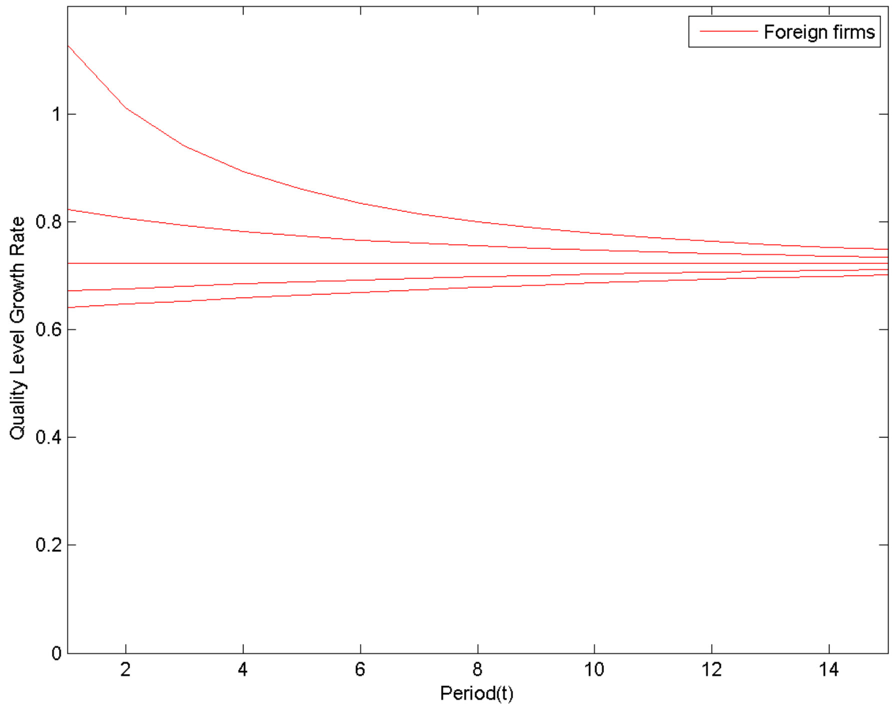

Figure 5.6: Growth Rate of Quality Levels in the Core-Periphery Steady State

Peripheral cluster

In this steady state, quality improvement is simulated where a minority share of sectors is located in a single region. The region with fewer sectors is described as the periphery as a sector in this region is described as a "peripheral cluster". Once again, there are 10 sectors but with three sectors in the home region (peripheral cluster) and seven sectors in the foreign region (core). Knowledge intensity levels are set to compare equivalent sectors in each region. The three sectors in the periphery (home region) have starting quality levels of 1 through 3 while in the agglomerated foreign region there are three firms with technology at 1, two firms at technology 2 and two firms at technology 3.

Figures 5.7, 5.8 and 5.9 show quality levels as in the above examples. It can be easily noticed that peripheral clusters grow at a lower rate than clusters in the agglomerated location. For example, for the highest quality sectors in each region these have the same starting values, but after 15 periods, the high knowledge intensity sectors in the home region has fallen to similar quality levels to the sector in the foreign region with second highest knowledge intensity. For each of the peripheral clusters, growth rates are below growth rates of comparable sectors in the agglomerated location. As a result, quality levels of peripheral clusters diminish relative to equivalent sectors in the the core because peripheral firms receive a lower level of total knowledge spillovers due to spatial externalities in knowledge transfer between locations.

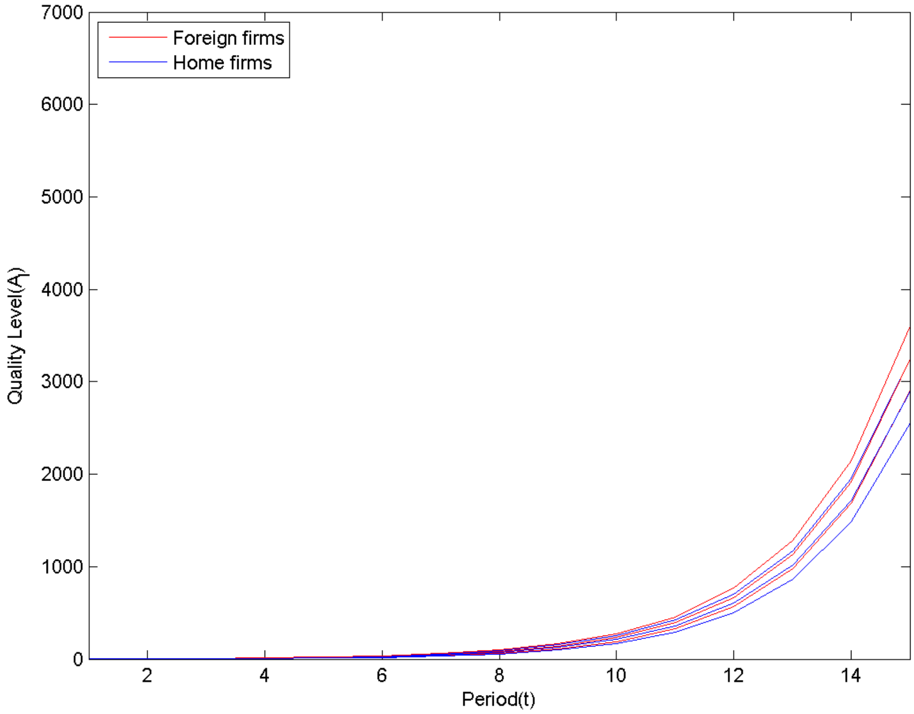

Figure 5.7: Quality levels of foreign firms in the peripheral cluster steady state

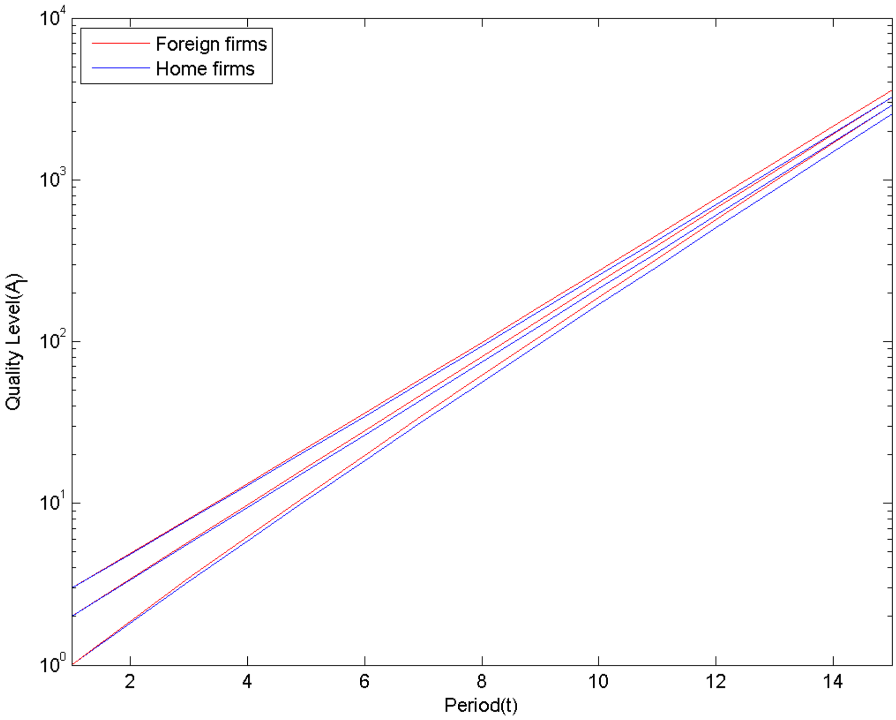

Figure 5.8: Log plot of quality levels in the peripheral cluster steady state

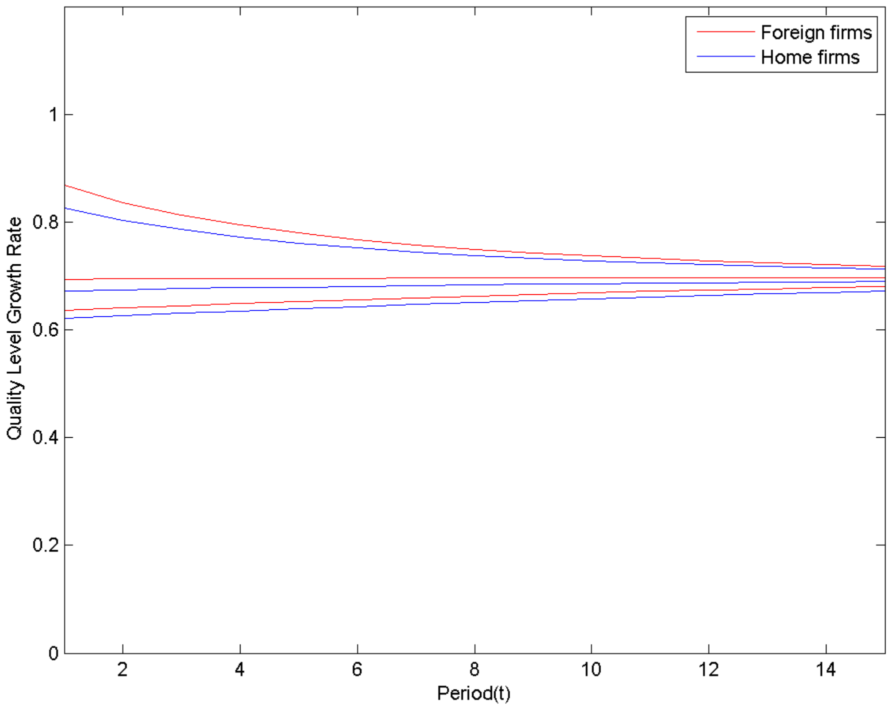

Figure 5.9: Growth rate of quality levels in the peripheral cluster steady state

As above, there is also convergence between sectors with different quality levels due to diminishing returns to relative scale. However, the home region converges to a lower growth rate in quality than the foreign region because of the weighted average of world knowledge intensities is lower than in the foreign region due to relatively greater spatial effects on knowledge transferred to firms in the periphery region.

5.4 Discussion

The model here suggests a sorting of firms between agglomerated and peripheral locations. While all firms can sustainably produce in the core so long as they continue investing in innovation, only sectors with a high enough ”own industry knowledge intensity” can produce sustainably in the periphery.

The model here offers several advantages over existing models of growth. The two-region endogenous growth models of Baldwin et al. (2003) inevitably result in catastrophic agglomeration at low transport costs if there is mobile labour. By developing a model with multiple industries and spillovers based on technological relatedness, firms balance the forces for clustering (also referred to as localisation) against forces for agglomeration (or urbanisation). The result is a model where industries can cluster in the periphery in the long term, provided they have sufficient own industry knowledge intensity. The ability to consider clusters within an endogenous growth model helps to explain the distribution of economic activity and economic growth such that, even with mobile labour, not all industry catastrophically agglomerates in a single region. The model suggests that with a quality ladders approach to modelling economic growth, the mobility of firms and sectors between locations is dependent on own sector knowledge intensity. Notably, it is the sectors with greater own sector knowledge intensity which are more sustainable in a peripheral location than industries with a lower own sector knowledge intensity which are more affected by forces for agglomeration.

In addition, the lower knowledge intensity sectors benefit more from industry spillovers resulting in higher growth rates. This has interesting implications for economic growth performance in different locations. In particular, the market failure of innovation (where firms underinvest in innovation, because they cannot appropriate the entire benefit of their innovations) is greater for sectors with higher own industry knowledge intensity. The implication is that high knowledge intensity sectors require greater support for innovation than low knowledge intensity sectors. Furthermore, peripheral regions require greater support for innovation, because they are unable to benefit fully from inter-industry knowledge spillovers and because the only sectors which can produce sustainably in the periphery are higher knowledge intensity sectors.

This section discusses further the characteristics of the model and its steady states, identifying implications for economic growth policy in different locations.

5.4.1 The additional market failure for innovation in related technologies

Endogenous growth theory describes a market failure of innovation (Aghion and Howitt, 1992; Grossman and Helpman, 1991b; Romer, 1990) where firms underinvest in innovation activities. Investors in R&D do not invest at the socially optimal level of innovation, because they cannot appropriate all of the future benefits generated from their investment. Those benefits accrue to future innovators. This feature is also present in the model here, but there is an additional market failure because of knowledge spillovers between related technology sectors.

Each innovation has benefits in terms of quality improvement in other varieties and sectors in subsequent periods. The revenue stream from these additional quality improvements does not accrue to the original innovators. Conversely, investors in innovation benefit from the innovations made by others, because these improve the level of quality improvement that is attainable. If, hypothetically, these costs and benefits were shared on an equal basis, firms would provide benefits to other firms of an equal amount to the benefits they receive from other firms' innovations. The market failure that investors in innovation do not receive the full benefit of their investment would be exactly offset by the benefits firms receive in knowledge externalities from other sectors. There could still be some role for policy as firms receive spillovers regardless of their investment in innovation, but policies could be non-discriminatory across sectors. However, the supply of knowledge to other sectors and benefits from other varieties' knowledge spillovers does not accrue evenly across sectors which has some interesting new implications for innovation policy.

Sectors with a high level of knowledge intensity have relatively lower rates of growth due to receiving relatively lower knowledge spillovers from other sectors. High knowledge intensity firms provide greater knowledge spillovers to other sectors than what is actually received by these firms in spillovers from other sectors. Alternatively, sectors with a low level of knowledge intensity have higher rates of growth, with a large share of quality improvement due to knowledge spillovers from other sectors. That is, lower knowledge intensity sectors have a smaller market failure of innovation, because they benefit more from knowledge spillovers. The knowledge spillovers from other sectors are similar to a subsidy for knowledge generation and the benefits of knowledge spillovers to other sectors are not retained by the innovator that generates the knowledge. Similarly, higher knowledge intensity sectors benefit other sectors by a greater amount and receive lower spillovers. The size of the market failure is the sum of these two components so higher knowledge intensity sectors have a greater overall size of market failure. As a result, high knowledge intensity sectors require greater support for innovation activities than low knowledge intensity sectors, because there is a greater market failure of innovation in these sectors.

In absence of government support these relative differences can be described as diminishing returns to relative scale. That is, sectors with relatively higher knowledge intensity have lower returns to investment in innovation while sectors with relatively lower knowledge intensity have greater returns to investment in innovation. The effect of diminishing returns to relative scale decreases over time for higher knowledge intensity sectors because of lower growth rates while the effect increases for lower knowledge intensity sectors because of higher growth rates.

The additional market failure of innovation is equal to the quality improvement directly achieved by the firm's investment in innovation, less the indirect quality improvement achieved through knowledge spillovers, plus the indirect quality improvement the firm provides to other varieties and sectors. The direct quality improvement due to the firm investing in innovation is the quality improvement specifically due to its decision to invest. It is the quality improvement if they invest in R&D $\Delta_t A_{i,j,t}$, less the quality improvement if they invest in innovation at the minimum fixed cost $(\Delta_t A_{i,j,t} - \big(\lambda_V \overline{A}_{i\forall(n_i+\widetilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1}\big))$. The indirect quality improvement the firm provides to other varieties is the quality improvement due to knowledge spillovers $(\frac{\lambda_V}{1+r}\Delta_t A_{i,j,t} + \frac{\lambda_M}{1+r}\Delta_t A_{i,j,t})$. Indirect spillovers firms are not realised until the subsequent period, so are discounted due to the rate of time preference. Hence, the additional market failure of innovation is given by:

$$\Delta_t A_{i,j,t} - \big(\lambda_V \overline{A}_{i\forall(n_i+\widetilde{n}_i),t-1} + \lambda_M \overline{A}_{i\forall M,t-1}\big) + \frac{\lambda_V}{1+r}\Delta_t A_{i,j,t} + \frac{\lambda_M}{1+r}\Delta_t A_{i,j,t} \qquad (5.20)$$

for each variety i, j, where the actual quality improvement achieved is $\Delta_t A_{i,j,t}$. By including related technology spillovers, this is the additional market failure of innovation representing the quality improvement across the economy that is achieved in all varieties where the benefits of the quality improvement do not accrue to the original innovator.

Sectors with a high level of knowledge intensity achieve quality improvements using a lower share of knowledge spillovers from related technology sectors and also provide a higher level of knowledge spillovers to other sectors. Therefore, this additional market failure of innovation is greater for high knowledge intensity sectors than low knowledge intensity sectors. In addition, the size of the actual benefits received from knowledge spillovers may be limited by the spatial distribution of innovation. In the equal distribution or peripheral cluster outcome, some quality improvement is lost in its transfer between locations. This has a greater impact on a peripheral location than on an agglomerated location.

5.4.2 Policy implications for innovation

This section considers how governments in different locations should respond with policy initiatives to support growth. The model describes how the benefits of innovation are greater than the quality improvements achieved by the firm that develops the innovation. This is the additional market failure of innovation from including related technology spillovers. Knowledge spillovers to technologically related varieties and less related varieties in the wider manufacturing sector lead to quality improvements in other varieties and sectors. Investors in innovation are not able to retain the entire benefit from their investment. Therefore, investment in innovation is less than the socially optimal level and there is a case of market failure in innovation. There is a need for policies or initiatives which increase investment in innovation and raise quality improvements. Typical policies include targeted R&D subsidies or R&D tax credits.

There are two key factors which lead to the policy implications for peripheral regions. Firstly, higher knowledge intensity sectors require greater support for R&D than low knowledge intensity sectors, because they have a greater additional market failure of innovation. Since only sectors with a knowledge intensity above the threshold can produce sustainably in the periphery, the sectors that are producing in the periphery require greater support for R&D than sectors producing in the agglomerated location. Secondly, sectors located in the periphery receive lower knowledge spillovers due to spatial externalities in knowledge spillovers, further suggesting additional support for innovation in the periphery.

With policies to support innovation, it is possible that some sectors which would otherwise switch to the agglomerated location, will now be sustainable in the long run in the periphery.[2] However, sectors on the margin would switch to the agglomerated location. This suggests that governments in smaller locations should have an innovation system that is more targeted to sectors with a level of own sector knowledge intensity above a specific threshold. These are the sectors that are less likely to leave the periphery. A targeted system would suggest supporting sectors that satisfy the threshold (Equation 5.19a) without the addition of spillovers. Governments of large agglomerated locations could have less targeted innovation policies.

Reconsider firm location decisions including a subsidy (or tax credit) of Z for R&D investment in both locations, but a greater (targeted) subsidy is available in a peripheral location. The requirement for a firm to remain located in the home region in the steady state becomes $\tilde{F} - \tilde{Z} \geq F - Z$. Substituting the costs of innovation in each location, the

[2]While not the focus of this chapter, funding of such policies should be done in the least perverse way. These could include income taxes (or land taxes if land was included as a factor of production), a consumption tax on the traditional good or a consumption tax on manufactured goods.

threshold rearranges to:

$$A_{i,t-1} \geq \frac{\left(\frac{ln\left(\tilde{Z}-Z\right)-2ln(\gamma)}{(\sigma-1)}+1\right)\left(\lambda_M \widetilde{\overline{A}}_{i\forall M,t-1}\right)-\lambda_M \overline{A}_{i\forall M,t-1}}{\left(1-\lambda_R-\lambda_R\frac{ln\left(\tilde{Z}-Z\right)-2ln(\gamma)}{(\sigma-1)}\right)(1+\lambda_V)}, \qquad (5.21)$$

in the same manner as Equation 5.19a was found. Peripheral economies should target innovation above a threshold that is dependent on the size of innovation subsidies in alternative locations for sectors above this intensity and on the distribution of manufacturing between locations. Note that if subsidies for knowledge intensive inovation in the periphery are greater than in foreign regions for those same sectors, this lowers the target threshold and broadens the range of knowledge intensive sectors that can now sustainably remain in the periphery.

In the model here, industries with a low own industry knowledge intensity are attracted to the additional knowledge spillovers available in an agglomerated location. Peripheral locations are therefore unable to sustain manufacturing in industries with a low own industry knowledge intensity. This is similar to the forces for agglomeration in existing models, where firms switch to the agglomerated location to take advantage of pecuniary or technical externalities.

The difference here is that the model suggests catastrophic agglomeration is not inevitable even with mobile capital and labour. Small or peripheral locations have an opportunity to retain manufacturing in industries that have a high enough own industry knowledge intensity because government action can lower the threshold level of knowledge intensity that determines in a sector is sustainable in the periphery. In particular, if governments delay implementing greater support for innovation activities and/or more targeted innovation policies, the effects could be long-lasting. Low knowledge intensive sectors may then have already switched to the agglomerated location and the location will have suffered some loss of manufacturing capability.

The model here had a very simplified version of knowledge spillovers between varieties produced by different firms. In this model, it is a one dimensional parameter with three step sizes. The greatest spillover is from another firm (or the same firm) that produced the same variety in the previous period, the second greatest spillover is from firms that are in the same sector but produce different varieties and the third spillover is from firms that produce varieties in different sectors. All are also weighted by the relative location of each firm. In reality, spillovers will not be so simply defined. Each pair of varieties will have a different spillover between them based on a proximity measure that may broadly reflect the arbitrary simplification here. However, some industries are also tied to local factor endowments such as resources, climate or even human capital that also add an additional dimension to knowledge spillovers. Therefore the relevant threshold could vary between sectors on the basis that the sources of some knowledge is actually immobile. As a result, peripheral governments could also have an opportunity to support industries that have knowledge inputs related to the industries in their region for which there is a local factor endowment determining the industry's location. For example, if the mining industry in Australia has similar knowledge inputs to innovation as other engineering fields, these engineering sectors, could have a lower threshold knowledge intensity, even if otherwise unrelated to mining. Similar conclusions could apply to industries with similar technologies to the technology providers of New Zealand's agricultural industry. This may imply a kind of economies of scope for innovation policy. The full explanation of this is beyond the scope of the thesis, but the model and discussion demonstrates how criteria might be determined for which sectors should be supported by innovation policy in peripheral locations and to what extent.

However, retaining industry in the periphery is globally diminishing for growth levels. Growth would eventually be higher if all industry agglomerated in the single location. Therefore, all locations will eventually benefit from closer integration, even if it triggers a loss of industry in the periphery. This loss of industry is reflected in skilled workers migrating to the agglomerated location and without transport costs the model doesn't result in an equity trade-off in the long run. But there is an equity trade off in the medium term, as peripheral regions have temporarily lower wages, that eventually equalise with migration. It is possible that this loss could even be greater than the eventual benefit to peripheral regions, presenting a strong barrier to peripheral locations integrating with larger economies. If the model were to be expanded to allow for sector specific skills, transport costs, migration costs and other spatial cost parameters, there is a strong case for the periphery being compensated by agglomerated locations for some loss of industry following the implementation of closer market integration. Similar cases could be made for compensating the periphery region in return for the periphery government choosing not to target knowledge intensive sectors under innovation policy. The model suggests that market integration is closely related and affected by systems of innovation and growth policy.

5.4.3 The impact of knowledge spillovers

Consider how varying the knowledge spillover parameters λ_R, λ_V and λ_M affects technology improvement and the distribution of economic activity. Increasing λ_R increases the level of knowledge transfer between locations. Economic integration which increases the ability to trade knowledge is growth-enhancing. If there is a peripheral cluster (either medium or long-term) or equal distribution steady state, firms benefit from the additional transfer of knowledge between locations which boosts all firms' abilities to improve technology. The impact is greater for a region with a smaller share of manufacturing, because a greater share of their technology improvement comes from inter-regional knowledge spillovers than for the agglomerated region. This result is consistent with results found by Baldwin et al. (2003) where knowledge spillovers are growth-enhancing.

Baldwin and Forslid (2000) found that increasing regional knowledge spillovers is stabilising for equal distribution outcomes, because it allows the equal distribution to remain a steady state for a larger range (at the lower end) of transport costs. Similarly it was found that regional knowledge spillovers are destabilising for the core-periphery outcome. Chapter 4 had a similar conclusion regarding stability with the addition that the consequences in the quality ladders model may be more catastrophic than in the product variety model because varieties switch location. Since transport costs are assumed zero in this model, stability is considered in terms of the effect on the steady state threshold of peripheral clusters, from changing each of the knowledge spillover parameters, λ. Changing λ_R has no effect on the steady state threshold as described in Equation 5.19a, so with the definition of stability used here, λ_R is neither stability enhancing or diminishing. However the threshold is affected by λ_V and λ_M.

λ_M increases the steady state threshold (Equation 5.19a). This implies that increases in λ_M are destabilising, because they could trigger a change in the steady state. Policies which increase the ability for knowledge to transfer between sectors has two effects, it makes both locations more attractive by increasing the knowledge available for technology improvement, but it has a greater effect on the larger region. Consequently, increases in λ_M reduce the relative own sector knowledge intensity of each sector, making lower knowledge intensive sectors more likely to switch to the agglomerated location.

λ_V decreases the threshold for a similar related reason. Increasing λ_V makes locating alongside other firms in their own sector more valuable. As a result, it increases the relative own sector knowledge intensity of each sector, making lower knowledge intensive sectors less likely to switch to the agglomerated location. Therefore, λ_V is stability-enhancing due to increasing the benefits from the clustering of related technology firms.

5.4.4 The model in relation to existing literature

The model presented here diverges from existing literature by considering related technology spillovers and the role of clustering for firm location and innovation decisions. There are a number of implications that are unique to this model. This section considers how the implications of this model are affected by alternate modelling techniques from other models of growth and location with knowledge externalities and what joint conclusions can be drawn from the model here in relation to the existing literature.

Stochastic innovation

If there were also a stochastic aspect to the model, such as a probability of R&D investment also developing a new variety or quality improvement in an alternative sector, in addition to the expected quality improvement in the firm's own variety, this could lead to the emergence of new peripheral clusters and the constant shifting of new peripheral clusters between peripheral locations and agglomerated locations. This is an example of using the modelling techniques in Duranton (2007) or Brezis and Krugman (1997) as an additional extension to the model presented here.

The stochastic emergence of alternative or replacement varieties, even in peripheral locations, can be thought of as the historical events that emerge prior to the model described here as well as an ongoing churn of industry as in the original models. Therefore the results of this hybrid model can be implied by the results of the three models. As with Duranton (2007) or Brezis and Krugman (1997) it could be expected that there will be switching of industry between locations but this switching is now partially endogenous switching from peripheral to core locations and partially stochastic churning of industry between locations. The model here is consistent with these papers, but provides an additional richness of an endogenous sorting of industries between peripheral and core locations due to technical externalities. Both of these stochastic models (Brezis and Krugman, 1997; Duranton, 2007) explain the rise and fall of locations through the stochastic emergence of new technologies in new locations, but fail to explain why a peripheral location might not be an optimal choice for some industries.

Combining Duranton (2007) or Brezis and Krugman (1997) with the model presented here, is expected to suggest that new industries are most likely to emerge in already agglomerated locations, but peripheral clusters will remain part of the economic landscape, developing new peripheral clusters, but at a lower frequency than core locations. Of the industries that emerge in the periphery, only the sectors with a level of own sector technology intensity greater than the relevant threshold can remain sustainable in the medium or long term.

But this does not mean that knowledge sectors typically cannot remain in the periphery indefinitely. Sectors with a high enough own industry knowledge intensity will never switch to the agglomerated location due to technical externalities in R&D. These sectors are dependent on knowledge spillovers from within their own firm or sector such that firms in these sectors face a high opportunity cost of switching location. As a result, these sectors are sustainable in a peripheral location.

New Economic Geography

If transport costs were included in the model here, similar to Krugman (1991b) or Chapters 3 and 4 then firms would also balance the market-access effect, the cost-of-living effect and the market-crowding effect with the location effects for clustering and agglomeration in this model. Firm allocation between regions would follow similar characteristics to the core-periphery model such as the break and sustain points that can be observed in the tomahawk diagrams in Figures 2.2 and 4.1. Cost-of-living and market-access effects would reinforce the agglomeration effect. Whether the market crowding effect reinforces or diminishes the agglomeration effect depends upon the overall distribution of economic activity. Similarly whether these three effects reinforce or diminish the clustering effect depends upon the distribution of economic activity in a firm's own sector as well as the distribution of all other economic activity. Notably, it is not expected that transport costs would have an effect on the knowledge intensity threshold in Equation 5.19a. Instead, cost-of-living, market-access and home market effects determine whether a sector is clustered or dispersed as in the core-periphery model while the knowledge intensity threshold determines whether a cluster is sustainable in a peripheral location.

5.4.5 Dynamics of sector migration

In the simulations above, the focus was on the steady state or long run growth rate. This section briefly comments on the dynamics of migration which in the model only occurs in unsteady states, but leads to the steady states already discussed.

In unsteady states, migration barriers and labour market frictions (represented by Equations 5.4a and 5.4b) would mean that when a firm (and the rest of the sector) decides to switch location, wages would rise in its new location and fall in its former location in order to clear the current labour market. Higher wages also push up the costs of innovation and production for all firms in the agglomerated location and lower the costs of innovation and production for firms in the periphery. As a result, it may be that the factor cost advantages of remaining in the peripheral location outweigh technical externalities in R&D for some sectors in the medium-term. Therefore, it is important to think about the dynamics of the transition to the steady state.

Starting in an unsteady state, one or more sectors do not meet the threshold given by Equation 5.19a. Sectors with an own industry knowledge intensity below the threshold will eventually switch to the agglomerated location. Each firm in these sectors has an incentive to switch regions. Some firms will want to switch, pushing up wages in the agglomerated location and down in the smaller location. The differential in wages will result in migration between regions, as workers follow higher wages. The size of the differential depends on the costs and benefits to the marginal sector in each region. When a firm switches location, it must offer higher wages than would otherwise be offered in the region if the firm had not switched, in order to attract workers away from incumbent firms and from the other location in subsequent periods.

With some firms switching location, the relevant threshold is recalculated each period, for the new distribution of industry and wages such that it may push the threshold for some sectors under their current own sector knowledge intensity. This means that not all sectors in the periphery will switch location immediately, but if a sector's own industry knowledge intensity falls below the threshold level, its firms will switch in future periods as migration eases the wage differential. Switching will start with the sectors that have the lowest knowledge intensity, as these are not able to overcome the threshold even with a lower wage in the periphery, because they benefit the most from spillovers from other sectors. Eventually every firm in the sector will have switched to the agglomerated location and enough workers will have migrated to equalise wages between locations. This will mean other sectors that were initially above the threshold, but only due to the differences in wages between locations, will now start to switch location. If so, these sectors will also switch between locations and workers will continue to migrate to the larger region, eventually equalising wages. This process will continue until only firms that meet the threshold (calculated under the new distribution of industry) remain in the smaller location. This is now the long-run steady state as simulated above.

While not the purpose of this chapter, future research could consider how industry location is affected in the medium-term by migration costs given the types of knowledge externalities that are explored here. For example, $s(1-s)$ in the migration Equation 5.4b could be replaced with $\delta L_{K,t}$ such that δ represents migration barriers. Varying these barriers to migration is similar to the different migration costs and requirements among countries. Reducing this barrier to migration means workers and firms can shift more quickly. With high barriers to migration, the small region will take a long time to transition to the long-run steady state. Greater barriers to migration results in a greater wage differential between locations. This allows some firms who take longer to switch location to benefit from lower wages by remaining in the smaller region. Firms in the larger region are penalised by higher wages in the short term and greater costs of innovation.

Alternatively, low barriers to migration will result in a faster transition to the steady state. Firms will switch quickly, encouraging workers to switch location and firms do not benefit as much from remaining in the periphery, because wage differentials are quickly mitigated by migration. This suggests that economic integration that reduces the barriers to migration could result in catastrophic agglomeration of industries with low own industry knowledge intensity. On the other hand, barriers to migration slow the transition process, meaning as sectors end up split across locations, there is a lower growth due to partially localised knowledge spillovers. The transition to the long-run steady state causes some cost on economic growth at the expense of firms in the agglomeration and skilled workers in the smaller region. While there may be some pain in the transfer of economic activity from the smaller region to the larger region, the overall cost is decreased by reducing barriers to migration.

5.5 Conclusion

This chapter developed and analysed an endogenous growth model where firms choose a variety, location and investment in innovation to develop quality improvements. With a quality ladders approach to innovation and a proximity approach to knowledge spillovers, a model is obtained with innovation clusters in which catastrophic agglomeration is no longer inevitable. Firms consider the technology levels and location of other firms when choosing their own location. The model allows unequal distributions of economic activity between locations even with mobile skilled labour.

Clusters of technologically related firms are sustainable in small or distant locations if they have a sufficient level of own sector knowledge intensity. Own sector knowledge intensity refers to the relative technology level of a firm's own sector as a source of knowledge for developing quality improvements compared to the technology levels of other sectors. A firm in a sector with a high level of own sector knowledge intensity sources a greater share of knowledge used for innovation activities from within its own sector. As a result, peripheral locations will be more specialised with clusters of firms that are in sectors with higher own sector knowledge intensities. On the other hand, agglomerated locations will be more diversified with a greater share of sectors with lower own sector knowledge intensity, because these sectors are more mobile between locations.

Future research should consider how "own sector knowledge intensity" can be assessed in an empirical model. Such a study would provide a more practical definition of knowledge intensity for policy-makers in peripheral locations to better understand the distribution of industry between locations. Innovation policy could more accurately target appropriate industries in peripheral locations.

The model helps explain lower growth rates in peripheral locations due to spatial externalities and industry selection effects. The key result of the model is that peripheral locations require greater and more targeted support for innovation than agglomerated locations. Innovation systems should be tailored to a region's unique economic geography that recognises the innovation systems of other locations, the distribution of economic activity and the own sector knowledge intensity of all industries.

This new approach using quality ladders, multiple industries, clusters and a proximity approach to knowledge spillovers provides a richer model of endogenous growth and the opportunity to consider a wider array of policy implications for small or distant locations. Peripheral locations require greater support for innovation than agglomerated locations, because they do not receive the full benefits of knowledge spillovers between sectors (and locations). In addition, innovation systems in peripheral locations should be more targeted to sectors above a specific threshold for own industry knowledge intensity. Agglomerated locations are more diversified and require a less targeted innovation system. Taking account of technical externalities in R&D now suggests a more tailored approach to innovation systems and economic growth policy.

Chapter 6

Discretely Innovating

This chapter introduces entry barriers to a quality ladders growth model with many separate market sectors such that entry via incremental innovation is only possible for a discrete number of firms in each sector. As a result, firms invest strategically in innovation to block the marginal entrant by leaving enough space in the market for just less than one firm. With homogeneous workers, those workers who are entrepreneurs are paid out all profits and can earn higher wages than workers employed in manufacturing, because there is not enough space in the market for an additional discrete profitable entrant. Overall, industries with few competitors have lower levels of investment in innovation and greater disparity between the wages of manufacturing workers and entrepreneurs. The model implies that countries and regions with economies characterised by small, isolated or remote markets will also have low innovation and growth, strengthening the need for more supportive innovation policy and strong competition law. Similarly, industries characterised by incremental innovations (rather than disruptive innovations) as modelled here, will sustain incumbent firms with low rates of innovation and growth. With increased competition, investment in innovation increases as the space in the market for less than one firm becomes a smaller share of the sector. The model offers further theoretical explanation for part of the inverted U relationship between competition and innovation with implications for policy particularly in small or peripheral countries.

6.1 Introduction

The industrial organisation literature has long considered the incentives for investing in innovation, particularly examining the role of rivals (both actual and potential)[1]. These microfoundations are fundamental to general equilibrium models of long-run growth. When

[1] Examples of industrial organisation literature on the incentives of firms to invest in innovation include but are not limited to Gilbert and Newbery (1982); Fudenberg et al. (1983); Reinganum (1983); Salant (1984); Harris and Vickers (1985); Vickers (1985); Reinganum (1989).

making the leap to general equilibrium models such as the endogenous growth literature, academics often rely on the role of potential rivals by assuming the market is contestable, free and continuous. However, it seems that growth theorists may be ignoring the strategic nature of competition at the micro level and therefore ignoring the possibility that the microfoundations of competition have wider economic implications on growth and productivity.

Market characteristics may be such that many incremental innovations simply cannot be implemented by a new entrant. Entrants join the market stochastically, because their only opportunity to enter is with significant innovations that change the entire landscape of the sector. Evidently, this inability for entrants to implement incremental innovations creates a market dynamic that assists incumbent firms to retain dominant market positions. This chapter investigates this incremental innovation process such that market entry is only available to discrete firms, allowing discrete incumbent firms to restrict innovation with incremental innovations to discourage entry of the marginal firm. As a result, innovation is lower than in continuous free entry models. These results are therefore extended to the general equilibrium in order to consider the effects on economic growth from discrete competition in individual sectors.

For industries where large stochastic innovations are rare, a few firms are able to retain dominant market positions and deter potential entrants, despite refraining from investing in innovation or competing aggressively on price. For example, in both Australia and New Zealand, retail supermarkets are dominated by only two firms. Despite some cost-saving innovations in the form of self-checkouts or online shopping, these innovations were not brought about by new entrants and have appeared slower than similar innovations in many other industries or countries. Similarly, a few airlines typically dominate each geographical market around the world, and they are able to deter entrants with little improvement to quality (or reduction in costs), up until the entry of "low cost carriers" in recent years. Both of these industries have high set-up costs, some network effects and perhaps they are industries where firms' reputations have a significant role. These may be characteristics which contribute to preventing entrants from developing incremental innovations and perpetuate the dominant market positions of incumbents.

Potential rivals stimulate incumbents to invest in innovation in order to retain their market share (Reinganum, 1983; Aghion et al., 2009). Capacity building allows market participants to alter conditions in a way that the outcome is to their advantage (Dixit, 1980). It is now standard in industrial organisation models that firms make price and output decisions based on their best response to competitors and rivals. Extending this idea to a general equilibrium application such as innovation in the endogenous growth literature seems a logical area to explore. In particular, it provides an additional theoretical explanation for at least part of the inverted U-relationship between competition and innovation (Aghion et al., 2005) with clear implications for competition and innovation policy. Most notably, the conclusions drawn here point to a more tailored and localised approach to competition and innovation policy based on country and market-specific characteristics.

Economic growth theory often assumes that innovation is possible for any potential entrant, but the actual market outcome may be limited to only a few firms. A barrier to entry via innovation could be justified on the basis that incumbent firms typically have a technological advantage for innovation activity in their own product, because they already have access to the product and sector-specific tacit knowledge that is not available to an entrant. A direct way to achieve this feature in a theoretical model could be to have a higher knowledge input to innovation for an incumbent firm, owing to firm specific tacit knowledge, but determining this knowledge and cost difference is rather artificial and arbitrary. Alternatively, limiting the market to a discrete number of firms is a more structural microfoundations approach that allows the model to consider the impact of different levels and modes of competition on innovation and consequently on economic growth. It is possible that the type of entry barriers described are more common in small or isolated markets, where competition from imported goods is relatively more expensive, a small market has only the capacity for a few firms and cultural barriers may make entry by foreign firms difficult. The model in this chapter is related to the findings of the models in Chapters 4 and 5 because it also provides implies policy implications for innovation and economic growth in small, isolated or peripheral regions and countries.

The simplest of entry barriers is assumed: firms can enter a market if there is enough space for an entire profitable firm. Intuitively, an entrepreneur's investment in innovation is only justified if they can develop a variety with a quality level such that an entire firm can profitably compete. Knowing this, firms that choose to invest in innovation, only need to invest in a way that is the best response to other participating firms, ignoring the marginal firm since the potential rival would not be profitable. As a result, firms make strategic decisions based on the actions of actual rivals, and not potential ones. There are no assumptions about markets being served by a continuum of firms or a portion of a firm. In the model here, each industrial sector has only a discrete number of firms who strategically use the discrete nature of market entry in making their decisions about investing in innovation.

The discrete entry assumption is relatively straightforward and common to the industrial organisation literature. For example, Reiss and Spiller (1989) and Berry (1992) both examine price competition in airline markets where the incumbent firms price strategically to deter discrete entrants. Barriers to entry have long been studied in the industrial organisation literature and typically rely on discrete entry (Bain, 1956). Models with discrete entry are also not completely new to macroeconomics or international trade.[2] For example, Eaton et al. (2013) connect the microeconomic characteristic of discrete entry within an international trade model to explain international trade patterns. Similarly, discrete entry is a characteristic of trade models that investigate capacity building prior to and during market integration. To the best of the author's knowledge, this is the first model where discrete entry has been applied to aggregate innovation or economic growth theory. However, the approach is well founded in the industrial organisation literature. Extending it to macroeconomic growth based on the same microfoundations is therefore an intuitive and relevant step that yields interesting and elegant insights.

Section 6.2 fully specifies the model including consumer preferences, technology, supply of labour and discrete market entry. Subsequently, Section 6.3 derives expressions for equilibrium prices, wages and production, steady state rates of innovation in each sector and economy-wide growth rates. Section 6.4 describes a series of simulations as a numerical exercise to further examine the effects of competition and discrete entry on innovation rates and economy-wide growth rates. Section 6.5 reconciles the model's findings with other literature and discusses the implications for markets affected by discrete entry and the characteristics of regions, countries and markets which could typically face these types of competition issues with clear implications for policy makers. Lastly, Section 6.6 concludes the chapter.

6.2 Model specification

Consider an economy with many sectors where discrete firms invest in R&D in order to produce quality improvements for entry and/or to maintain their sector position producing a discrete variety in one sector. Only discrete firms compete in each sector, such that firms take into account their own impact on the sector index of price and quality responding to the perceived elasticity of substitution. Market entry at a quality-leading level is only available to a discrete number of firms such that potential rivals facilitate strategic responses from market participants only in the sense that the potential rival may replace a marginal rival. In practice, the opportunity to replace an existing firm is not available, because existing firms can undertake a large enough incremental innovation to deter a potential marginal rival. Firms maximise profits by investing as a best response to rivals. Firms are able to charge higher prices, because competition is restricted to only a few discrete firms. Firms are able to invest strategically such that their investment in research is only in response to

[2] The idea for discrete entry in relation to growth was developed following a discussion with Prof Sam Kortum who used a trade model with an integer number of firms to explain the zeros commonly found in trade data while maintaining other desirable model characteristics (Eaton et al., 2013), even though typical trade models would predict at least some minuscule level of trade with every country.

the actual number of rivals (whether existing or potential) and not in response to marginal rivals, since the marginal firm cannot profitably enter.

For simplification, labour is the only factor of production and each sector has its own specialised labour supply. This provides an opportunity, similar to Krugman's (1982) study of comparative advantage, to calibrate the supply of labour in each sector for a wide range of competition levels in order to examine the impact of competition and discrete entry on innovation activity. Consider an economy populated by L consumers/workers. Workers differ by their sector-specific skills and supply their labour inelastically to firms in their own sector only, working in either manufacturing or R&D. Workers in R&D are developing quality improvements to enter the market each period so the term R&D workers is used interchangeably with entrepreneurs, although entrepreneur more correctly refers to the initial R&D worker who makes the decision to invest in innovation. It is the entrepreneur who decides whether more R&D workers are to be employed in order to increase the firm's quality target, although it requires sharing profits with these additional R&D workers. Entrepreneurs face a tradeoff between sharing profits with more R&D workers or allowing more rival firms to profitably enter the market. R&D workers may earn greater wages than manufacturing workers as discrete entry creates a barrier to workers participating in the labour market for R&D workers, with manufacturing wages falling until the sector's labour market clears.

All consumers are assumed to have the same preferences, even if workers have different wages. The representative consumer has a taste for varieties produced in many different sectors and a taste for variety in each sector.

6.2.1 Preferences

As in Chapter 5 the quality ladders approach of Grossman and Helpman (1991b) is extended by using a CES utility function in each sector nested in an upper level Cobb-Douglas function to allow for multiple industrial sectors of several varieties each. Unlike the previous chapters since the model considers a single region only, there is no need to include a trandtional goods sector to facilitate trade. Consumers have the following intertemporal preferences:

$$U = \sum_{t=0}^{\infty} \alpha^t \ln Q_t, \quad \alpha = 1/(1+\rho), \tag{6.1}$$

where ρ represents the consumers rate of time preference and Q_t is Cobb-Douglas consumption of manufactured goods from N sectors in period t:

$$Q_t = \prod_{i=1}^{N} c_{i,t}^{\frac{1}{N}}. \tag{6.1a}$$

Varieties in alternate sectors are neither complements nor substitutes (an elasticity of one from Cobb-Douglas utility in Equation 6.1a) such that each sector has a constant expenditure share of $1/N$. Consumers have a constant elasticity of substitution between varieties in each sector:

$$c_{i,t} = \left[\sum_{j \in i} (A_{i,j,t} c_{i,j,t})^{\frac{\sigma_i - 1}{\sigma_i}} \right]^{\frac{\sigma_i}{\sigma_i - 1}}, \quad \sigma_i > 1 \forall i = 1, ..., N. \tag{6.1b}$$

The elasticity of substitution between varieties in the same sector is σ_i for industry i. Each sector i has a fixed number of firms determined by consumer preferences (σ_i), the supply of specialised labour to the sector (L_i) and the criterion that only a discrete number of firms may enter each market. Each variety j is produced by a single firm and there are no economies of scope. For simplicity, variety j in sector i is referred to as variety i, j.

Intertemporal utility optimisation implies the transversality condition and Euler equation $\frac{E_{t+1}}{E_t} = \frac{1+r}{1+\rho}$, where E_t is expenditure in period t, ρ is the rate of time preference and r is the rate of return on savings between period t and $t+1$. Rearranging gives:

$$\frac{E_{t+1}}{1+r} = \frac{E_t}{1+\rho} = \alpha E_t. \tag{6.2}$$

Expenditure is normalised to $E_t = 1 \forall t$. The subscript t is suppressed where the time dimension is the same for all variables.

6.2.2 Technology

As in the previous two chapters growth is introduced using the quality ladders approach of Young (1998). This model without scale effects is used such that smaller or larger sectors, lower or higher quality sectors, holding all else constant, all have the same rate of quality improvement. Differences in rates of innovation arise because of differences in the nature and level of competition, the labour supply and consumer preferences. This form of growth allows the model to exclusively examine the impact of competition and discrete entry upon innovation activity.

Each variety is produced by the quality leading firm that developed a quality improvement for that variety in the previous period. In each period, the quality leader produces variety i, j as a monopolist and, if it is going to enter in the following period, the firm conducts R&D to ensure a quality improvement large enough to maintain its niche monopoly position. If the firm's quality improvement is too small such that there is enough space in the market for a profitable discrete firm, a potential entrant could develop a quality improvement for a variety in sector i reducing profits and prices with additional competition. This is the partially contestable nature of discrete entry.

In the deterministic model here, the firm knows a quality improvement will occur at the end of each period and that it is required to enter in the coming period. So, for the purposes of the economic model, it does not matter whether the quality improvement is developed by an incumbent or entrant. Alternatively, Duranton (2007) examined industry churn using a stochastic quality ladders model, where the firm chooses a probability of achieving a fixed level of quality improvement. An incumbent never innovates in their own variety because of creative destruction; the firm does not want to destroy the value of their existing patent by developing a quality improvement. In another alternative stochastic approach a firm could choose the size of innovation and a corresponding probability of achieving the targeted innovation, where size of innovation and probability of innovating are inversely related to one another. Reinganum (1983) reconciles these stochastic and deterministic approaches by suggesting that the deterministic model is most appropriate for incremental innovations where the incumbent can minimise risk by aiming for incremental innovations with a high probability. For example, it is common in many industries for firms to add minor new features to new generations of otherwise similar models. Industry churn would occur on the rare occasion that an entrant achieved a drastic innovation to enter the market with a substantial quality improvement. Industry churn was the focus of Duranton (2007), so it is not readdressed in the model here. For simplicity, the model maintains the deterministic approach for modelling elegance, following the incremental innovation technology path.

As in the previous two chapters, the labour requirement as an input to R&D in the previous period $t-1$ to achieve the targeted quality level $A_{i,j,t}$ for variety i,j and the fixed cost of production in period t is given by:

$$F_{i,j,t-1}(A_{i,j,t},\bar{A}_{i,j,t-1}) = \begin{cases} \gamma e^{\eta A_{i,j,t}/\overline{A}_{i,j,t-1}} & \text{if } A_{i,j,t} \geqslant \overline{A}_{i,j,t-1} \\ \gamma e^{\eta} & \text{otherwise} \end{cases}. \tag{6.3}$$

The parameters γ and η are constants used for calibration. $\overline{A}_{i,j,t-1}$ is an index of technological opportunity for variety i,j, representing the intertemporal spillover of knowledge available to variety i,j researchers. In contrast to previous chapters the index of technological opportunity used here is from Young (1998) and is simply the highest existing quality level for variety i,j. Each sector has symmetric firms such that all varieties in a single sector have the same quality level. The impact of knowledge is not the focus of this chapter so simplicity is maintained by avoiding additional inter-firm knowledge linkages or other sources of knowledge that were already discussed extensively in Chapters 4 and 5. The fixed cost can be thought of as two components: a standard fixed cost of γe^{η} irrespective of quality improvement and a research cost of $\gamma e^{\eta A_{i,t}/\bar{A}_{i,j,t-1}} - \gamma e^{\eta}$. Following their fixed investment in period $t-1$, firms may produce any quantity of their product in period t at a constant marginal (labour) cost of $w_i\beta$. The parameter β is the marginal cost of production and is assumed constant for all varieties in sector i. For simplicity and to make a fair comparison between sectors, the parameters γ, η and β are assumed to be the same for every industry and use only the supply of specialised labour is used to vary competition in different industries.

6.2.3 Labour

Following the approach of Krugman (1982) each industrial sector is assumed to have an exclusive labour supply where workers are immobile between sectors, but those workers are mobile between firms (varieties) within their sector and mobile between employment in R&D or manufacturing. This assumption implies workers' skills are industry-specific and it is difficult for workers to re-skill for employment in other sectors. All workers inelastically provide one unit of labour per period in their own specialised sectors.

In each industrial sector, the labour supply is written as

$$L_{i,t} = \sum_{j \epsilon n_{i,t}, n_{i,t+1}} l_{i,j,t} \quad \forall i = 1, ..., N, \tag{6.4}$$

where $l_{i,j,t}$ is the labour employed in R&D and manufacturing in period t by the firm producing variety i, j and $n_{i,t}$ and $(n_{i,t+1})$ are the discrete number of firms in that sector in periods t and $t+1$ respectively. Both the number of varieties in sector i in period t $(n_{i,t})$ and period $t+1$ $(n_{i,t+1})$ are required since firms employ workers in R&D to produce quality improvements for the following period and workers in manufacturing to produce current versions. A single firm producing variety i, j could employ only manufacturing workers (withdrawing from the market in period $t+1$), only R&D workers (in order to enter the market in period $t+1$) or both (in order to produce in both period t and $t+1$).

Since the focus of this chapter is on the relationship between growth/innovation and competition it is assumed that the wages of workers in a particular industry are not affected by the market power of employers in the employee-employer relationship. Of course, it is entirely possible that workers who have very specific skills in an industry dominated by only a few employers could face non-competitive wages (Bhaskar et al., 2002), but monopsony or oligopsony pressure on the labour market is beyond the scope of this chapter. However, wages are affected by the nature of competition in product markets as this determines the revenue to the firm and the productivity of the marginal worker. The calibration of each specialised labour supply can be used to alter competition between sectors and compare the level of innovation in different industries. Industries with a large labour supply will have a greater number of firms competing. This allows a direct comparison of the level of innovation and competition, without altering demand parameters such as elasticities.

Production involves a fixed labour cost $F_{i,j,t-1}$ in the period prior to production and a constant marginal cost of β. For all sectors $i = 1, 2, 3, ..., N$ and varieties $j = 1, 2, 3, ..., n_i$, the labour required by each firm in period t is given by:

$$\begin{aligned} l_{i,j,t} &= 0 & if\, c_{i,j,t} = 0\, and\, c_{i,j,t+1} = 0 \\ l_{i,j,t} &= \beta c_{i,j,t} & if\, c_{i,j,t} > 0\, and\, c_{i,j,t+1} = 0 \\ l_{i,j,t} &= F_{i,j,t} & if\, c_{i,j,t} = 0\, and\, c_{i,j,t+1} > 0 \\ l_{i,j,t} &= F_{i,j,t} + \beta c_{i,j,t} & if\, c_{i,j,t} > 0\, and\, c_{i,j,t+1} > 0. \end{aligned} \tag{6.4a}$$

where $F_{i,j,t}$ is the fixed cost of developing a quality improvement for production in period $t+1$ and $c_{i,j,t}$ is the period t production of variety i, j at its existing quality level that was developed in the previous period.

6.2.4 Market entry

The number of sectors is assumed to be fixed such that there are always N sectors, but firms (producing a single variety each) can enter each sector. The entry criteria of typical growth models is modified such that free entry is not continuous (or that the number of firms is assumed to be large). In this model, free entry is only possible for a discrete number of profitable firms. That is, an additional firm can only enter the market if there is enough space for an additional discrete firm to profitably enter. As a result, each industrial sector has a discrete number of firms and participating firms have a strategic incentive to underinvest in innovation by making strategic decisions in response to actual rivals, rather than potential rivals who are unable to profitably enter the market.

Typically general equilibrium models using Dixit-Stiglitz competition assume either the number of firms is continuous or that the number of firms is large enough that it approximates a continuous product space. But, with disaggregating the market into many sectors, the assumption of entry for only a discrete number of firms is much more realistic. If there is not enough space in the market for a whole profitable firm, they are unable to enter the market and produce a variety in this industry. The market does not reflect production by some portion of a firm. In particular, by disaggregating manufacturing into many industrial sectors, there will only ever be a few competitors in each industry. As a result, participating firms can innovate strategically to only respond to the competitive pressures of actual rivals, rather than potential rivals, limiting fixed costs for market participants and increasing wages of entrepreneurs. If participants fail to innovate by enough such that a marginal firm entered the market it is possible that the new competitive position would result in lower profits for participating firms and lower wages for entrepreneurs. If participants then attempted to innovate at the profit-maximising level in future periods, R&D workers would find themselves with lower wages than manufacturing workers, switching to supplying their labour as a manufacturing worker and at least one firm would exit the market. This characteristic can be thought of as the market having partially free entry, in that contestability and free entry are still characteristics, but subject to strategic investment by discrete participating firms. As described above, stochastic innovation could allow for industry churn (Duranton, 2007), but sectors would be still be dominated by a few firms. To maintain a monopoly in their own niche variety, firms can invest strategically by investing in response to actual rivals rather than potential marginal competitors. Partially free entry and strategic innovation (i.e. discrete entry) allows participating firms to have higher prices than in typical continuous free entry models of monopolistic competition.

To facilitate the potential for new entrant firms in new varieties, a comparable innovation cost is assumed. If there is space in the market, it is assumed that the new entrant creates a new variety, where the knowledge input to innovation is an average of quality levels for its particular industry. This maintains symmetry in each industrial sector even if a new variety is introduced. The index of technological opportunity for a firm with a new variety is given by $\overline{A}_{i,j,t-1} = \frac{\sum_{j \in n_{i,t-1}} A_{i,j,t-1}}{n_{i,t-1}}$. In the model, the only practical opportunity for a new variety is an exogenous shock, such as the removal of trade barriers or a change in the elasticity of substitution, that creates enough space for a new variety. Alternatively, variety i, j could be replaced by variety i, k in a stochastic model such as Duranton (2007), but this is essentially treated as exogenous by participating firms. Since the focus of this chapter is on the impact of competition upon innovation activity, simplicity and elegance is maintained by using the deterministic approach of Young (1998).

6.3 Equilibrium and the steady state

Equilibrium wages, prices and production follow from typical optimisation of Cobb-Douglas and nested CES utility. Firms choose prices or the quantity supplied to the market and a level of investment in quality-improving innovations based on the consumers taste for quality and the expected actions of other participating firms. Usually in CES models, monopolistic competition means both Cournot and Bertrand competition yield the same results, because the number of firms is assumed very large or continuous. However, in the discrete model here with only a few firms in each market, optimisation requires that firms take note of their perceived elasticity of substitution and competitive outcomes therefore also depend upon the form of competition. Therefore the analysis considers equilibrium outcomes under both Cournot and Bertrand competition.

Each firm's profit-maximising investment in innovation is determined. This requires a new approach where the firm's innovation decisions are based on the nature of competition in its own sector. In Cournot competition firms maximise profits by choosing output, given consumers' willingness to pay in order to clear the market. Alternatively, Bertrand competition means firms maximise profits by choosing prices, given consumers' quantity demanded at each price. These approaches are extended to also determine the profit maximising level of quality improvement for both types of competitive equilibria. If the profit maximising output is determined by Cournot competition, it is assumed that firms choose a level of quality improvement given consumers' willingness to pay for higher quality products. Alternatively, if the profit maximising price is determined by Bertrand competition, firms choose a level of quality improvement given consumers' quantity demanded in response to quality improvements.

6.3.1 Demand

Consumers allocate expenditure across sectors and varieties subject to the budget constraint $\sum_{i \in N} P_i c_i \leq E$, where P_i is the price index of sector i to be defined in Equation 6.6a. By two-stage optimisation a demand function is derived for each variety. From Cobb-Douglas utility across sectors expenditure per sector is found to be:

$$c_i P_i = \frac{1}{N} E, \tag{6.5}$$

such that the consumer spends $\frac{1}{N}$ of her expenditure on each variety i. Similar to Equations 4.5 and 5.7, maximising the utility function finds that the direct demand function for each variety is given by:

$$c_{i,j} = \frac{E}{N} A_{i,j}^{\sigma_i - 1} P_{i,j}^{-\sigma_i} P_i^{\sigma_i - 1} \tag{6.6}$$

where $P_{i,j}$ and $A_{i,j}$ are the price and quality level of variety i, j respectively. P_i is an index of price and quality defined by:

$$\begin{aligned} P_i^{\sigma_i - 1} &= \left[\sum_{j \in i,} A_{i,j}^{\sigma_i - 1} P_{i,j}^{1-\sigma_i} \right]^{-1} \\ P_i &= \left[\sum_{j \in i,} A_{i,j}^{\sigma_i - 1} P_{i,j}^{1-\sigma_i} \right]^{\frac{1}{1-\sigma_i}} . \end{aligned} \tag{6.6a}$$

The derivation of this demand function and sector price index is a simple extension of the method described in Appendix A, dropping transport costs and making each manufacturing sector represent a $\frac{1}{N}$ share of expenditure.

Quantity demanded is inversely related to the price of each good and positively related to quality. Similarly, quantity demanded is inversely related to the quality of other goods within the sector and positively related to the prices of other goods, but if there are more than two firms, demand is more responsive to changes in the price and quality of the actual variety than to changes in prices and qualities of other individual varieties.

Alternatively, the inverse demand function can be found by optimising for the indirect utility function such that inverse demand in each sector is given by:

$$P_{i,j} = A_{i,j}^{\frac{\sigma_i - 1}{\sigma_i}} c_{i,j}^{\frac{-1}{\sigma_i}} \frac{E}{N} \left[c_i^{\frac{\sigma_i - 1}{\sigma_i}} \right]^{-1}, \tag{6.7}$$

where c_i is specified in consumer preferences. Consumer's willingness to pay is inversely related to the quantity supplied to the market and positively related to quality. Similarly, willingness to pay is inversely related to the quality and quantity supplied of other goods in the sector, but if there are more than two firms, consumers are again more responsive to changes in the quantity and quality of the actual variety than to changes in the quantities and qualities of other individual varieties.

In either form of competition, the relevant price index is the same for both Bertrand and Cournot competition, since it is defined by the budget constraint $\sum_{i \epsilon N} c_i P_i \leq E$. Similarly in both forms of competition varieties do not compete with goods outside their own sector due to the nature of Cobb-Douglas preferences in the upper level of the utility function.

The firm's profit function is given by:

$$\pi_{i,j,t} = \frac{(P_{i,j,t} - \beta w_{i,M,t}) \, c_{i,j,t}}{1 + r_t} - w_{i,R,t-1} F_{i,j,t-1} \left(A_{i,j,t} \right) = 0, \tag{6.8}$$

subject to the appropriate demand function. $w_{i,R,t-1}$ is the wage of entrepreneurs or R&D workers in period $t-1$ and $w_{i,M,t}$ is the wage of workers employed in manufacturing in period t in order to clear the specialised sector i labour market. All profits are paid as wages to entrepreneurs and research workers such that no remaining profit remains with the firm. Period t revenues are discounted, since investment in innovation occurs in the period prior to production.

Since research workers can potentially earn more than manufacturing workers and the entrepreneur determines the firm's strategic decision to hire more research workers, the optimisation decision is modified such that entrepreneurs attempt to maximise their own wage $w_{i,R,t-1}$. Rearranging, the firms output, wage, innovation and pricing decisions are governed by:

$$\max w_{i,R,t-1} = \frac{(P_{i,j,t} - \beta w_{i,M,t}) \, c_{i,j,t}}{(1 + r_t) \, F_{i,j,t-1} \left(A_{i,j,t} \right)}. \tag{6.9}$$

Cournot prices

In Cournot competition, the firm makes strategic choices of output in response to consumers willingness to pay. Entrepreneurs maximise their own wages given the inverse demand function in Equation 6.7. Differentiating Equation 6.9 with respect to the decision variables of output and quality, the first order conditions are given by:

$$\frac{\partial w_{i,R,t-1}}{\partial c_{i,j,t}} = \frac{P_{i,j,t} - \beta w_{i,j,t}}{(1+r_t)\, F_{i,j,t-1}} + \frac{c_{i,j,t}}{(1+r_t)\, F_{i,j,t-1}} \frac{\partial P_{i,j,t}}{\partial c_{i,j,t}} = 0 \qquad \text{and} \tag{6.10a}$$

$$\frac{\partial w_{i,R,t-1}}{\partial A_{i,j,t}} = \frac{c_{i,j,t}}{(1+r_t)} \left(\frac{\frac{\partial P_{i,j,t}}{\partial A_{i,j,t}} F_{i,j,t-1} - (P_{i,j,t} - \beta w_{i,M,t}) \frac{\partial F_{i,j,t-1}}{\partial A_{i,j,t}}}{(F_{i,j,t-1}(A_{i,j,t}))^2} \right) = 0. \tag{6.10b}$$

Rearranging Equation 6.10a gives:

$$P_{i,j,t} = \frac{1}{\left(1 + \frac{c_{i,j,t}}{P_{i,j,t}} \frac{\partial P_{i,j,t}}{\partial c_{i,j,t}}\right)} \beta w_{i,j,t} \tag{6.10c}$$

Price is a function of marginal cost and the elasticity of substitution as in typical Dixit-Stiglitz models. However, if there is only one or a few discrete firms in each sector, the firm's perceived elasticity of substitution must also account of its own effect on the price quality index within its own sector. In contrast to previous chapters, it is no longer assumed there are a enough firms for the firm producing variety i, j to not notice the effect of its own prices on the sector's price index. While the exact number of firms in each sector n_i is still to be determined, symmetry in each sector means the inverse demand function in Equation 6.7 for variety i, j can be written as:

$$P_{i,j}(n_i) = \frac{E}{N} A_{i,j}^{\frac{\sigma_i - 1}{\sigma_i}} c_{i,j}^{\frac{-1}{\sigma_i}} \left(A_{i,j}^{\frac{\sigma_i - 1}{\sigma_i}} c_{i,j}^{\frac{\sigma_i - 1}{\sigma_i}} + (n_i - 1) \left[A_{i,k}^{\frac{\sigma_i - 1}{\sigma_i}} c_{i,k}^{\frac{\sigma_i - 1}{\sigma_i}} \right] \right)^{-1}, \tag{6.10d}$$

where variety i, k represents each of the other symmetrical varieties in sector i. Differentiating with respect to consumption of variety i, j and simplifying the differential using symmetrical varieties, the equilibrium price is evaluated by substituting the differential into Equation 6.10c:

$$P_{i,j,t} = \left(\frac{n_{i,t} \sigma_i}{(n_{i,t} - 1)(\sigma_i - 1)} \right) \beta w_{i,j,t}, \qquad n_i \geq 2. \tag{6.10e}$$

This is the relationship between price and competition (by number of firms) with Cournot competition in a CES utility function with symmetry. As n_i increases, prices for a firm's own variety have less effect on the perceived elasticity of substitution until the familiar CES pricing rule is reached that is common to both Cournot and Bertrand competition $P_{i,j,t} = \beta w_{i,j,t} \left(\frac{\sigma_i}{\sigma_i - 1} \right)$.

In the case of $n_i = 1$, both Cournot and Bertand price functions break down, because firms are able to charge any price if they are monopolists with Cobb-Douglas utility between sectors (varieties). Because this chapter focusses on the effect of increasing competition on innovation, it is only necessary to focus on the $n \geq 2$ scenario. Including the $n = 1$ scenario would require the upper-level of the utility function to be an alternative such as a CES function. While this still provides elegant solutions for the relationship between price and competition even for $n_i = 1$, it makes labour market clearing less elegant, because determining wages in one sector would require including the prices and elasticities of all sectors. In the Cobb-Douglas version used here, clearing the labour market only requires coinsideration of prices in the same sector but there is no loss of generality for the conclusions regarding the effects of competition.

Bertrand prices

In Bertrand competition, entrepreneurs make strategic choices in response to the quantity consumers demand. Entrepreneurs maximise their own wages given the direct demand function in Equation 6.7. Differentiating Equation 6.9 with respect to the decision variables of price and quality, the first order conditions are given by:

$$\frac{\partial w_{i,R,t-1}}{\partial P_{i,j,t}} = \frac{c_{i,j,t}}{(1+r_t)\, F_{i,j,t-1}\left(A_{i,j,t}\right)} + \frac{\left(P_{i,j,t} - \beta w_{i,j,t}\right)}{(1+r_t)\, F_{i,j,t-1}\left(A_{i,j,t}\right)} \frac{\partial c_{i,j,t}}{\partial P_{i,j,t}} = 0, \tag{6.11a}$$

$$\frac{\partial w_{i,R,t-1}}{\partial A_{i,j,t}} = \frac{c_{i,j,t}}{(1+r_t)} \left(\frac{\frac{\partial P_{i,j,t}}{\partial A_{i,j,t}} F_{i,j,t-1} - \left(P_{i,j,t} - \beta w_{i,M,t}\right) \frac{\partial F_{i,j,t-1}}{\partial A_{i,j,t}}}{\left(F_{i,j,t-1}\left(A_{i,j,t}\right)\right)^2} \right) = 0. \tag{6.11b}$$

Rearranging Equation 6.11a gives:

$$P_{i,j,t} \left(\frac{c_{i,j,t}}{P_{i,j,t}} \frac{1}{\frac{\partial c_{i,j,t}}{\partial P_{i,j}}} + 1 \right) = \beta w_{i,j,t}. \tag{6.11c}$$

With discrete firms and symmetry in each sector, the demand function can be written as:

$$c_{i,j}\left(n_i\right) = \frac{E}{N} A_{i,j}^{\sigma_i - 1} P_{i,j}^{-\sigma_i} \left[P_{i,j}^{1-\sigma_i} A_{i,j}^{\sigma_i - 1} + \left(n_i - 1\right) P_{i,k}^{1-\sigma_i} A_{i,k}^{\sigma_i - 1} \right]^{-1}. \tag{6.11d}$$

Differentiating with respect to the price of variety i, j and simplifying the differential using symmetrical varieties, price can be evaluated by substituting the differential into Equation 6.11c:

$$P_{i,j}(n) = \left(\frac{\sigma_i - 1 - \sigma_i n_i}{n_i + \sigma_i - 1 - \sigma_i n_i} \right) \beta w_{i,j} = \left(\frac{\sigma_i - 1 - \sigma_i n_i}{\left(n_i - 1\right)\left(1 - \sigma_i\right)} \right) \beta w_{i,j}, \qquad n \geq 2. \tag{6.11e}$$

This is the relationship between price and competition (by number of firms) with Bertrand competition. As with Cournot competition, as n_i increases, firms gradually take less notice of their own effect on the price index. As the number of firms tends to infinity, the familiar CES pricing rule $P_{i,j,t} = \beta w_{i,j,t} \left(\frac{\sigma_i}{\sigma_i - 1} \right)$ is reached that was also found in earlier chapters.

6.3.2 Labour market clearing

Since β units of labour are required for each unit of production, the labour used in manufacturing in sector i simply equals the consumer expenditure in sector i divided by the symmetrical price per unit and multiplied by β:

$$l_{i,m,t} = \frac{E_t}{NP_{i,j,t}}\beta. \tag{6.12a}$$

Labour used in R&D in period t equals the number of entrants in the coming period multiplied by the investment per firm,

$$l_{i,I,t} = n_{i,t+1}F_{i,j,t}. \tag{6.12b}$$

Labour market clearing requires that the total amount of labour used in period t in manufacturing and R&D in each sector is equal to the amount of specialised sectoral labour available. Total labour in each sector is therefore equal to:

$$L_{i,t} = n_{i,t+1}F_{i,j,t} + \frac{E_t}{NP_{i,j,t}}\beta. \tag{6.12c}$$

Manufacturing labour

Substituting Cournot prices into Equation 6.12c, the labour employed in sector i, governed by Cournot innovation, is given by:

$$L_{i,t} = n_{i,t+1}F_{i,j,t} + \frac{E_t\left(n_{i,t}-1\right)\left(\sigma_i-1\right)}{Nn_{i,t}\sigma_i w_{i,M,t}}. \tag{6.13}$$

Solving for Cournot manufacturing wages gives:

$$w_{i,M,t} = \frac{E_t\left(n_{i,t}-1\right)\left(\sigma_i-1\right)}{Nn_{i,t}\sigma_i\left(L_{i,t}-n_{i,t+1}F_{i,j,t}\right)}. \tag{6.13a}$$

Alternatively substituting Bertrand prices into Equation 6.12c, the labour employed in sector i, governed by Bertrand innovation, is given by:

$$L_{i,t} = n_{i,t+1}F_{i,j,t} + \frac{E_t\left(n_{i,t}-1\right)\left(\sigma_i-1\right)}{N\left(\sigma_i n_{i,t}-\left(\sigma_i-1\right)\right)w_{i,M,t}}. \tag{6.14}$$

Rearranging, Bertrand manufacturing wages are given by:

$$w_{i,M,t} = \frac{E_t\left(n_{i,t}-1\right)\left(\sigma_i-1\right)}{N\left(\sigma_i n_{i,t}-\left(\sigma_i-1\right)\right)\left(L_{i,t}-n_{i,t+1}F_{i,j,t}\right)}. \tag{6.14a}$$

Manufacturing wages depend on demand for varieties in this sector, the number of competitors, the elasticity of substitution and which portion of labour is already occupied in the research sector.

6.3.3 Strategic innovation and endogenous variety

All profits above manufacturing costs are paid in wages to research workers, such that there is zero profit remaining with the firm. Workers are mobile between manufacturing and research and there are no barriers to entering the labour market as a manufacturing worker. However, there are barriers to working in R&D, since this is restricted to a discrete number of firms. Therefore, R&D workers will never earn less than manufacturing workers, because they would switch to the higher paid manufacturing role until research and manufacturing wages equalise. Research workers could be paid a greater wage rate than manufacturing workers, because the reverse process is only possible if there is space for an additional discrete firm. Therefore, in equilibrium, labour market clearing and free entry for discrete firms requires $w_{i,M,t} \leq w_{i,R,t}$.

Entrepreneurs can strategically keep research wages to a higher level than manufacturing workers, because participating firms (or which the entrepreneur determines it's employment of R&D or manufacturing workers) have an advantage in terms of only needing to invest enough to block the marginal firm, by leaving space for just less than one discrete firm. Entrepreneurs will hire additional research workers to maximise their own wage, but not as many as the continuous free entry model would suggest, as this will erode research wages, as profits are shared with a greater number of research workers. The level of competition is determined by the discrete number of firms that can enter the market with free entry, but once the discrete number of firms is determined, participating firms can invest strategically in response to their competitors strategic actions, ignoring the potential for additional entrants.

Just as with determining prices, Cournot and Bertrand approaches are used to determine innovation, where firms maximise profits (research wages) subject to inverse demand and direct demand respectively. It is assumed that if the firm makes output decisions in response to inverse demand (Cournot), they will also make innovation decisions in response to inverse demand. Similarly if the firm makes pricing decisions in response to the regular demand curve (Bertrand), they will also make innovation decisions in response to the regular demand curve. That is, a combined approach of Cournot prices and Bertrand innovation or Bertrand prices and Cournot innovation, is not explored. To clarify, a firm in a Cournot market optimises profits (entrepreneurial wages) by choosing a particular quantity and quality to supply to the market based on the consumers' willingness to pay for that quantity and quality level, while a firm in a Bertrand market chooses a particular price and quality based on the consumers' quantity demanded at that price and quality level. There is potential for a third competitive option where firms choose a particular price and output based on the consumers' preference for quality improvement to clear the market at that price and output, but since consumers have no way of choosing quality levels (i.e. quality improvements are not flexible by consumer decisions in order to clear the market), this alternative competitive equilibrium is considered unlikely. For the sake of future research, it is possible that the quality improvement of some varieties in the services sector (e.g. taxi services) are flexible and this competitive strategy and equilibrium could be the subject of future research.

Research labour and Cournot innovation

The first order condition was given by Equation 6.10b. Rearranging and dividing both sides by $A_{i,j,t}$ gives:

$$\frac{F_{i,j,t-1}}{A_{i,j,t}} \frac{1}{\frac{\partial F_{i,j,t-1}}{\partial A_{i,j,t}}} = \frac{(P_{i,j,t} - \beta w_{i,M,t})}{A_{i,j,t}} \frac{1}{\frac{\partial P_{i,j,t}}{\partial A_{i,j,t}}} \quad . \tag{6.15}$$

As expected entrepreneurs select a quality level where the elasticity of the research cost with respect to quality equals the elasticity of demand (in terms of willingness to pay) with respect to quality. However, with a few discrete firms, each firm responds to its perceived elasticity of substitution and actual rivals. Differentiating Equations 6.3 and 6.10d, substituting into Equation 6.15 and evaluating with Cournot prices, the firms' profit-maximising quality target is given by:

$$A_{i,j,t} = \overline{A}_{i,j,t-1} \frac{(\sigma_i - 1)(n_{i,t} - 1)}{\eta(n_{i,t} + (\sigma_i - 1))}. \tag{6.15a}$$

As in Young (1998) parameters are assumed such that there is always growth in the competitive model. That is $\frac{\sigma_i - 1}{\eta} > 1$. However, it is possible under Cournot competition that strategically responding to a few discrete rivals would result in no growth in the steady state. That is $\frac{(\sigma_i-1)(n_{i,t}-1)}{\eta(n_{i,t}+(\sigma_i-1))} \leq 1$ for some low levels of n_i. Solving for n_i, growth will not occur if $n_{i,t} \leq \frac{(\eta+1)(\sigma_i-1)}{(\sigma_i-1)-\eta}$. In these cases, participating firms employ the minimum number of research workers to ensure their continued market position and there is no quality improvement. Just by entering the market at the minimum fixed cost participating firms are able to deter the marginal entrant, even without developing a quality-improving innovation. In these cases, the quality target is given by the existing quality level $\overline{A}_{i,j,t-1}$ and there is zero growth. Substituting the quality target in Equation 6.15a into the innovation function (Equation 6.3), the number of research workers required per firm is given by the discrete function:

$$F_{i,j,t-1} = \begin{cases} \gamma e^{\frac{(\sigma_i-1)(n_{i,t}-1)}{n_{i,t}+(\sigma_i-1)}} & \text{for } n_{i,t} > \frac{(\eta+1)(\sigma_i-1)}{(\sigma_i-1)-\eta} \\ \gamma e^{\frac{1}{\eta}} & \text{otherwise.} \end{cases} \tag{6.15b}$$

Investment in innovation is positively related to the number of firms. As n_i increases, the innovation target and investment in innovation tend towards the competitive level found in Young (1998).

Substituting research cost into Equation 6.9, the wages of entrepreneurs and research workers in sectors with Cournot innovation are given by:

$$w_{i,R,t-1} = \begin{cases} \frac{(n_{i,t}+(\sigma_i-1))\alpha E_{t-1}}{N n_{i,t}^2 \sigma_i \gamma e^{\frac{(\sigma_i-1)(n_{i,t}-1)}{n_{i,t}+(\sigma_i-1)}}} & \text{for } n_{i,t} > \frac{(\eta+1)(\sigma_i-1)}{(\sigma_i-1)-\eta} \\ \frac{(n_{i,t}+(\sigma_i-1))\alpha E_{t-1}}{N n_{i,t}^2 \sigma_i \gamma e^{\eta}} & \text{otherwise.} \end{cases} \tag{6.16}$$

Research labour and Bertrand innovation

The first order condition was given by Equation 6.11b. Rearranging and dividing both sides by $A_{i,j,t}$ gives:

$$\frac{F_{i,j,t-1}\left(A_{i,j,t}\right)}{A_{i,j,t}} \frac{1}{\frac{\partial F_{i,j,t-1}}{\partial A_{i,j,t}}} = \frac{c_{i,j,t}}{A_{i,j,t}} \frac{1}{\frac{\partial c_{i,j,t}}{\partial A_{i,j,t}}}. \tag{6.17}$$

As with Cournot innovation, entrepreneurs select a quality level where the elasticity of the research cost with respect to quality equals the elasticity of demand (in terms of quantity demanded) with respect to quality. Differentiating Equations 6.3 and 6.11d, substituting into Equation 6.17 and rearranging, the firms' profit-maximising quality target is given by:

$$A_{i,j,t} = \overline{A}_{i,j,t-1} \frac{(\sigma_i - 1)(n_{i,t} - 1)}{\eta n_{i,t}}. \tag{6.17a}$$

As long as $\sigma_i > 2\eta + 1$, Bertrand competition always results in growth for any level of competition $n_i \geq 2$. For $1 < \sigma_i < 2\eta + 1$, the supply of labour must be large enough that $n_i > \frac{\sigma_i - 1}{(\sigma_i - 1) - \eta}$ for growth to occur. Substituting the quality target into Equation 6.3, the number of research workers required by the firm is given by the discrete function:

$$F_{i,j,t-1} = \begin{cases} \gamma e^{\frac{(\sigma_i-1)(n_{i,t}-1)}{n_{i,t}}} & \text{for } n_{i,t} > \frac{\sigma_i-1}{(\sigma_i-1)-\eta} \\ \gamma e^{\eta} & \text{otherwise.} \end{cases} \tag{6.17b}$$

Similar to Cournot innovation, as n_i increases, the innovation target and investment in innovation tends toward the competitive level found in Young (1998).

Substituting into Equation 6.9, the wages of entrepreneurs and research workers in markets governed by Bertrand innovation are given by:

$$w_{i,R,t-1} = \begin{cases} \frac{\alpha E_{t-1}}{N(\sigma_i n_{i,t} - (\sigma_i - 1))\gamma e^{\frac{(\sigma_i-1)(n_{i,t}-1)}{n_{i,t}}}} & \text{for } n_{i,t} > \frac{\sigma_i - 1}{(\sigma_i - 1) - \eta} \\ \\ \frac{\alpha E_{t-1}}{N(\sigma_i n_{i,t} - (\sigma_i - 1))\gamma e^{\eta}} & \text{otherwise.} \end{cases} \tag{6.18}$$

As with prices, the form of competition does not matter when the market is assumed large and contestable, but with few firms the level of innovation and the wages of entrepreneurs vary between Cournot and Bertrand competition.

Endogenous variety

As described above, free entry with discrete firms requires:

$$w_{i,M,t} \leq w_{i,R,t}. \tag{6.19}$$

Substituting Cournot manufacturing and research wages (Equations 6.13a and 6.16 respectively), the wage inequality (Equation 6.19) can be rearranged to:

$$L_{i,t} \geq \begin{cases} \left(\frac{(n_{i,t}-1)(\sigma_i-1)n_{i,t+1}^2}{n_{i,t}(n_{i,t+1}+(\sigma_i-1))\alpha} + n_{i,t+1} \right) \gamma e^{\frac{(\sigma_i-1)(n_{i,t+1}-1)}{n_{i,t+1}+(\sigma_i-1)}} & for\ n_{i,t} > \frac{(\eta+1)(\sigma_i-1)}{(\sigma_i-1)-\eta} \\ \\ \left(\frac{(n_{i,t}-1)(\sigma_i-1)n_{i,t+1}^2}{n_{i,t}(n_{i,t+1}+(\sigma_i-1))\alpha} + n_{i,t+1} \right) \gamma e^{\eta} & \text{otherwise.} \end{cases} \tag{6.20}$$

In the steady state, the number of firms is unchanging such that the largest integer $n_i \geq 2$ that satisfies

$$L_i \geq \begin{cases} \left(\frac{(n_i-1)(\sigma_i-1)n_i}{(n_i+(\sigma_i-1))\alpha} + n_i \right) \gamma e^{\frac{(\sigma_i-1)(n_i-1)}{n_i+(\sigma_i-1)}} & \text{for } n_i > \frac{(\eta+1)(\sigma_i-1)}{(\sigma_i-1)-\eta} \\ \\ \left(\frac{(n_i-1)(\sigma_i-1)n_i}{(n_i+(\sigma_i-1))\alpha} + n_i \right) \gamma e^{\eta} & \text{otherwise.} \end{cases} \tag{6.20a}$$

implicitly defines the steady state number of firms with Cournot competition. The Cournot model is now complete. The n_i which satisfies Equation 6.20a can be used to solve any variable.

Similarly, the wage inequality in Equation 6.19 can be rearranged by substituting Bertrand manufacturing and research wages (Equations 6.14a and 6.18) such that:

$$L_{i,t} \geq \begin{cases} \left(\frac{(n_{i,t}-1)(\sigma_i-1)(\sigma_i n_{i,t+1}-(\sigma_i-1))}{\alpha(\sigma_i n_{i,t}-(\sigma_i-1))} + n_{i,t+1} \right) \gamma e^{\frac{(\sigma_i-1)(n_{i,t+1}-1)}{n_{i,t+1}}} & \text{for } n_{i,t} > \frac{\sigma_i-1}{(\sigma_i-1)-\eta} \\ \\ \left(\frac{(n_{i,t}-1)(\sigma_i-1)(\sigma_i n_{i,t+1}-(\sigma_i-1))}{\alpha(\sigma_i n_{i,t}-(\sigma_i-1))} + n_{i,t+1} \right) \gamma e^{\eta} & \text{otherwise.} \end{cases} \tag{6.21}$$

With a constant number of firms in the steady state, the largest integer $n_i \geq 2$ that satisfies

$$L_i \geq \begin{cases} \frac{(n_i-1)(\sigma_i-1)+\alpha n_i}{\alpha}\gamma e^{\frac{(\sigma_i-1)(n_i-1)}{n_i}} & \text{for } n_i > \frac{\sigma_i-1}{(\sigma_i-1)-\eta} \\ \\ \frac{(n_i-1)(\sigma_i-1)+\alpha n_i}{\alpha}\gamma e^{\eta} & \text{otherwise.} \end{cases} \tag{6.21a}$$

implicitly defines the steady state number of firms. In the non-growth scenario of Bertrand competition, the number of firms can be explicitly defined as the largest integer $n_i \geq 2$ which satisfies:

$$\frac{\alpha L_i + (\sigma_i - 1)\gamma e^{\eta}}{((\sigma_i - 1) + \alpha)\gamma e^{\eta}} \geq n_i \qquad \text{for } n_i < \frac{\sigma_i - 1}{(\sigma_i - 1) - \eta} \tag{6.21b}$$

As above, the n_i which satisfies Equations 6.21a or 6.21b can be used to solve for any variable.

6.3.4 Steady-state growth

The measure of growth for the economy as a whole is the rate at which total output increases. In this model, total output is made up of two components: production/consumption of manufactured goods and production of quality-improving innovations.

Defining growth in production/consumption of manufactured goods as growth in the ability to gain utility from consumption, utility increases at the rate that real wages increase. Real wages describe the ability for workers to spend wages to gain utility in consumption. Real wages are defined by $\omega_{i,t} = \frac{w_{i,t}}{P}$, where P describes the perfect price index such that it recognises the changes in costs and quality across all sectors. Since nominal wages are unchanging, real wages are increasing at the rate the price index declines. So far, price indices have been defined for each sector only. The perfect price index describes the overall cost of living. It is the P that buys one unit of Q:

$$P = \frac{1}{N}\prod_{i=1}^{N}\left(P_i^{\frac{1}{N}}\right): \tag{6.22}$$

where P_i is given by Equation 6.6a.

The first step to determine growth in consumption, is to first determine how the price index changes between periods. Defining technology growth for each industrial sector i as the rate that the quality level increases, the steady-state technology growth rate with Cournot competition is given by:

$$g_{i,A} = \frac{A_{i,j,t} - A_{i,j,t-1}}{A_{i,j,t-1}} = \frac{\varepsilon(\sigma_i - 1)(n_i + \varepsilon - 1)}{\eta(\sigma_i(\varepsilon + 1) + n_i - 1)} - 1. \tag{6.23}$$

The technology growth rate with Bertrand competition is given by:

$$g_{i,A} = \frac{\varepsilon\left(\sigma_i - 1\right)\left(n_i + \varepsilon - 1\right)}{\eta\left(n_i + \varepsilon\right)} - 1. \tag{6.24}$$

The sector i price index is falling at a rate of $\left(g_{i,A}\right)^{\sigma-1}\frac{1}{1-\sigma}$. The growth rate of consumption/production is given by the rate the perfect price index declines:

$$g_Q = \frac{1}{N\left(1-\sigma\right)}\prod_{i=1}^{N}\left(g_{i,A}\right)^{\sigma-1} \tag{6.25}$$

Since production/consumption is one portion of the economy and the other portion of the economy, innovation, is constant, the overall growth in GDP means each sector must be multiplied by the percentage of the workforce employed in manufacturing. Growth in consumption/production and GDP growth depend upon the particular make-up of the economy. In particular, the form of competition in each market and the level of competition determine growth in technology in each sector and if these vary between sectors and countries, so will long-run growth rates.

6.4 Long-run growth and simulations

Simulations show the impact on the general equilibrium growth rate that results from discrete entry in individual markets. The simulations in this chapter use hypothetical combinations of sectors with varying levels of competition by varying the distribution of the specialist supply of labour across sectors to describe the overall effect of competition on innovation and economy-wide growth. Simulations are a numerical exercise to demonstrate the impact of differing competition levels in some or many sectors on economy-wide growth and not intended to represent an actual economy. Calibrating the model to a real world example is beyond the scope of this thesis as it requires an empirical foundation for calibration. Instead, the results of this simulation exercise offer implications for growth and innovation policy in economies based on the mode and level of competition in some or many sectors.

Prior to simulating the general equilibrium, the relationship between competition and innovation is examined by considering the relationship between the supply of labour in each sector and firms' investment in innovation. This conveniently isolates the impact of competition on innovation by adjusting the labour supply on a sector basis and also allows for a fair comparison between Bertrand and Cournot competition. Increasing competition encourages firms to invest in innovation as their own impact on market elasticities diminishes with a greater number of firms. Furthermore, the use of a growth model without scale effects means any difference in investment in innovation is purely a result of competition effects and discrete entry rather than market size. The impact of competition on innovation can be considered both in terms of the number of firms n_i and in terms of labour market size L_i. While Bertrand competition results in greater innovation than Cournot competition when there are only a few firms, Cournot competition results in a greater number of firms given an identical labour supply. Despite this greater number of firms, under these simulated examples, Cournot equilibria results in lower growth given the same labour supply.

Secondly this section compares growth rates in different simulated examples. Labour is distributed across 100 sectors. The base scenario distributes labour in a linear fashion, and four other scenarios are built from here. The labour supply is distributed across sectors rather than distributing competition, because competition is a result of the labour supply rather than vice versa. These five scenarios are examined to consider the effect of discrete entry and competition on overall growth. Initially γ is calibrated such that the highest level of competition resulting from labour is 29 firms in the 100^{th} sector of the base scenario under Bertrand competition. A range of competition levels up to 29 firms is satisfactory to demonstrate the role of individual sectoral competition in macroeconomic growth with Bertrand competition. The simulations are then repeated using Cournot competition under the same calibration, so a fair comparison can be made. Furthermore, it is this calibration that is used for the examination of innovation rates with increasing competition/labour supply. Lastly, γ is recalibrated for the Cournot model to be able to examine it closer at low levels of competition.

6.4.1 Parameters

Simulation parameters are used similar to those in Chapter 5. The initial set of parameters are developed for the base model and these parameters are used to develop other models for comparison. The base model describes a market where the labour supply is distributed in a linear fashion across 100 sectors.

CES simulations such as in Krugman (1991b) use an elasticity of $\sigma = 4$ or as in Baldwin and Forslid (2000) using $\sigma = 5$. Since the economy is disaggregated to a much greater extent than other CES models, this justifies a higher elasticity of substitution between varieties in the same sector. However, an extremely high σ means varieties in the same sector are very close substitutes. As a result, consumers are more responsive to changes in quality. That is, a high σ overwhelms other factors in determining the quality target. Therefore, a fixed elasticity is chosen for all simulations of $\sigma = 25$ (i.e., higher than Krugman (1991b) or Baldwin and Forslid (2000)), because there is greater disaggregation, but it is not so high as to overwhelm calculations.

As in the previous chapter the rate of time preference is set by following Baldwin and Forslid (2000) with $\alpha = \frac{1}{2}$, which implies an annual discount rate of approximately 7% when periods represent 10 years. η is calibrated such that the expected free entry growth rate equals the rate of time preference, i.e. $\eta = \frac{(\sigma-1)}{2} = 12$. Initial values for A_i are set to 1 such that all growth in technology can be easily compared to initial technology levels. Simulations run for 15 periods. Simulations can be run indefinitely, but in the deterministic model used here, they do not show any further detail over the long run. γ is calibrated in order to set the total labour supply to 1 in the base model and the maximum competition to 29 firms under Bertrand competition. Setting the total labour supply to 1 allows the share of employment that is allocated to each sector to be easily interpreted. In the other simulations, the same value for γ is used in order to fairly compare growth rates between different simulations. The simulations are repeated with Cournot competition under the same calibration to compare growth rates. However, under the initial calibration, the smallest sector still has 74 firms with Cournot competition and 6 firms with Bertrand competition. Lastly γ is recalibrated such that competition is a maximum of 29 firms under Cournot competition so that the Cournot model can be examined at lower levels of competition.

6.4.2 Competition and innovation

In order to fully understand the macroeconomic growth effects of competition levels in each sector of the economy, it is vital to first understand the relationship between competition and innovation with discrete firms on a sector basis. The following figures describe this relationship. As the number of firms in a sector rises, innovation increases with each additional firm due to competitive pressure, as firms respond less to their own effect on the price index, because they become a smaller overall share of the sector. The figures also examine the relationship between innovation and the supply of the factor of production (labour) which is directly related to competition in this model. As the availability of the factor of production rises, competition increases if it leads to an additional discrete firm. As a result, innovation increases stepwise, with each step due to the addition of another discrete firm.

With a growth rate of 1 being the expected innovation rate under the fully competitive limit model, Figure 6.1 describes the innovation rates for competition levels of 2 to 100 firms in each sector under Bertrand competition. As competition increases, innovation rates quickly tend towards the expected rate under the fully competitive or continuous model. With the assumption of discrete entry, innovation rates never quite reach this continuous limit such that there are always costs to the market from this barrier to entry.

Alternatively, Figure 6.2 describes innovation rates for competition levels of 2 to 100 firms in each sector under Cournot competition. As competition increases, innovation rates also rise. However, for low levels of competition, there is zero growth, as simply paying

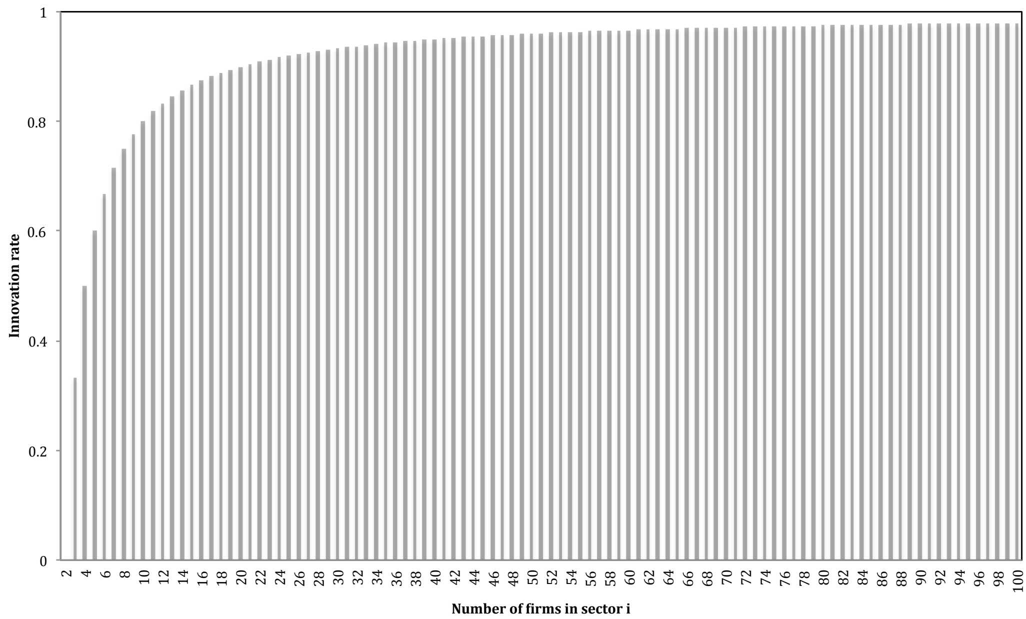

Figure 6.1: Number of discrete firms vs innovation rate under Bertrand competition

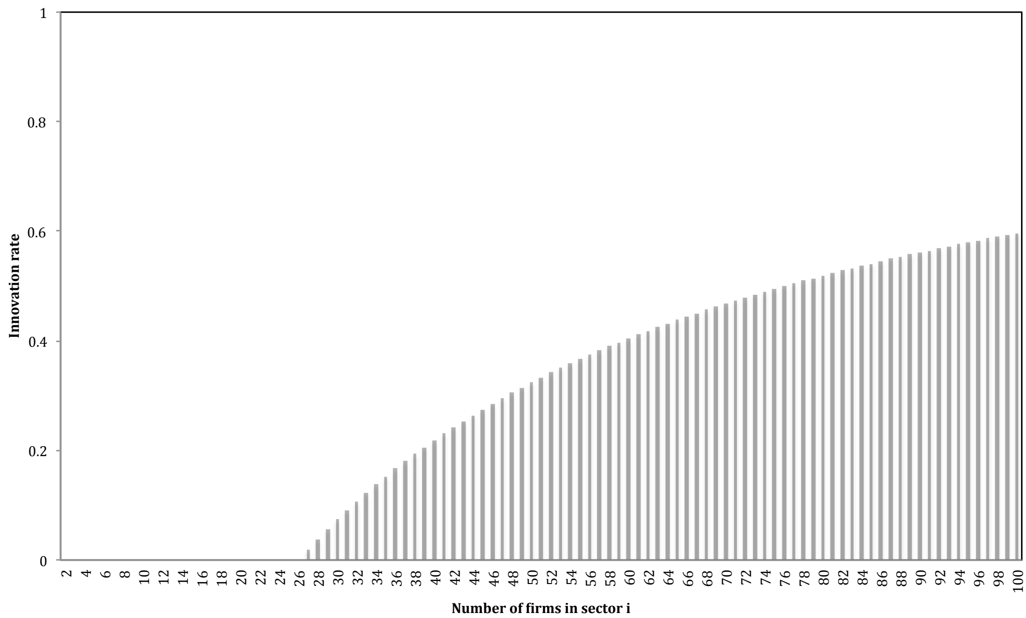

Figure 6.2: Number of discrete firms vs innovation rate under Cournot competition

the minimum fixed cost without any quality improvement is enough to deter entry of the marginal firm. It is not until there are 27 firms in the sector before there is positive growth.

However, this does not tell the full story. The comparison on on a competition level basis only is misleading, because firms in Bertrand and Cournot sectors require different levels of the factor of production, particularly in research. When a single firm employs fewer research workers as in Cournot competition, many more discrete firms can enter the market and compete. For example, using the parameters above, a labour supply that results in 3 firms (and an innovation rate of 0.33) in a Bertrand sector results in 34 firms (and an innovation rate of 0.14) if it were a Cournot sector.

Using a gamma calibrated for the Cournot simulations, the minimum and maximum labour supply is calculated that would result in 2, 3, and up to 100 firms in the sector for both Cournot and Bertrand innovation and the figures below consider the innovation rates in relation to this labour requirement. Under the initial calibration, Cournot competition requires a labour supply of 60 for 100 firms while Bertrand required a supply of 7200 for 100 firms. With one being the expected innovation rate for the continuous model, Figures 6.3 and 6.4 describe how innovation rates rise with increases in the supply of labour (and consequent increases in competition) for Cournot and Bertrand innovation respectively.

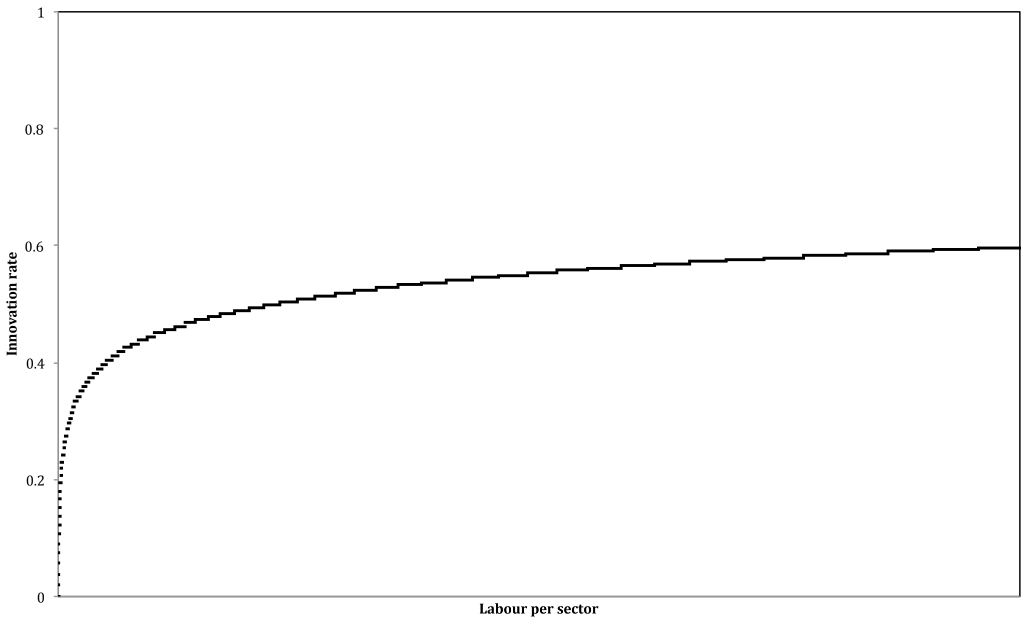

Figure 6.3: Labour supply vs innovation rate under Cournot competition

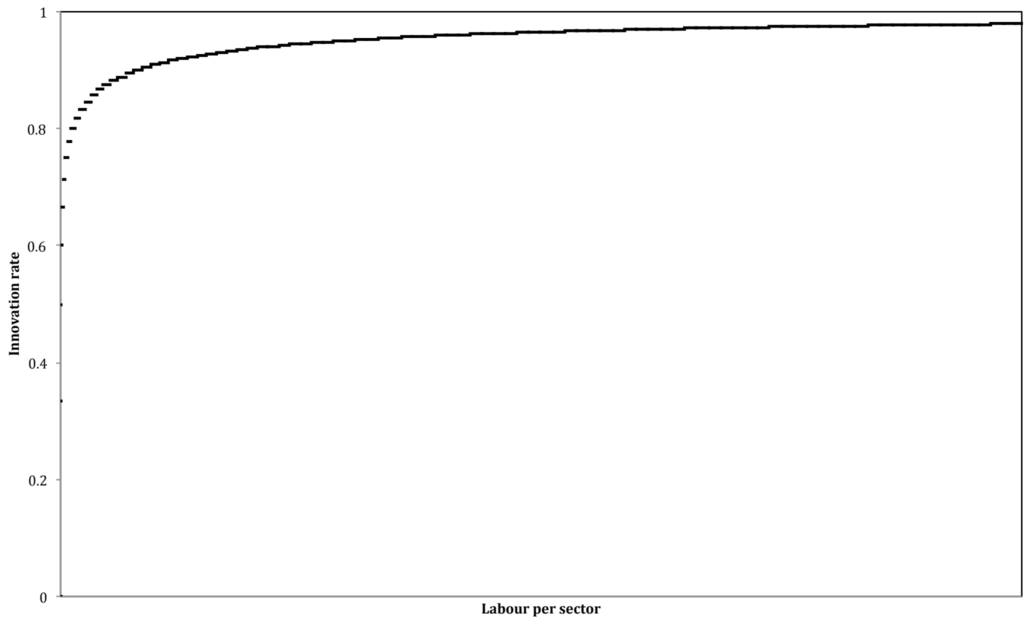

Figure 6.4: Labour supply vs innovation rate under Bertrand competition

Innovation rates increase stepwise as additional labour allows an additional firm to enter the market at each “step”. Each step is discrete and tends towards the growth rate of the continuous model. These steps are larger in Bertrand competition resulting in higher innovation rates for the same labour supply (keeping all other parameters the same). But these higher rates occur from competition between substantially fewer firms in Bertrand competition, meaning the two alternatives are not as substantially different as it appears in Figures 6.1 and 6.2. Comparing the innovation rates under Bertrand and Cournot competition on a labour supply basis, Figure 6.5 combines Figures 6.3 and 6.4 and uses a smaller scale. The lower line represents Cournot innovation (from Figure 6.3) and the upper line represents Bertrand innovation (from Figure 6.4). Notably the trend towards the continuous innovation rate is much closer when compared on a labour supply basis rather than a pure competition basis. For policy makers, the mode of competition is particularly important when considering the effect of competition on innovation rates. Firms that compete fiercely on price in sectors under Bertrand competition may be particularly innovative, even with only a few firms, while Cournot sectors could contain many firms but not achieve high rates of innovation.

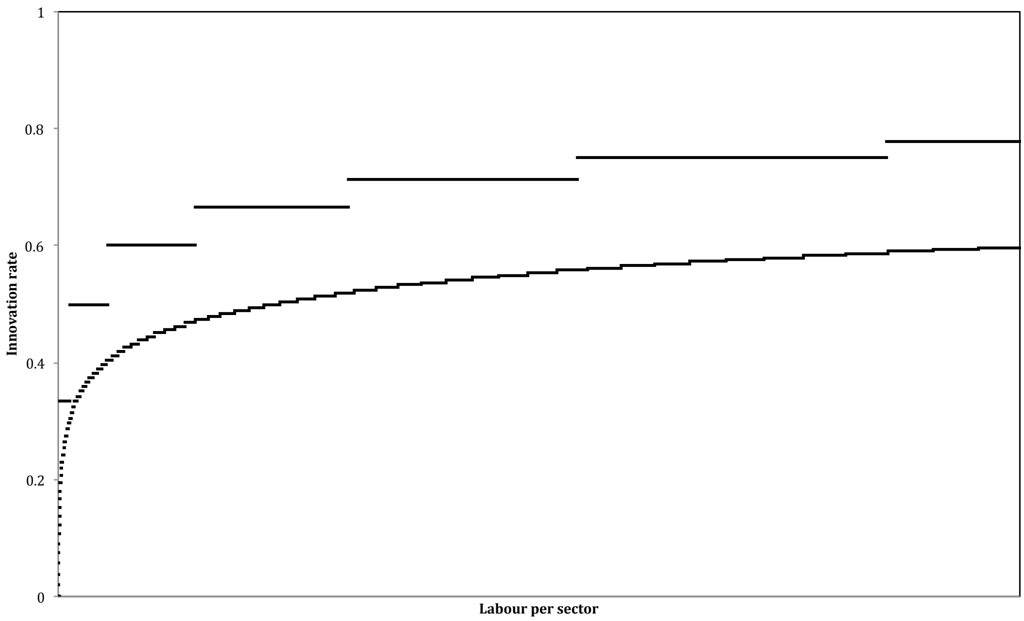

Figure 6.5: Labour supply vs innovation rate under Cournot and Bertrand competition

6.4.3 Simulations: Competition and growth

Over an entire economy, it is possible for a few or many (or somewhere in between) sectors to have low levels of competition. The overall effect on economy-wide growth depends upon the portion of the economy made up of low competition sectors. To examine the effects of competition issues on macroeconomic growth, growth is modelled under five scenarios where the factor of production (labour) which determines competition is distributed in a linear pattern across 100 sectors. Simulations are repeated under Cournot and Bertrand competition. Initially, the model is calibrated for the Bertrand model and the base labour distribution generates competition ranging from six to 29 firms in each sector in the Bertrand simulation and from 74 to 201 firms in each sector for the Cournot simulation. The exercise is repeated, calibrating γ for Cournot competition to range from three to 29 firms in each sector.

For each of the five different scenarios, labour is distributed across 100 sectors. Labour supply is distributed across sectors rather than across levels of competition, because competition is a result of the labour supply rather than vice versa. The base scenario distributes labour linearly and the four other scenarios are developed from this base scenario. Ordering 100 sectors from smallest to largest, total labour supply of one is distributed across the sectors in a linear relationship such that the supply of labour is 0.0002 in the smallest sector, increasing by 0.0002 in each sector until the largest sector is reached with a labour supply of 0.0198. The second scenario examines the situation where a large part of the economy has a low supply of labour and low competition. In this scenario, the lowest supply of labour at 0.0002 (and consequently low competition) is allocated to the 20 smallest sectors before the linear relationship is introduced. As a result, Bertrand competition ranges from 6 firms in the first 20 sectors to 26 firms in the 100^{th} sector. In the third simulation the alternative is considered where many sectors have the higher competition level. In this scenario the highest supply of labour at 0.0198 is allocated to the 20 biggest sectors with competition ranging from 14 firms in the smallest sector to 29 firms in the largest 20 sectors. In the last two scenarios every sector is allocated with the lowest (scenario 4) and highest (scenario 5) supplies of labour respectively. In scenario 4, all 100 sectors have 6 firms and in scenario 5, all 100 sectors have 29 firms. This is clearly extreme, but these numerical exercises include the more extreme examples to emphasise the effects of discrete entry on economy-wide growth. These five scenarios are examined to consider the effect of discrete entry and competition effects on overall growth. A summary of the five scenarios is described in Table 6.1.

Scenario	Labour distribution	Number of firms in each sector
One	Linear	6 to 29 firms
Two	20 sectors with small labour supply	20 sectors with 6 firms, 80 with 7-26 firms
Three	20 sectors with large labour supply	20 sectors with 29 firms, 80 with 14 to 28 firms
Four	100 sectors with small labour supply	100 sectors with 6 firms
Five	100 sectors with large labour supply	100 sectors with 29 firms

Table 6.1: Summary of simulation scenarios

γ is calibrated such that the highest level of competition is 29 firms in the 100^{th} sector of the base scenario under Bertrand competition. Simulations are repeated using Cournot competition under the same calibration, so a fair comparison can be made. It is this calibration that is used above for the earlier examination of innovation rates with increasing competition/labour supply where the minimum and maximum labour required was examined for each competition level from 2 to 100 firms per sector. Figures 6.6 and 6.7 describe growth rates under scenarios one to five for Cournot and Bertrand competition respectively.

Increasing competition clearly leads to higher growth rates. This provides an additional theoretical explanation for the upward sloping (and diminishing slope) portion of the inverted U relationship between competition and innovation (Aghion et al., 2005). In that model the shape owes itself to the escape competition effect being dominant for competitors that have "neck and neck" technology levels. Each firm's "distance to frontier" (Acemoglu et al., 2006) technology determines the extent to which competition has an increasing or diminishing effect on innovation. The model here is simpler by assuming symmetry, but a similar "escape competition" effect is still present such that firms have to invest more in innovation to block the marginal entrant when there is greater competition.

The higher competition levels in simulations 3 and 5 result in significantly higher growth rates than the low competition levels in simulations 2 and 4. But the existence of some low

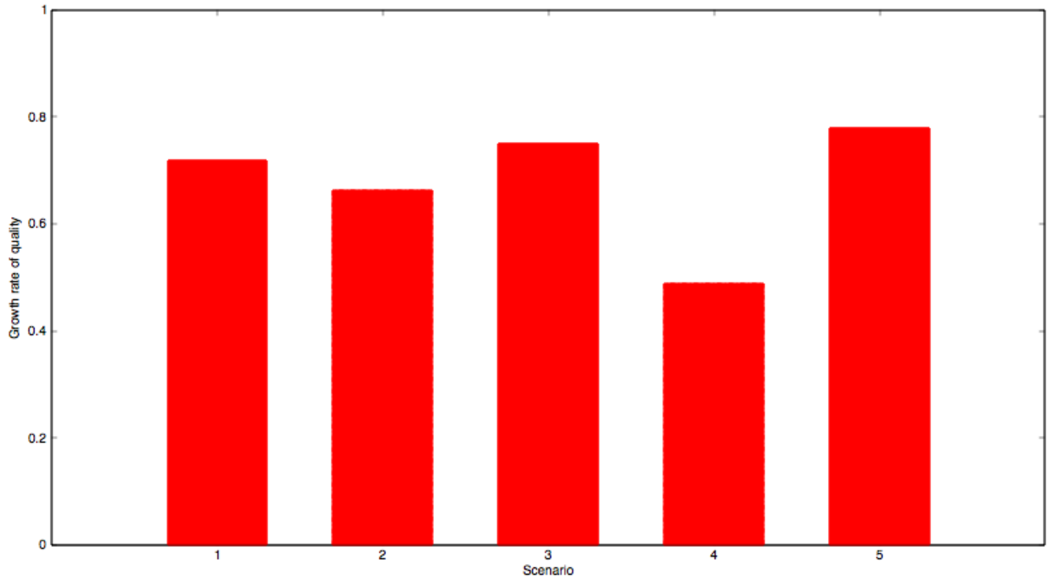

Figure 6.6: Growth rates for five simulated scenarios under Cournot competition

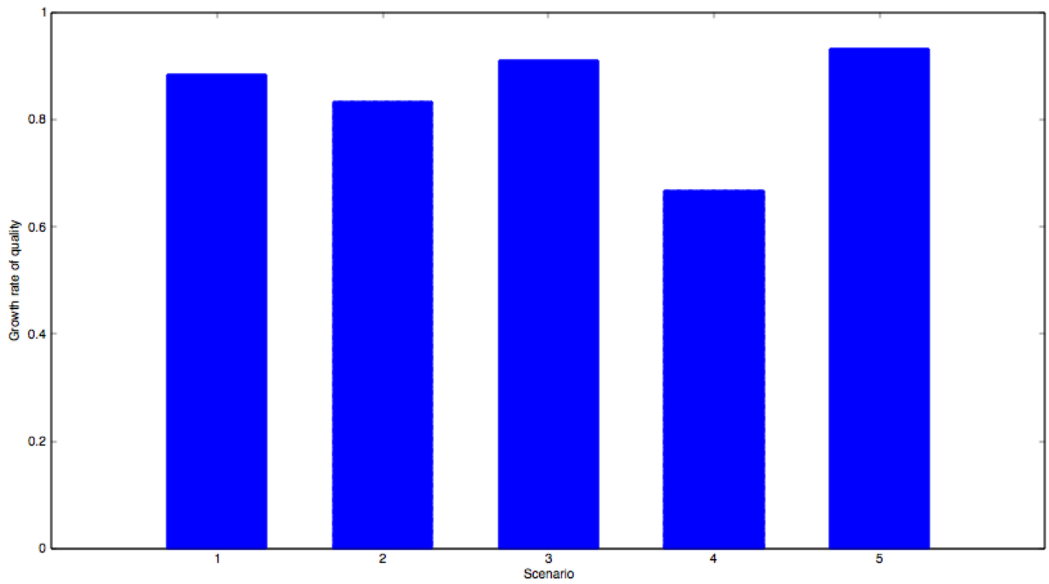

Figure 6.7: Growth rates for five simulated scenarios under Bertrand competition

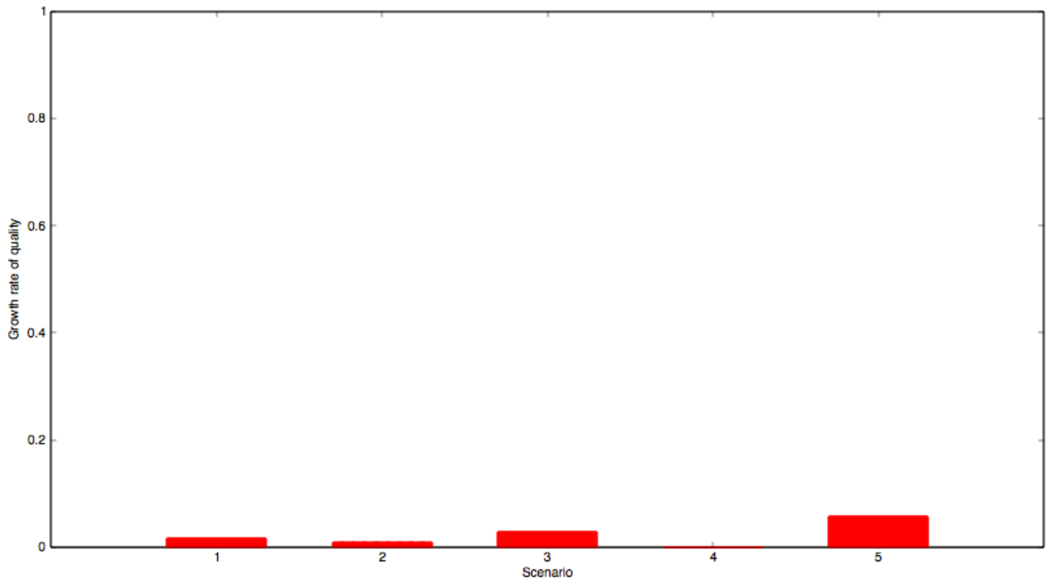

Figure 6.8: Growth rates for five simulated scenarios (recalibrated) under Cournot competition

competition sectors does not always have a substantial effect on the economy-wide growth rate. Each sector only makes up 1% of expenditure and unless the economy is dominated by low competition as in simulation 4, the effects are not substantial. The simulations suggest that there are implications for innovation and growth policy at a sector level. Hence, it is important to not only consider growth policy as a macroeconomic problem that requires economy-wide solutions. As can be seen above problems that impede competition (such as discrete entry) have a negative effect on growth. With Cournot competition, there are considerably more firms per sector (201 compared to 29), but the additional competition results in lower growth due to the nature of innovation under Cournot competition.

Investigating this further, γ is recalibrated for the Cournot model to examine it closely at low levels of competition such that the highest level of competition in the Cournot model is 29 firms. The results presented in 6.8 are extremely low growth as discrete Cournot firms are able to protect their market position without innovating unless there are at least 27 firms. Even with this many firms, market participants only have to innovate slightly to maintain their market position. While this figure is representative of the parameters chosen here rather than representing a specific economy, it strongly indicates that few firms and Cournot competition can be significantly detrimental to innovation and growth. The exercise is not repeated with Bertrand competition under this second calibration as there would be less than 2 firms in each sector.

6.5 Discussion of the model

Discrete entry results in firms investing in innovation in response to actual competitors rather than also to potential entrants. Firms conduct incremental innovations, but quality improvements are lower than in the continuous free entry model. However, additional competitors mean innovations quickly tends towards the continuous growth outcome under Bertrand innovation. While additional firms in the Cournot model also increase innovation, it takes many more firms to reach the continuous growth outcome. The comparison is not completely fair, because holding all other parameters constant, Cournot firms allow substantially more firms to enter the market using the same factors of production. Despite these additional firms, in this calibration Cournot innovation also lags behind Bertrand innovation on a labour supply basis in terms of tending towards the growth outcome for the continuous free entry model.

Notably, the existence of a few small markets does not condemn an economy to lower innovation and growth under Bertrand competition. In the simulation, each sector only makes up a $\frac{1}{100}$ portion of consumption expenditure. So only when the economy is dominated by small markets is there a significantly greater effect. Furthermore, the exact distribution of sector size is important such that average market size is not the important factor, but to what extent the overall market is made up of small sectors. Cournot competition is more detrimental to growth on a number of firms basis, but lower innovation rates are mitigated by an increase in the number of firms able to enter the market with the same factors of production. When Cournot firms strategically underinvest in innovation, they allow additional firms to enter the market. Despite this additional competition, in this calibration Cournot competition with discrete entry results in lower growth than Bertrand competition. It is possible this is the case for all possible calibrations, or that some specific criterion determines whether Cournot or Bertrand competition leads to higher innovation rates but the proof of this is left for future research.

Low competition or discrete entry as modelled in this chapter may be a more common characteristic of small or isolated markets. The size of these markets may not be attractive to a multi-national or new entrant such that these markets can remain isolated from the competitive forces in other larger agglomerations. In this way, the model developed by this chapter provides further evidence for localised and industry-specific innovation and growth policy, particularly in small, isolated or non-tradables markets where discrete entry is more likely to be an issue. There is no one-size-fits-all approach to stimulating innovation and growth. Economic policy must be tailored to an industry, country or region's specific characteristics.

6.5.1 The additional market failure of discrete entry

There is a market failure for innovation in endogenous growth models (Aghion and Howitt, 1992; Grossman and Helpman, 1991b; Romer, 1990) such that firms underinvest in innovation activities because of inter-temporal spillovers. Investors in R&D do not invest at the socially optimal level of innovation, because they cannot keep all of the future benefits generated from their investment. Those benefits accrue to future innovators. This feature is also present in the model here, but there is an additional market failure, because firms respond to actual competitors, including their own effect on the price quality index. When a single firm is a large part of the sector price quality index, there is an additional market failure due to imperfect competition. While the continuous model assumes that a large number of entrants has made this negligible, in the disaggregated discrete model, this market failure can be significant. The market failure appears most clearly through prices as is already well understood in the industrial organisation literature. What has been less clear up until now is the extent to which market failure from imperfect competition could be present in the market for quality improving innovations.

The size of the market failure is determined by the extent that the observed elasticity of substitution is affected by the firms' own prices and quality levels and by the mode of competition. In Bertrand competition, even if a firm is one of only a few competitors, its effect on the perceived elasticity may be limited. Adding firms to the sector results in the perceived elasticity quickly tending towards the CES elasticity from the continuous model. However, in Cournot competition, the perceived elasticity tends towards the CES rate at a much slower rate. Once again, this is also related to the factors of production (labour) such that the same factors of production result in many more firms entering the market under Cournot competition than under Bertrand competition. Nonetheless, the additional firms are not enough to mitigate the Cournot effect on innovation. Innovation is always lower under Cournot competition for the same factors of production. Therefore, the size of the innovation market failure is greater under Cournot competition. Examining Figures 6.1 to 6.5, the size of the additional market failure in each sector is the difference between the innovation level on the chart and the continuous innovation rate at 1. Notably, the size of the additional market failure diminishes with additional firms.

6.5.2 Competition, innovation and growth

While low competition levels results in lower levels of innovation, the existence of a few low competition sectors does not always have a substantial effect on the economy-wide growth rate. In the simulations here, each sector only makes up 1% of expenditure and unless many sectors are dominated by low competition, the effects are not necessarily substantial. However, the causes of low competition may be a property of a specific economy, region, location or industry. It is therefore possible that some sectors will be highly competitive in some countries or regions and uncompetitive in others. When considering how competition policy affects growth and innovation policy, policy-makers must examine local characteristics which will vary among different cities, regions and nations. Furthermore, the mode of competition is a particularly important market characteristic. Cournot sectors have significantly less innovation per firm, although this allows many more firms to enter the market. However, additional firms do not necessarily lead to high growth rates, because Cournot models require significantly more firms to enter the market before growth rates trend towards the continuous free entry rate. In considering the effect of competition and market entry characteristics on innovation, each region and industry will have unique characteristics which determine the overall effect on growth and the appropriate policy response to support innovation.

Examining Figures 6.3 and 6.4, low competition dominates simulations 2 and 4 resulting in lower growth. While low competition is also present in simulations 1 and 3, the effect is diminished substantially, because these low competition sectors only make up a small portion of the economy. Similarly, simulations 3 and 5 are dominated by sectors with high competition, which substantially overcomes the additional market failure of innovation caused by discrete entry. Notably, growth rates are considerably lower or even zero when simulations are considered with Cournot competition (Figures 6.4 and 6.5). Low competition in a Cournot sector has a significantly negative impact on economic growth. When comparing Cournot and Bertrand innovation on a production factor basis (labour), Cournot sectors have considerably lower innovation rates and consequently growth rates, even though Cournot competition allows significantly more firms to enter the market. As a result, policy-makers examining competition and innovation policy need to consider both the number of firms and the mode of competition. In response to barriers to entry such as the discrete entry assumption used in this chapter, policy makers should focus on removing or reducing these barriers, perhaps by offering assistance to new entrants and greater assistance for Cournot markets.

The results here are consistent with Aghion et al. (2005) and Aghion and Griffith (2005) who also examined the relationship between competition and innovation, finding an inverse U-relationship. The upward sloping portion of the inverted-U can be observed in the increasing rate of innovation in response to a greater number of firms and the diminishing size of the increase at higher growth rates. Aghion and Griffith's (2005) research describes a relationship between technology-leading and following firms where competition discourages following firms from innovating, but encourages leading firms who are attempting to "escape competition". In this model, firms are also more innovative as they respond to greater competition, but following firms are unable to make a profit, as an entrant could incrementally innovate by a greater amount and replace the following firm as a discrete participant. Escaping competition the Aghion et al. (2005) model comes in two forms. Firstly, firms develop innovations to ensure that they are the technology leading firm in the coming period. This is known as "escaping competition" *for* the market. Contestability as in the continuous growth model stimulates innovation by both incumbents and potential entrants. The model here also includes the "escape competition" effect but it is dominated by its second form, that is, competition *in* the market. Discrete entry creates a barrier to entry reducing the escape competition *for* the market effect, but maintaining the escape competition *in* the market effect. The second half of the inverted U relationship is not seen because discrete entry prevents the escape competition *for* the market effect from eventually dominating the escape competition *in* the market effect.

Disaggregating the model into many sectors, this chapter has extended the positive relationship between competition and innovation to a relationship between competition and economic growth. If some sectors suffer from a lack of competition resulting in lower innovation, there is a flow-on effect that results in lower overall economic growth.

6.5.3 Discrete and stochastic innovation

In stochastic models of innovation, initial market entry is attained by investing in R&D to achieve an innovation stochastically. Each time a new stochastic innovation arrives, a new firm enters the market replacing an existing incumbent. The stochastic and deterministic approaches are reconciled by Reinganum (1983) by suggesting the deterministic model is most appropriate for incremental innovations, where the incumbent can minimise risk by aiming for incremental innovations with a high probability. The stochastic mechanism creates industry churn which would occur on the rare occasion that an entrant achieved a drastic innovation to enter the market with a substantial quality improvement. Industry churn was the focus of Duranton (2007), so it is not readdressed in the model here. If stochastic processes were included, it is possible, or even likely in some industries, that innovations occur with a very low probability. This would result in an incumbent maintaining incremental innovations for a long time as in the model here, responding only to actual market participants rather than potential ones. Responding to potential entrants is unnecessary. This would suggest that the R&D investment cycle of a firm is a process of researching large stochastic innovations to enter the market, followed by incremental innovations to maintain a market position, until a new competitor takes the market with a new stochastic innovation. At this point, the original firm either exits completely or returns to investing in stochastic innovation to be a potential competitor in future periods. This idea essentially combines the idea of industry churn from large stochastic innovations with discrete incumbents developing incremental innovations.

If stochastic innovations occur with a high probability, the additional market failure of discrete entry may be overcome if there are few barriers to firms investing in these stochastic innovations. This essentially describes a situation where the market is large and contestable. However, if stochastic innovations occur with a low probability and/or if there are barriers to firms investing in stochastic innovation, the additional market failure of discrete entry described above will be sustained. Barriers that may prevent R&D investment by stochastic entrants could include sector or variety-specific knowledge which is embedded in people and consequently with their employers and locations. The inability for new entrants to access the necessary knowledge to invest in innovation as a stochastic entrant prevents them from developing the stochastic innovations which result in industry churn. Policy-makers must also consider on an industry/location/country basis if there are barriers to firms investing in stochastic innovation. For sectors where stochastic innovation is rare, the additional market failure of discrete entry could be significant, while for sectors with a high rate of churn, stochastic innovation mitigates this market failure.

6.5.4 Policy implications for competition, innovation and growth

This research produces a series of implications for policy-makers. Characteristics of individual sectors, economies or regions need to be monitored which could signal to policy-makers that discrete entry is causing an additional market failure for innovation. Markets which are characterised by few competitors, high margins and low industry churn are likely to also have low rates of innovation, as market participants respond only to actual competitors rather than potential competitors. This is also more likely to occur where the factors of production for innovation are particularly specialised, in short supply or largely controlled by one or a few firms. This factor of production does not have to be limited to labour, but could include GPTs, R&D facilities, capital or any other factor that was left out of the model here. Where there are barriers to new entrants producing innovations in order to enter the market, discrete incumbent firms can maintain only incremental innovations responding to actual competitors without concern for potential competitors.

To address these symptoms, there are a number of options to facilitate more socially optimal outcomes. Firstly, policy responses need to focus on removing barriers to entry which sustain large and dominant firms that do not face a threat of entry. It is this threat of entry which leads to additional firms, industry churn by stochastic innovation (Duranton, 2007) and the continuous free entry growth outcome (Young, 1998). Secondly, the potential for trade may also compensate for low competition or supplement local competition. Removing trade barriers and investing in infrastructure which reduces transport costs (by making a market more accessible) increases the size of the market. Most importantly, it increases the supply and availability of factors of production leading to greater firm entry overall and more competition to encourage firms to invest further in innovation. While this will not be possible for non-tradables, the characteristic of being a non-tradabile good or servive is precisely a sign for policy-makers to monitor the sector's innovation performance and potentially address any issues with industry-specific innovation policies. As an addition, other market-building policies such as those which encourage agglomeration will also result in more competitive pressure to develop innovation, mitigating the innovation market failure of discrete entry.

Policy makers should also consider factors which influence the ability to develop innovations. Innovation requires a number of factors to take fundamental research through to invention and eventually innovation. In this model, these factors are represented by the parameters γ and η. They include R&D-related infrastructure (such as the tools and equipment used by primary researchers), the depth of local capital markets, the availability or thickness of local labour markets and other factors considered by the innovation economics literature (Nelson, 1993). These factors can be location or industry-specific such that they can only be addressed by location or industry-specific policy. By disaggregating the model into many sectors the model here allows policy makers to consider sector-specific characteristics of innovation and apply sector-specific policy to address the broader policy portfolio of economic growth.

6.6 Conclusion

Essentially the model here is able to combine the partial equilibrium characteristics of individual sectors such as barriers to entry, Cournot or Bertrand competition, and imperfect markets for innovation with the broader general equilibrium features from endogenous growth theory. This encourages a revision of economic growth policy that tailors a policy approach to local, regional or industry characteristics. The model finds that a lower supply of the factors of production for innovation, results in lower competition also results in an even lower investment in innovation because of the discrete nature of market entry. Firms investing in incremental innovations are able to invest in innovation in response to actual market participants rather than potential entrants.

Entrepreneurs are paid out in firm profits, but, with only discrete entry, entrepreneurs earn higher wages than labour employed in manufacturing. As a result, industries with few competitors have lower levels of investment in innovation and a greater disparity between the wages of manufacturing workers and entrepreneurs. Lower innovation because of discrete entry has a flow on effect to overall economic growth, even when it occurs in just a few sectors. This lower competition is a characteristic of individual sectors and therefore, policies to stimulate growth should be more focused on policies that target innovation, competition or firm entry in these individual sectors. Economic growth policy is therefore not a generic approach, but a localised, industry and economy-specific policy problem.

Future research on understanding innovation and its link to economic growth should therefore also focus on localised, industry or economy-specific characteristics that may affect competition levels or the ability to produce innovations in individual sectors. Economies of scale can result in a dominant incumbent, restricting innovation as the market is dominated by only a few firms. Fields such as industrial organisation, labour economics, economic geography, and regional and urban economics have much to contribute to the study of the contribution of local or industry factors to competition, innovation and subsequently economic growth. Innovation is much more complex than endogenous growth theories have suggested and the general equilibrium model here adds to the growing body of evidence that growth and innovation policy should focus on localised and industry factors of innovation.

Chapter 7

Conclusion

7.1 Findings of the thesis

Innovation is the fundamental engine of economic growth that cannot be understood without considering the location of economic activity, the characteristics of knowledge spillovers and the nature of competition. It is the characteristics of different locations and industries that have strong implications for the economic outcomes and well-being of the people who participate in these economies. Even small differences between locations cumulate over many years and thereby have much greater consequences for long-run standards of living. It is these initially small differences between core and periphery regions which may have left peripheral locations with a lower standard of living than core locations. With a better understanding of innovation, there is the potential for more effective approaches to policy such that innovation-based growth of peripheral regions can be sustained.

By reconsidering how innovation is understood within theoretical models of economic growth, this thesis provides evidence for tailored policy solutions that support economic growth and the long-term well-being of peripheral regions, that may differ from a universal approach to economic policy. Endogenous growth theory is based on an assumption that the cost of innovation depends on available knowledge. New knowledge is an input to future innovation activities. This thesis explores models of growth that incorporate a better understanding of innovation within the framework of endogenous growth theory. Contributions from the NEG acknowledge that knowledge available to the firm is affected by spatial externalities. Furthermore, there is a technological dimension to the transfer of knowledge between firms. Knowledge diminishes with distance and with decreasing technological relatedness. This leads to an additional force for firm location: the clustering effect.

As a background, Chapter 2 explores the development of the fields of endogenous growth, economic geography and innovation as well as the similarities and differences between these fields. The compatibility of modelling techniques, such as Dixit-Stiglitz competition (in endogenous growth and NEG) and the spatial nature of innovation (in economic geography) suggests these three areas offer avenues for understanding innovation and its consequences for growth. It is this foundation on which the NEG has developed new models that incorporate spatial factors into theories of economic growth, as discussed in Chapter 3.

Making a new competition contribution to the literature, this thesis develops three endogenous growth models with two of these supported by simulations. This thesis makes key extensions to understanding innovation within endogenous growth theory and discusses the consequences for growth and the implications for economic policy.

7.1.1 Three new models

Chapter 4 presents a quality ladders growth model with additional knowledge spillovers between varieties such that the location of firms is related to the location of other manufacturing firms. Furthermore, the ability for manufacturing activity to change location and the likelihood of manufacturing agglomerating in a single region is related to the mobility of knowledge. In this model, greater knowledge spillovers make firms more likely to shift away from the periphery towards a large agglomeration (or be replaced by an alternative firm in the agglomeration). The model demonstrates how workers face trade-offs between location choices and has consequences for firm location choice. Firms balance the clustering effect alongside the traditional forces for agglomeration and dispersion from the NEG that affect worker location: the home market effect, the cost of living effect and the market crowding effect.

In Chapter 5, the market is disaggregated into many industrial sectors, where the level of knowledge spillovers is based on the technological relatedness of different varieties. This obtains a model that implies that the clustering of firms with related technologies could be a characteristic of economic activity in peripheral locations. Multiple industries account for the related technology dimension of knowledge transfer. This chapter refrains from modelling the impact of transport costs as in traditional NEG models and focuses solely on technical externalities in R&D. Firms are attracted to locate alongside their own industry due to the technological relatedness of innovation inputs creating the clustering effect. However, firms are also attracted to a larger agglomeration, even if the firms in the agglomeration are less technologically related. Even with the omission of transport costs, full catastrophic agglomeration is not inevitable. Firms in a peripheral cluster balance the trade-off between sourcing knowledge from a specialised peripheral cluster (the clustering effect) and a diversified agglomeration (the agglomeration effect). As a result, clusters of firms within the same technological sector can sustainably produce in the periphery and resist the agglomeration effect, because each firm in the cluster depends upon the continued co-location of other firms in the cluster. Including transport costs in the model would result in firms also balancing these clustering and agglomeration effects with the traditional location effects from the NEG: the market-access effect, the cost-of-living effect and the market-crowding effect.

Lastly, the model developed in Chapter 6 considers the proposition that market entry could be a prerequisite for innovation within an endogenous growth model with many sectors, such that the opportunity to innovate is only available to discrete entrants. The characteristics of market entry may mean the competitive profit-maximising response of these discrete firms can take the form of Bertrand or Cournot competition. As a result of discrete entry, the participating firms can only enter if there is space for a whole firm. Participants can invest in innovation strategically, so that there is space for less than one firm to enter the market. Hence, there is partial contestability such that an incumbent can remain the producer if they produce a large enough quality improvement to prevent the marginal discrete firm from entering the market. As a result, industries with few competitors have lower levels of investment in innovation, higher prices and greater returns to shareholders at the expense of consumer welfare. With increased competition, investment in innovation increases, as the space for less than one entrant becomes a smaller and smaller share of the market. While the mode of competition did not have any effect on a model with continuous varieties or a large number of firms, with discrete competition and only a few firms, the form of competition has a dramatic effect on the firms' incentives to invest in innovation. Small markets are likely to be a characteristic of some industries in isolated, distant and peripheral locations. Therefore, the model provides interesting implications for policy-makers in peripheral locations.

7.1.2 Implications for growth and policy

The thesis suggests a number of implications for economic growth and policy. In particular, the externalities and characteristics of innovation affect different locations and industries in different ways. As a result, distant or peripheral economies are likely to experience lower rates of innovation and economic growth. With market integration, growth can recover in the long run, but potentially to the medium-term detriment of peripheral industry if integration triggers catastrophic agglomeration. Furthermore, small markets may be subject to particular underinvestment in innovation, because discrete firms hold a dominant position in knowledge markets.

The model in Chapter 4 suggests that integration could be particularly costly for peripheral locations in the medium term. In the model, innovation activity shifts to the location that provides the greatest source of knowledge spillovers. Innovators that remain in the periphery will be overtaken by greater knowledge generation and innovation in the agglomeration. As a result, peripheral locations could suffer from a reduction in innovation activity due to increased market integration and a medium-term loss of growth. In the long run, economic growth may recover as trade enables the peripheral region to benefit from innovations that occur in the core region. But the cumulative effect of lower growth in the short term and the burden of transport costs results in peripheral locations lagging behind core locations. Instead, integration policy for peripheral governments should focus on the ability to transfer and trade in knowledge.

The implications from the model in Chapter 5 are not so pessimistic for the periphery. The model suggests that industrial clusters may provide a solution to retain knowledge industries in peripheral locations. Specialised industrial clusters can continue to produce sustainably in the long run. As a result, integration could still result in catastrophic agglomeration, but it is not inevitable. This means the periphery will be less likely to suffer from the medium-term loss of growth due to closer integration, as predicted in Chapter 4. While growth is slightly lower for each of these peripheral clusters than if the cluster were in the core region, quality improvement is higher for each individual firm by remaining in the cluster than by switching.

In particular, it is sectors where firms source knowledge more intensively from within their own sector that are more able to exist in the periphery compared to less knowledge-intensive sectors. If a cluster exists in the periphery it can remain in the periphery if its own industry knowledge intensity is greater than a critical level. That is, it relies more on spillovers from its own industry than spillovers from other industries. This critical level depends on the relationship between inter-industry spillovers and intra-industry spillovers. This suggests a role for peripheral governments in the establishment, monitoring and support of knowledge-intensive peripheral clusters to support regional economic growth.

In the model developed in Chapter 6, the overall impact of discrete entry upon economic growth depends upon the extent to which small markets are a characteristic of the economy as well as on the form of competition. Smaller industries have lower rates of innovation and competition, but this only has a significant impact on growth if these markets are a substantial part of the economy. While additional firms under Bertrand competition mean innovation quickly tends towards the innovation rate in the continuous model, Cournot competition requires many more entrants to tend toward this limit. At the same time and holding all other factors constant, Cournot firms are considerably smaller and a greater number of firms can enter a small market. But the positive effect on growth from these additional firms is never able to overcome the negative effect on growth due to lower rates of innovation in Cournot markets. By exploring the effect of discrete entry upon innovation and growth, this chapter finds that both the mode and level of competition are particularly important in determining economic growth outcomes. As a result, small or distant markets require more thorough monitoring and management of competition for the purpose of encouraging innovation and economic growth.

While there are many more dimensions to innovation, the models explored here show how innovation and growth are distinctly local phenomena. The outcomes for different locations depend upon the location of economic activity, the distribution of industries, the relatedness of technologies, the nature of market entry and the mode and level of competition in each market. Each of these factors has local impacts on growth that require a tailored, localised response from policy-makers.

7.2 Where to from here?

It is clear from the thesis that closer inspection of the factors affecting innovation is a fruitful direction for research. As with every research project, there were many possible directions this research could have headed. The models developed here therefore offer an interesting agenda for further research.

The extent to which knowledge spillovers from related technologies are useful to innovation activity is an important factor that determines the size of the market failure of innovation and the sustainability of peripheral clusters. While in the models developed in Chapters 4 and 5 these external sources of knowledge interact with a firm's own knowledge in a purely additive way, innovation and growth theory would benefit from empirically considering other forms of interaction among knowledge inputs. The models in this thesis assume additive knowledge inputs for simplicity. However, other forms of interaction between knowledge inputs should be explored further.

Similarly, there is potential to consider inter and intra-related technology spillovers as a continuous scale rather than the two or three-step scale used here. This extension is particularly complex but may offer further insights. There is a trade-off for policy-makers in peripheral regions between developing policies that encourage diversification and policies that encourage region specialisation, that could be thoroughly explored. In addition, the characteristics of an industry that determine the nature of knowledge spillovers for firms in that industry could determine which sectors are more likely to cluster in a peripheral region. While 5 considers a measure of own industry knowledge intensity this approach suffers from certain modelling limitations. Knowledge spillovers are much more complex in the real world. Research that considers the factors affection how knowledge from different sectors transfers between firms and locations would have significant implications for regional policy.

An extension to the model in Chapter 6 could consider whether incumbents can innovate strategically to limit competition. In this case, if it is assumed that an incumbent can block each potential entrant by innovating, this leads to the profit maximising firm blocking every potential entrant such that the model reaches a corner solution with fewer and fewer competitors. The reduction in competition in turn reduces the incentives to innovate. Such a model would require a stochastic approach to firm entry. Stochastic firm entry would encourage a strategic response where new entrants reduce competition with an initially drastic innovation to achieve market entry well ahead of competitors, followed by incremental innovations to sustain their market position. Such a model would suggest the focus of innovation policy should be on stimulating innovation for all firms by first encouraging innovations by new entrants. These policies could focus on reducing barriers to entry and assisting new entrants, particularly in markets with few competitors or Cournot competition.

While the model in Chapter 6 considered Cournot and Bertrand competition, there is potential to consider a new type. Under Cournot competition, firms choose the quantity and quality they will produce to maximise profits based on the responsiveness of consumers to market prices. Similarly under Bertrand competition, firms choose the price and quality they will offer to maximise profits based on the response of consumer demand and therefore the quantity supplied that is required to clear the market. Alternatively, a new form of competitive equilibrium might exist where firms choose a particular price and output based on the consumers' preference for quality improvement to clear the market at that price and output. It is possible that the quality improvement of some varieties in the services sector are flexible and this competitive strategy and equilibrium could be the subject of future research.

The models in this thesis focused on understanding innovation within economic growth theory and its policy implications to support growth. However, perhaps an equally important topic of research is the welfare effects of these models and the welfare impacts of the policy recommendations. Spatial problems have particularly interesting welfare implications and the welfare distributions from the models in Chapters 4 and 5 would be important extensions of the research here. While the models lead to expectations that peripheral locations would typically be poorer than agglomerated regions, the ability to sustain specialised clusters allows portions of the population to maintain a higher income. Similarly, the inequality arising from imperfect markets in Chapter 6 would also suggest that smaller markets might be more unequal than larger markets. This also provides an interesting question for empirical research.

There are many other factors affecting innovation that could be examined that appear to support the findings in this thesis, steering innovation and economic policy in a distinctly localised direction. Therefore, this research agenda is of particular importance to researchers who are interested in the economic policy implications for peripheral locations. This is only a selection of related research questions that could be pursued following on from the work in this thesis.

7.3 Understanding innovation is the key to understanding economic growth

Modelling additional dimensions of innovation and discussing the implications for growth, economic policy and future research in this thesis also links into a wider academic debate about whether development should focus on places or people. A focus on people attempts to treat space as neutral. People are assumed to be mobile. This implies policy-makers should maximise opportunities for individuals, even if that means they move to alternative locations. However people are also not homogenous and their mobility behaviour shapes the places they choose to live. A focus on place recognises that distance and location are important and that not all people are mobile workers so emphasises creating opportunities in locations where people already live. The World Bank (2009) argues that policy approaches should be people-focused, because there is no compelling evidence to demonstrate that spatial equality leads to personal equality. Policy-makers that advocate place-based approaches argue that geography matters in all its dimensions including cultural, organisational and institutional aspects. In particular, they argue that a space-neutral approach to policy may have unintended spatial implications. Efficiency is boosted by policies which encourage agglomeration. On the other hand, peripheral locations may require policies which draw on local characteristics to address their development challenges. This suggests a need for localised, tailored policy based on a location's economic characteristics as do the models developed in this thesis. A policy focus on spatial characteristics should ultimately contribute to improving personal well-being.

On a similar note, the role of competition, related technology, education, trade, infrastructure and other economic characteristics also provide fruitful areas for research on innovation and growth. A combination of location, people and institutional approaches to understanding the complexity of innovation within the context of growth will provide policy-makers with a plethora of policy options to stimulate growth based on the characteristics, needs and preferences of their regions' residents. Overall, it is clear from the models developed in this thesis and from the research agenda above that understanding innovation really is the key to understanding economic growth.

References

Abramovitz, M. (1956). Resource and output trends in the united states since 1870. *American Economic Review*, 46:5–23.

Acemoglu, D., Aghion, P., and Zilibotti, F. (2006). Distance to frontier, selection, and economic growth. *Journal of the European Economic Association*, 4(1):37–74.

Acs, Z. and Varga, A. (2002). Geography, endogenous growth and innovation. *International Regional Science Review*, 25:132–148.

Acs, Z. J. (2002). *Innovation and the Growth of Cities*. Edward Elgar.

Acs, Z. J. (2009). *Universities, knowledge transfer and regional development: Geography, entrepreneurship and policy*, chapter Jaffe-Feldman-Varga: The search for knowledge spillovers. New Horizons in Regional Science. Edward Elgar.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and innovation: An inverted-u relationship. *The Quarterly Journal of Economics*, 120(2):701–728.

Aghion, P., Blundell, R., Griffith, R., Howitt, P., and Prantl, S. (2009). The effects of entry on incumbent innovation and productivity. *The Review of Economics and Statistics*, 91(1):20–32.

Aghion, P. and Griffith, R. (2005). *Competition and growth: reconciling theory and evidence*. MIT Press.

Aghion, P., Harris, C., Howitt, P., and Vickers, J. (2001). Competition, imitation and growth with step-by-step innovation. *Review of Economic Studies*, 68:467–492.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2):323–51.

Aghion, P. and Howitt, P. (1998). *Endogenous Growth Theory*, volume 1 of *MIT Press Books*. The MIT Press.

Aghion, P. and Howitt, P. (2005). Growth with quality-improving innovations: An integrated framework. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1 of *Handbook of Economic Growth*, chapter 2, pages 67–110. Elsevier.

Aghion, P. and Howitt, P. (2009). *The Economics of Growth*, volume 1 of *MIT Press Books*. The MIT Press.

Andrew, J. P., DeRocco, E. S., and Taylor, A. (2009). The innovation imperative in manufacturing: How the united states can restore its edge. Technical report, Boston Consulting Group.

Arrow, K. (1962). Economic welfare and the allocation of resources for inventions. In Nelson, R., editor, *The rate and direction of inventive activity*, chapter Economic Welfare and the Allocation of Resources for Inventions,. Princeton University Press.

Asheim, B. and Gertler, M. (2005). The geography of innovation: Regional innovation systems. In Fagerberg, J., Mowery, D., and Nelson, R., editors, *The Oxford Handbook of Innovation*. Oxford University Press.

Audretsch, D. B. and Feldman, M. P. (1996). R&d spillovers and the geography of innovation and production. *American Economic Review*, 86(3):630–40.

Audretsch, D. B. and Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In Henderson, J. V. and Thisse, J. F., editors, *Handbook of Regional and Urban Economics*, volume 4 of *Handbook of Regional and Urban Economics*, chapter 61, pages 2713–2739. Elsevier.

Bain, J. S. (1956). *Barriers to New Competition*. Harvard University Press, MA.

Baldwin, R. E. and Forslid, R. (2000). The core-periphery model and endogenous growth: stabilizing and destabilizing integration. *Economica*, 67(267):307–24.

Baldwin, R. E. and Forslid, R. (2010). Trade liberalization with heterogeneous firms. *Review of Development Economics*, 14(2):161–176.

Baldwin, R. E., Forslid, R., Martin, P., Ottaviano, G., and Robert-Nicoud, F. (2003). *Economic Geography and Public Policy*. Princeton University Press.

Baldwin, R. E. and Martin, P. (2004). Agglomeration and regional growth. In Henderson, J. V. and Thisse, J. F., editors, *Handbook of Regional and Urban Economics*, volume 4 of *Handbook of Regional and Urban Economics*, chapter 60, pages 2671–2711. Elsevier.

Baldwin, R. E., Martin, P., and Ottaviano, G. I. (2001). Global income divergence, trade and industrialisation: The geography of growth take-offs. *Journal of Economic Growth*, 6:5–37.

Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 82:1095–1117.

Barro, R. J. and i Martin, X. S. (2003). *Economic Growth*, volume 1 of *MIT Press Books*. The MIT Press, 2nd edition edition.

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica*, 60(4):889–917.

Bhaskar, V., Manning, A., and To, T. (2002). Oligopsony and monopsonistic competition in labor markets. *Journal of Economic Perspectives*, 16(2):155–174.

Bond-Smith, S. and McCann, P. (2014). Incorporating space in the theory of endogenous growth: Contributions from the new economic geography. In Fischer, M. M. and Nijkamp, P., editors, *Handbook of Regional Science*, pages 213–236. Springer Berlin Heidelberg.

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, 39(1):61–74.

Boschma, R. and Frenken, K. (2009). Technological relatedness and regional branching. *Papers in Evolutionary Economic Geography (PEEG)*, (0907).

Brakman, S. and Heijdra, B. J., editors (2001). *The Monopolistic Competition Revolution in Retrospect*. Cambridge University Press. Cambridge Books Online.

Brezis, E. S. and Krugman, P. R. (1997). Technology and the life cycle of cities. *Journal of Economic Growth*, 2(4):369–83.

Cappellin, R. and Wink, R. (2009). *International Knowledge and Innovation networks: Knowledge Creation and Innovaiton in Medium-technology clusters*. Edward Elgar.

Christopherson, S., Garretsen, H., and Martin, R. L. (2008). The world is not flat: putting globalization in its place. *Cambridge Journal of Regions, Economy and Society*, 1(3):343–349.

Coe, D. and Helpman, E. (1995). International r&d spillovers. *European Economic Review*, 39:859–887.

Crescenzi, R. and Rodríguez-Pose, A. (2013). R&d, socio-economic conditions, and regional innovation in the u.s. *Growth and Change*, 44(2):287–320.

Davis, D. R. (1998). The Home Market, Trade, and Industrial Structure. *American Economic Review*, 88(5):1264–76.

Desmet, K. and Rossi-Hansberg, E. (2009). Spatial growth and industry age. *Journal of Economic Theory*, 144(6):2477–2502.

Dinopoulos, E. and Thompson, P. (1999). Scale effects in schumpeterian models of economic growth. *Journal of Evolutionary Economics*, 9(2):157–185.

Dixit, A. (1980). The role of investment in entry-deterrence. *Economic Journal*, 90(357):95–106.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review*, 67(3):297–308.

Domar, E. (1946). Capital expansion, rate of growth, and employment. *Econometrica*, 14(2):137–147.

Dosi, G. (1988). Sources, Procedures, and Microeconomic Effects of Innovation. *Journal of Economic Literature*, 26(3):1120–71.

Duranton, G. (2007). Urban evolutions: The fast, the slow, and the still. *American Economic Review*, 97(1):197–221.

Dutta, S. and Lanvin, B., editors (2013). *The Global Innovation Index 2013: The Local Dynamics of Innovation*. Cournell University, INSEAD and WIPO, Geneva, Ithaca and Fontainebleau.

Eaton, J. and Kortum, S. (1999). International technology diffusion: Theory and measurement. *International Economic Review*, 40:537–570.

Eaton, J., Kortum, S. S., and Sotelo, S. (2013). International trade: Linking micro and macro. In Acemoglu, D., Arellano, M., and Dekel, E., editors, *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress, Volumes I, II, III.*, chapter II, 8. Cambridge University Press.

Eaton, M. L. and Kennedy, D. L. (2007). *Innovation in medical technology: Ethical Issues and challenges*. The John Hopkins University Press.

EIU (2009). *A new ranking of the world's most innovative countries*. Economist Intelligence Unit, London.

Fagerberg, J. (2005). Innovation: A guide to the literature. In Fagerberg, J., Mowery, D. C., and Nelson, R. R., editors, *The Oxford Handbook of Innovaiton*. Oxford University Press.

Fagerberg, J., Mowery, D. C., and Nelson, R. R., editors (2005). *The Oxford Handbook of Innovation*. Oxford University Press.

Faggian, A. and McCann, P. (2009). Human capital, graduate migration and innovation in british regions. *Cambridge Journal of Economics*, 33(2):317–333.

Feldman, M. P. and Audretsch, D. B. (1999). Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review*, 43:409–429.

Florida, R. (1995). Toward the learning region. *Futures*, 27(5):527 – 536.

Forslid, R. and Ottaviano, G. I. (2003). An analytically solvable core-periphery model. *Journal of Economic Geography*, 3(3):229–240.

Frankel, M. (1962). The production function in allocation of growth: A synthesis. *American Economic Review*, 52:995–1022.

Freeman, C. and Soete, L. (1997). *The Economics of Industrial Innovation, 3rd Edition*, volume 1. The MIT Press.

Frenken, K. and Boschma, R. A. (2007). A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process. *Journal of Economic Geography*, 7(5):635–649.

Fudenberg, D., Gilbert, R., Stiglitz, J., and Tirole, J. (1983). Preemption, leapfrogging and competition in patent races. *European Economic Review*, 22(1):3–31.

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

Fujita, M. and Thisse, J.-F. (2003). Does geographical agglomeration foster economic growth? and who gains and loses from it? *The Japanese Economic Review*, 54(2):121–145.

Garnaut, R. (2013). *Dog Days*. Black Inc, Australia.

Gaspar, J. and Glaeser, E. L. (1998). Information technology and the future of cities. *Journal of Urban Economics*, 43(1):136–156.

Gilbert, R. J. and Newbery, D. M. G. (1982). Preemptive patenting and the persistence of monopoly. *American Economic Review*, 72(3):514–26.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, 100(6):1126–52.

Grimes, A., Le Vaillant, J., and McCann, P. (2011). Auckland's knowledge economy: Australasian and European comparisons. Occasional Papers 11/2, Ministry of Economic Development, New Zealand.

Grossman, G. M. and Helpman, E. (1991a). *Innovation and Growth in the World Economy*. MIT Press Books. The MIT Press.

Grossman, G. M. and Helpman, E. (1991b). Quality ladders in the theory of growth. *Review of Economic Studies*, 58(1):43–61.

Grossman, G. M. and Helpman, E. (1991c). Trade, knowledge spillovers, and growth. *European Economic Review*, 35(2-3):517–526.

Grossman, G. M. and Helpman, E. (1994). Endogenous innovation in the theory of growth. *Journal of Economic Perspectives*, 8(1):23–44.

Grossman, G. M. and Helpman, E. (1995). Technology and trade. In Grossman, G. M. and Rogoff, K., editors, *Handbook of International Economics*, volume 3 of *Handbook of International Economics*, chapter 25, pages 1279–1337. Elsevier.

Hall, B. (2005). Innovation and diffusion. In *The Oxford Handbook of Innovation*. Oxford University Press.

Harris, C. and Vickers, J. (1985). Perfect equilibrium in a model of a race. *Review of Economic Studies*, 52(2):193–209.

Harrod, R. F. (1939). An essay in dynamic theory. *Economic Journal*, 49:14–33.

Helpman, E. (1992). Endogenous macroeconomic growth theory. *European Economic Review*, 36(2-3):237–267.

Hendy, S. and Callaghan, P. (2013). *Get off the Grass: Kickstarting New Zealand's Innovation Economy*. Auckland University Press.

Hollanders, H. and Es-Sadki, N. (2014). *Innovation Union Scorecard 2014*. European Commission.

Hoover, E. M. (1948). *The Location of Economic Activity*. McGraw-Hill.

Hunt, S. (1983). *Marketing Theory: the Philosophy of Marketing Science*. Richard D. Irwin, Inc., New York.

Isard, W. (1956). *Location and Space-economy; a General Theory Relating to Industrial Location, Market Areas, Land Use, Trade, and Urban Structure*. Technology Press of MIT and John Wiley, New York.

Jacobs, B., Nahuis, R., and Tang, P. J. (2002). Sectoral productivity growth and r&d spillovers in the netherlands. *De Economist*, 150(2):181–210.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–98.

Jones, C. I. (1995a). R&d-based models of economic growth. *Journal of Political Economy*, 103(4):759–84.

Jones, C. I. (1995b). Time Series Tests of Endogenous Growth Models. *The Quarterly Journal of Economics*, 110(2):495–525.

Jones, C. I. (1999). Growth: With or Without Scale Effects? *American Economic Review*, 89(2):139–144.

Kaldor, N. (1970). The case for regional policies. *Scottish Journal of Political Economy*, 17(3):337–348.

Keller, W. (2002). Geographic localization of international technology diffusion. *American Economic Review*, 92(1):120–142.

Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42(3):752–782.

Keller, W. and Yeaple, S. R. (2013). The gravity of knowledge. *American Economic Review*, 103(4):1414–44.

Kendrick, J. W. (1956). Productivity trends: Capital and labour. *Review of Economics and Statistics*, 38:248–257.

Kline, S. and Rosenberg, N. (1986). An overview of innovation. In Rosenberg, N., editor, *The Positive Sum Strategy: Harnessing Technology for Economic Growth*, pages 275–304. National Academy Press, Washington DC.

Krugman, P. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics*, 9(4):469–479.

Krugman, P. (1982). Trade in differentiated products and the political economy of trade liberalization. In *Import Competition and Response*, NBER Chapters, pages 197–222. National Bureau of Economic Research, Inc.

Krugman, P. (1991a). *Geography and Trade*, volume 1 of *MIT Press Books*. The MIT Press.

Krugman, P. (1991b). Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–99.

Krugman, P. and Venables, A. J. (1995). Globalization and the inequality of nations. *The Quarterly Journal of Economics*, 110(4):857–80.

Krugman, P. and Venables, A. J. (1996). Integration, specialization, and adjustment. *European Economic Review*, 40(3-5):959–967.

Lichtenberg, R. M. (1960). *One-tenth of a nation: National forces in the economic growth of the New York Region*. Harvard University Press, Cambridge, MA.

Lipsey, R., Bekar, C., and Carlaw, K. (1998a). The consequences of changes in gpts'. In Helpman, E., editor, *General Purpose Technologies and Economic Growth*, pages 15–54. MIT Press, Cambridge, MA.

Lipsey, R., Bekar, C., and Carlaw, K. (1998b). What requires explanation? In Helpman, E., editor, *General Purpose Technologies and Economic Growth*, pages 15–54. MIT Press, Cambridge, MA.

Lucas, R. J. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.

Marshall, A. (1890). *The Principles of Economics*. MacMillian, London, 1 edition.

Martin, P. (1999). Public policies, regional inequalities and growth. *Journal of Public Economics*, 73(1):85 – 105.

Martin, P. and Ottaviano, G. I. (1999). Growing locations: Industry location in a model of endogenous growth. *European Economic Review*, 43(2):281 – 302.

Martin, P. and Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review*, 42(4):947–68.

Martin, R. and Sunley, P. (1998). Slow convergence? the new endogenous growth theory and regional development. *Economic Geography*, 74(3):201–227.

McCann, P. (2007). Sketching out a model of innovation, face-to-face interaction and economic geography. *Spatial Economic Analysis*, 2(2):117–134.

McCann, P. (2008). Globalization and economic geography: the world is curved, not flat. *Cambridge Journal of Regions, Economy and Society*, 1(3):351–370.

McCann, P. (2009). Economic geography, globalisation and New Zealand's productivity paradox. *New Zealand Economic Papers*, 43(3):279–314.

McCraw, T. (2007). *Prophet of Innovation: Joseph Schumpeter and Creative Destruction*. The Belknap Press of Harvard University Press, Cambridge, MA.

McKeown, M. (2008). *The Truth about Innovation*. Pearson/Prentice Hall.

Nelson, R., editor (1993). *National Innovation Systems: A Comparative Analysis*. Oxford University Press, New York NY.

Nelson, R. and Winter, S. (1982). *An Evolutionary Theory of Economic Change*. Belknap of Harvard University Press.

Nijkamp, P. and Poot, J. (1993). Endogenous technological change, innovation diffusion and transitional dynamics in a nonlinear growth model. *Australian Economic Papers*, 32(61):191–213.

Nijkamp, P. and Poot, J. (1998). Spatial perspectives on new theories of economic growth. *The Annals of Regional Science*, 32(1):7–37.

OECD/Eurostat (2005). *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data*. The Measurement of Scientific adn Technological Activities. OECD Publishing, 3rd edition.

Ottaviano, G., Tabuchi, T., and Thisse, J.-F. (2002). Agglomeration and trade revisited. *International Economic Review*, 43(2):409–436.

Peretto, P. (1998). Technological change and population growth. *Journal of Economic Growth*, 3(4):283–311.

Rebelo, S. (1991). Long-Run Policy Analysis and Long-Run Growth. *Journal of Political Economy*, 99(3):500–521.

Reinganum, J. F. (1983). Uncertain innovation and the persistence of monopoly. *American Economic Review*, 73(4):741–48.

Reinganum, J. F. (1989). The timing of innovation: Research, development, and diffusion. In Schmalensee, R. and Willig, R., editors, *Handbook of Industrial Organization*, volume 1 of *Handbook of Industrial Organization*, chapter 14, pages 849–908. Elsevier.

Reiss, P. C. and Spiller, P. T. (1989). Competition and entry in small airline markets. *Journal of Law and Economics*, 32(2):S179–202.

Rodríguez-Pose, A. and Crescenzi, R. (2008). Mountains in a flat world: why proximity still matters for the location of economic activity. *Cambridge Journal of Regions, Economy and Society*, 1(3):371–388.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5):1002–37.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5):S71–102.

Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives*, 8(1):3–22.

Rosenberg, N. (1972). Factors affecting the diffusion of technology. *Explorations in Economic History*, 10(1):3–33.

Salant, S. W. (1984). Preemptive patenting and the persistence of monopoly: comment. *American Economic Review*, 74(1):247–250.

Samuelson, P. (1952). The transfer problem and transport costs: The terms of trade when impediments are absent. *Economic Journal*, 62:278–304.

Schumpeter, J. A. (1934). *The theory of economic development*. Harvard University Press, Cambridge, MA. Translated by Redvers Opie.

Schwab, K. and Sala-i-Martín, X., editors (2013). *The Global Competitiveness Report 2013-2014*. Geneva.

Segerstrom, P. S. (1998). Endogenous Growth without Scale Effects. *American Economic Review*, 88(5):1290–1310.

Simonen, J. and McCann, P. (2008). Firm innovation: The influence of R&D cooperation and the geography of human capital inputs. *Journal of Urban Economics*, 64(1):146–154.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1):65–94.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39(3):pp. 312–320.

Storper, M. and Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, 4(4):351–370.

Swan, T. W. (1956). Economic growth and capital accumulation. *Economic Record*, 32(2):334–361.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, 15(6):285 – 305.

Tirole, J. (1988). *The Theory of Industrial Organization*, volume 1 of *MIT Press Books*. The MIT Press.

Varga, A., editor (2009). *Universities, Knowledge Transfer and Regional Development: Geography, Entrepreneurship and Policy*. Edward Elgar, Cheltenham, UK.

Venables, A. J. (1996). Equilibrium locations of vertically linked industries. *International Economic Review*, 37(2):341–59.

Venables, A. J. (1999). The international division of industries: Clustering and comparative advantage in a multi-industry model. *Scandinavian Journal of Economics*, 101(4):495–513.

Vickers, J. (1985). Pre-emptive patenting, joint ventures, and the persistence of oligopoly. *International Journal of Industrial Organization*, 3(3):261–273.

World Bank (2009). World development report: Reshaping economic geography. Technical report, The World Bank, Washington DC.

Yamamoto, K. (2003). Agglomeration and growth with innovation in the intermediate goods sector. *Regional Science and Urban Economics*, 33(3):335–360.

Young, A. (1993). Invention and bounded learning by doing. *Journal of Political Economy*, 101(3):443–72.

Young, A. (1998). Growth without scale effects. *Journal of Political Economy*, 106(1):40–63.

Appendix A

Demand functions

This appendix derives the demand functions from Chapter 4. The demand functions in Chapters 3, 5 and 6 can be found using a similar approach modified for the specifications of each model. Chapter 3 does not include quality parameters because growth is in the form of an increasing variety of manufactured goods. Chapter 5 assumes zero transport costs for manufacturing goods and has multiple manufacturing sectors. Lastly, Chapter 6 also has multiple manufacturing sectors, involves a single region only and has no traditional goods sector.

Returning to the utility function from Chapter 4, consumer utility per period is given by the Cobb-Douglas function:

$$Q = C_M^{\mu} C_T^{1-\mu}, \qquad 0 < \mu < 1, \tag{A.1}$$

where C_T is the consumption in period t of the traditional good and C_M is the consumption of n and $\tilde{n}$ differentiated varieties of manufactured goods from the home and foreign regions respectively:

$$C_M = \left[\sum_{i=1}^{n} (A_i c_i)^{\frac{\sigma-1}{\sigma}} + \sum_{i=1}^{\tilde{n}} \left(\tilde{A}_i \tilde{c}_i \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \qquad \sigma > 1. \tag{A.2}$$

The variables with a tilde(˜) represent the characteristics of varieties produced in the foreign region and A_i represents the current quality of variety i. Each period consumers select the consumption bundle of all goods that maximises Q, subject to the budget constraint,

$$E \geq P_T C_T + P_M C_M, \tag{A.3}$$

where P_T is the price of traditional goods and P_M is a price index of manufactured goods.

The first stage of optimisation allocates expenditure between traditional and manufactured goods. The lagrangian is:

$$\mathcal{L} = C_M^{\mu} C_T^{1-\mu} - \lambda \left(P_T C_T + P_M C_M - E \right). \tag{A.4}$$

Differentiating, the first order conditions are given by:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial C_T} &= (1-\mu) C_M^{\mu} C_T^{-\mu} - \lambda P_T = 0, \\ \frac{\partial \mathcal{L}}{\partial C_M} &= \mu C_M^{\mu-1} C_T^{1-\mu} - \lambda P_M = 0, \qquad \text{and} \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= E - P_T C_T - P_M C_M = 0. \end{aligned} \tag{A.5}$$

The first two first order conditions imply:

$$\frac{1-\mu}{\mu} \frac{C_M}{C_T} = \frac{P_T}{P_M}. \tag{A.6}$$

Rearranging gives,

$$\frac{1-\mu}{\mu} P_M C_M = P_T C_T. \tag{A.7}$$

Substituting into the budget constraint yields how expenditure is allocated between manufactured and traditional goods:

$$\mu E = P_M C_M, \tag{A.8}$$

and

$$(1-\mu) E = P_T C_T. \tag{A.9}$$

The consumer spends a constant portion of her budget, μ, on manufactured goods and a constant portion of expenditure, $1-\mu$, on traditional goods.

The second stage of optimisation allocates expenditure between varieties of manufactured goods. The consumer now attempts to maximise equation A.2 subject to the budget constraint

$$\mu E \geq \sum_{i=1}^{n} P_i c_i + \sum_{i=1}^{\tilde{n}} \frac{\tilde{P}_i}{\tau} \tilde{c}_i, \tag{A.10}$$

where P_i represents the price of a variety in its location of production and c_i represents consumption of each individual variety i. In the models from chapters 3 and 4, transport costs are applied to manufactured goods that are exported to the other region. $\tau < 1$ represents the proportion of the final good that arrives at the destination. The remaining portion is used up during transportation. τ is a measure of the freeness of trade or an inverse index of transport costs. The lagrangian is given by:

$$
\begin{aligned}
\mathcal{L} &= C_M - \lambda\left(\sum_{i=1}^{n} P_i c_i + \sum_{i=1}^{\tilde{n}} \frac{\tilde{P}_i}{\tau}\tilde{c}_i - \mu E\right) \\
\mathcal{L} &= \left[\sum_{i=1}^{n} (A_i c_i)^{\frac{\sigma-1}{\sigma}} + \sum_{i=1}^{\tilde{n}} \left(\tilde{A}_i \tilde{c}_i\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} - \lambda\left(\sum_{i=1}^{n} P_i c_i + \sum_{i=1}^{\tilde{n}} \frac{\tilde{P}_i}{\tau}\tilde{c}_i - \mu E\right).
\end{aligned}
\tag{A.11}
$$

Differentiating using the chain rule, the first order conditions are:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial c_i} &= C_M^{\frac{1}{\sigma}} A_i^{\frac{\sigma-1}{\sigma}} c_i^{\frac{-1}{\sigma}} - \lambda P_i = 0, \\
\frac{\partial \mathcal{L}}{\partial \tilde{c}_i} &= C_M^{\frac{1}{\sigma}} \tilde{A}_i^{\frac{\sigma-1}{\sigma}} \tilde{c}_i^{\frac{-1}{\sigma}} - \lambda \frac{\tilde{P}_i}{\tau} = 0, \qquad \text{and} \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= \mu E - \sum_{i=1}^{n} P_i c_i - \sum_{i=1}^{\tilde{n}} \frac{\tilde{P}_i}{\tau}\tilde{c}_i = 0.
\end{aligned}
\tag{A.12}
$$

Solving the first two of the first order conditions for c_i and $\tilde{c}_i$ respectively yields:

$$
\begin{aligned}
c_i &= \lambda^{-\sigma} C_M A_i^{\sigma-1} P_i^{-\sigma} \qquad \text{and} \\
\tilde{c}_i &= \lambda^{-\sigma} C_M \tilde{A}_i^{\sigma-1} \left(\frac{\tilde{P}_i}{\tau}\right)^{-\sigma}.
\end{aligned}
\tag{A.13}
$$

Substituting into the third of the first order conditions and solving for $\lambda^{-\sigma}$ gives

$$
\lambda^{-\sigma} = \frac{\mu E}{C_M}\left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} - \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{1-\sigma}\right]^{-1}. \tag{A.14}
$$

Substituting into equation A.13 yields the demand functions of home region consumers for home and foreign goods respectively:

$$
\begin{aligned}
c_i &= \mu E A_i^{\sigma-1} P_i^{-\sigma}\left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} - \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{1-\sigma}\right]^{-1}, \qquad \text{and} \\
\tilde{c}_i &= \mu E \tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{-\sigma}\left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} - \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{1-\sigma}\right]^{-1}.
\end{aligned}
\tag{A.15}
$$

The remainder of this appendix defines the price index. Multiplying these demand functions by A_i and $\tilde{A}_i$ respectively and taking each to the power of $\frac{\sigma-1}{\sigma}$ gives:

$$
\begin{aligned}
(c_i A_i)^{\frac{\sigma-1}{\sigma}} &= (\mu E)^{\frac{\sigma-1}{\sigma}} A_i^{\sigma-1} P_i^{1-\sigma}\left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} - \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{1-\sigma}\right]^{-\frac{\sigma-1}{\sigma}}, \qquad \text{and} \\
\left(\tilde{c}_i \tilde{A}_i\right)^{\frac{\sigma-1}{\sigma}} &= (\mu E)^{\frac{\sigma-1}{\sigma}} \tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{1-\sigma}\left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} - \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{1-\sigma}\right]^{-\frac{\sigma-1}{\sigma}}.
\end{aligned}
\tag{A.16}
$$

Summing demand of all home and foreign varieties yields:

$$
\sum_{i=1}^{n} (c_i A_i)^{\frac{\sigma-1}{\sigma}} + \sum_{i=1}^{\tilde{n}} \left(\tilde{c}_i \tilde{A}_i\right)^{\frac{\sigma-1}{\sigma}} = (\mu E)^{\frac{\sigma-1}{\sigma}}\left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} - \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1}\left(\frac{\tilde{P}_i}{\tau}\right)^{1-\sigma}\right]^{\frac{1}{\sigma}}. \tag{A.17}
$$

Taking to the power of $\frac{\sigma}{\sigma-1}$ gives:

$$C_M = \mu E \left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} + \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1} \left(\frac{\tilde{P}_i}{\tau} \right)^{1-\sigma} \right]^{\frac{1}{\sigma-1}}, \tag{A.18}$$

which can be rearranged to:

$$\left(\frac{C_M}{\mu E} \right)^{\sigma-1} = \left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} + \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1} \left(\frac{\tilde{P}_i}{\tau} \right)^{1-\sigma} \right]. \tag{A.19}$$

Recalling the relationship in equation A.8, the price index is defined accordingly:

$$P_M = \left[\sum_{i=1}^{n} A_i^{\sigma-1} P_i^{1-\sigma} + \sum_{i=1}^{\tilde{n}} \tilde{A}_i^{\sigma-1} \left(\frac{\tilde{P}_i}{\tau} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{A.20}$$

Substituting into the demand functions from equation A.15 yields the demand functions given by equation 4.5 in Chapter 4.

These are the demand functions of home consumers for varieties produced in the home and foreign regions in response to prices in the region of production. However, the home region firm responds to the demand function of home consumers for home varieties and the demand function of foreign consumers for home varieties according to local prices. These two functions are

$$\begin{aligned} c_i &= \mu E A_i^{\sigma-1} P_i^{-\sigma} P_M^{\sigma-1}, \qquad \text{and} \\ c_i &= \mu E A_i^{\sigma-1} \left(\tilde{P}_i \right)^{-\sigma} \tilde{P}_M^{\sigma-1}, \end{aligned} \tag{A.21}$$

where foreign consumers respond to a local price index, $\tilde{P}_M$, $\tilde{P}_i$ now represents the price of variety i produced in the home region and exported to the foreign region and transport costs are treated as part of the cost of producing goods for export.

The remaining demand functions from the thesis can be found by following a similar approach, with the alterations that Chapter 3 does not include the quality parameters, Chapter 5 assumes zero transport costs, Chapters 5 and 6 hav multiple manufacturing sectors and Chapter 6 includes a single region only.