UNIVERSITY OF SOUTHAMPTON

Openness to Trade, Research and Development and Growth

Duncan McVicar

A thesis submitted in accordance with the requirement for the degree of Doctor of Philosophy

Department of Economics

September 1999

This thesis was submitted for examination in September 1999. It does not necessarily represent the final form of the thesis as deposited in the University after the examination.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL SCIENCE

ECONOMICS

Doctor of Philosophy

OPENNESS TO TRADE, RESEARCH & DEVELOPMENT AND GROWTH
By Duncan T. McVicar

This thesis comprises of four major papers concerning the effects of international trade on research and development (R&D) and growth. Taken together, the papers suggest the presence of significant links between openness to trade and growth through R&D. I argue that it is critical to establish the nature of such relationships if we are to correctly understand economic growth at a macroeconomic level. I contribute to the existing literature both by suggesting an alternative mechanism driving an integration/growth relationship and by providing new empirical evidence on some major questions of theoretical ambiguity.

By drawing on the industrial organisation (IO) literature concerning innovation and market structure, I demonstrate how economic integration can affect R&D and growth purely through intensity of competition effects. This contributes to the trade/growth literature by adding a new mechanism that does not rely on the usual scale effects and also provides further support for the IO literature that suggests there is a negative relationship between market concentration and R&D expenditure.

The overall theoretical ambiguity concerning the integration/R&D relationship coupled with the importance placed on R&D in many models of trade and growth implies a critical role for direct empirical evidence of the R&D/openness relationship. Given the notable lack of such evidence in the literature, I make a substantial contribution by examining the effect of openness to trade on R&D expenditure at industry level across a panel of OECD countries. This new evidence generally supports the existence of a positive openness/R&D relationship, although there is considerable variation across sectors. In addition, the exposition of estimation issues arising from the nature of the data coupled with the dynamic specification of the R&D equation is likely to prove a useful contribution to the foundations for further research.

The existing literature identifies knowledge spillovers as a key mechanism through which trade can affect technological progress and growth. I provide new evidence of the nature of knowledge spillovers in UK manufacturing, allowing for spillovers through foreign direct investment (FDI) in addition to trade in goods. FDI has not previously been included in such a sectoral-level study of spillovers and the discussion of the difficulties in estimating the model are likely to provide a useful starting point for further research. My results are somewhat counter-intuitive in that they suggest foreign R&D may have a negative effect on UK productivity in some sectors. A number of possible explanations for this are suggested, providing a further contribution to the burgeoning spillovers literature.

Contents

List of Figures

List of Tables

Preface

The vast majority of this work took place at the University of Southampton, where I was registered full-time and supported by the Economic and Social Research Council (Award No. R00429534230). This support is gratefully acknowledged. The latter stages of the research were carried out while I was at the Northern Ireland Economic Research Centre, Queen's University of Belfast.

An earlier version of Chapter 3 of this thesis was presented at the University of Southampton in 1997 and the 1998 International Conference on Economic Integration and Transformation at the University of Toronto. Material from Chapter 4 was presented at seminars at the Queen's University of Belfast in 1998 and the 1999 Royal Economic Society Conference in Nottingham. Some parts of Chapter 5 were presented at the 1999 Econometric Society European Meeting in Santiago de Compostela. I am grateful for participants at these seminars for their interest and comments.

Acknowledgements

First and foremost, I would like to thank John Driffill, my supervisor, for his patient advice and guidance over the last four years. His attention to detail and his early encouragement to write things down as I went along have been of invaluable help throughout the research.

A large number of other members of the Economics Department at Southampton have also helped with advice at various stages of the research, for which I am very grateful. In particular, I must mention Jan Podivinsky for his thirst to advise on matters of econometrics and Sally Srinivasan for advice on panels.

Bronwyn Hall and Clint Cummins were very helpful with programming queries, despite the likelihood that they receive thousands of such questions every year. Thanks are also due to those at NIERC who have had to put up with the occasional bout of thesis-related distraction on my part over the last year.

I am especially grateful to Jess for being the perfect antidote to late night research-related worries and for proof reading. Finally, the support and occasional nagging of my parents has been a constant help, and the phrase 'just get on with it' is now firmly lodged in my memory. Thanks.

Chapter 1

Introduction

In little more than a decade, the economic growth literature has moved from a position of relative stagnation to being a vibrant and increasingly fruitful field of research. Widespread dissatisfaction with the neoclassical growth model has led to the adoption of the new paradigm of endogenous growth alongside. The usual mechanism through which long run growth is achieved in this literature is technological progress driven by innovation carried out by profit maximising agents. Subsequent advances in this literature, such as the consideration of knowledge spillovers and imitation, have brought us to the position of having models that can both explain long run growth and predict convergence.

The background to the development of the endogenous growth literature is one of increasing world trade and increasing openness to trade, both globally and regionally. The literature has not been slow to incorporate this international trade environment in models of R&D driven growth and in more general studies of growth rates.

This thesis provides an assessment of where we stand in terms of explaining economic growth through technological progress and its relationship with international trade. Whilst a great deal of progress has been made over the last decade in this respect, I argue in Chapter 2 that a number of significant gaps remain in our understanding along the path of the development of this literature. In addition to driving this literature forward, I argue that it is also necessary to back up and fill in some of these gaps in order to provide a firmer foundation of understanding on which to proceed. The next three Chapters attempt this for selected topics within the framework of the trade/R&D literature. Chapter 3 considers one such issue in how economic integration can affect R&D expenditure purely through competition intensity effects. Chapters 4 and 5 identify critical parts of the literature where further empirical evidence is essential and attempt to make a contribution to this evidence, firstly on the nature of the relationship between R&D incentives and openness to trade, and secondly on the nature of knowledge spillovers. Chapter 6 concludes.

In the next Chapter I present a comprehensive review of the endogenous growth literature based on R&D undertaken by profit maximising agents. This literature is set against the background of dissatisfaction in the 1970s and 1980s with the neoclassical growth model due to Solow and built upon by others. The primary cause of this dissatisfaction was the failure of the neoclassical model to predict long run growth from within the model. In short, the model implies that once an economy has reached its steady state level of capital, further growth is only possible by assuming *exogenous* technological progress.

The pioneering research heralding the onset of endogenous growth through R&D was carried out by Romer in the 1980s, although it has recently been suggested that the fundamental breakthrough had been made in the literature twenty years previously.[1] Romer's first model had growth driven by knowledge created as an external by-product of the accumulation of broad capital. Romer's subsequent papers built on this breakthrough by introducing intentional knowledge creation, resulting from investment in R&D by profit maximising agents in return for monopoly rents on new product varieties. The crucial step was to drop the assumption of perfect competition.

Two characteristics of the Romer model have attracted a great deal of attention in the literature. Firstly, the model fails to predict convergence, which is one of the most attractive features of the neoclassical model and has considerable empirical support. It is not until international trade and the possibility of imitation are introduced to the basic model that the model can predict convergence. Secondly, the model predicts that larger economies will grow faster. These *scale effects* have little empirical support. They also have very important implications in the context of international trade.

Two special features of the technological knowledge considered by Romer are that it is not perfectly appropriable and that it is not perfectly excludable. In other words, an agent cannot completely prevent another agent from also using his or her ideas at any

[1] Aghion and Howitt (1998), whilst not detracting from the importance of the contribution made by Romer (1986), argue that in many respects, the Romer paper was preceded by Frankel (1962).

point in time. This is the basis by which the introduction of an international dimension can lead to convergence, by allowing follower countries to catch up by temporarily growing at a faster rate than leading countries. This can occur either by the follower directly imitating the leaders' designs or by learning from the knowledge capital accumulated from R&D in the leading countries. I discuss both the knowledge spillovers literature and the imitation literature in some detail, given their critical role in reconciling the Romer model with evidence of conditional convergence. I argue that our understanding of knowledge spillovers is limited, particularly by the failure of the majority of the empirical literature to allow for alternative spillover mechanisms in addition to trade in physical goods. I return to this topic in Chapter 5.

The Romer model's prediction of scale effects has interesting implications when international trade is introduced, particularly when economic integration is considered. I assess the state of this literature in Chapter 2, in the light of the lack of firm empirical evidence in support of such scale effects. Two alternative conclusions can be reached. Firstly, scale effects cannot be found because they don't exist in practice. Secondly, scale effects cannot be found because the available data are not of sufficient quality to do so. The former hypothesis has critical implications for models of R&D driven growth and economic integration. In Chapter 3, I present an alternative model in which integration can affect R&D driven growth purely through its effects on the intensity of competition.

In defence of models that predict faster growth with increased trade, through the mechanism of R&D, I discuss the empirical growth literature. Although this literature is generally supportive of a positive growth/trade liberalisation relationship, consistent with the internationalised Romer model, I argue in Chapter 4 that in addition, more *direct* evidence of such a relationship is needed.

Chapter 3 begins from the premise that economic integration might lead to dynamic efficiency gains, which potentially could imply higher long run growth. The most generally used arguments for this hypothesis are firstly, that integrating a number of economies enlarges the size of the market available to successful innovators and secondly, that integration increases the size of the knowledge base available to firms. These are both examples of *scale effects*, where a larger economy will grow faster than a smaller one, other things being equal. As discussed in Chapter 2 however, the evidence for the existence of scale effects is weak.

At the same time as enlarging the market available to successful innovators and increasing the knowledge base to which a firm has access, integration will also alter the nature of competition facing firms within the integrating markets. Although such *dynamic* intensity of competition effects have not gone unnoticed by the integration literature, they have not received the amount of attention they deserve, especially in the context of the lack of evidence for scale effects. Such effects are not scale effects in the sense that market enlargement is a scale effect. In other words, there is no reason to suppose a larger country will grow faster than a smaller one. Rather, integrating two existing economies, with an existing number of firms, will intensify the degree of competition faced by these firms because they are left with a smaller share of a larger market.

The relationship between R&D and the intensity of competition has received a great deal of attention in the IO literature, however. Chapter 3 draws on this literature by building a model of an imperfectly competitive (differentiated) industry in which to examine the effects of the intensity of competition, but spread over two economies. Increased competition resulting from integration of these two economies makes a firm's market share more sensitive to prices, whilst not necessarily affecting the firm's actual size. A symmetric equilibrium results where R&D is increasing with integration. This is because of the increased competitive threat, or intensified *strategic* incentives to both steal competitor's market share and to protect own market share.

Two alternative scenarios are then considered as extensions to the basic model. Firstly, co-operation in R&D in the form of a research joint venture (RJV) removes strategic incentives and induces firms to hold back the pace of innovation. Secondly, if the extent of knowledge spillovers between firms across countries is positively related to the degree of integration of those countries, then the relationship between R&D incentives and integration becomes non-monotonic. In this case, there is an optimal degree of integration.

The model is then extended to an n-sector general equilibrium model in which integration speeds up long run growth purely because of this competition effect. I argue that this model, although somewhat specialised, makes a significant contribution to the literature on integration, R&D and growth by moving away from the reliance on scale effects. In the opposite direction, the model contributes to the IO literature by extending a standard Cournot model in a novel way to allow for integration across markets.

Chapter 4 is a major empirical study which aims to address an apparent gap in the literature between empirical evidence from macro growth regressions and micro models of integration and R&D. An extensive literature exists providing evidence of a generally positive relationship between growth rates and openness to trade. There are also several models of integration and R&D that provide a strong potential explanation for this relationship. However, there is little *direct* empirical evidence of the nature of the relationship between R&D incentives and openness to trade. If the R&D story is to be widely accepted as an important explanation of the apparent positive relationship between growth and openness to trade, then I argue that such direct evidence is an essential part of the picture.[2]

A significant part of Chapter 4 discusses the difficulties involved in attempting to establish empirical evidence for such a relationship. The first problem is that moving from models of R&D and trade to estimable empirical models is far from straightforward as the theoretical models in the literature tend to be highly specialised. Therefore empirical models must be ad hoc to some extent, although there are precedents which can provide guidance. Secondly, R&D incentives and openness to trade are both inherently difficult to measure. Whilst I am happy to blur the distinction between incentives and easily measurable expenditure in this context, accurately measuring openness is a more immediate difficulty, and one that has attracted considerable attention in the literature.

[2] Henrekson et al (1997), for example, find a positive long run growth effect of membership of the European Union. Their explanation for this largely relies on models predicting a stimulus to R&D incentives from economic integration.

I use a time-series cross-section data set at manufacturing sector level across 14 OECD countries between 1973 and 1992. Treating each sector as a separate panel of countries, I estimate a simple dynamic R&D equation with openness to trade as an explanatory variable, proxied by trade intensity. The nature of the data (narrow cross section and fairly short time dimension) coupled with the dynamic nature of the R&D equation implies that no estimator will be ideal in the sense of being both unbiased and efficient. This problem with dynamic panel data models has attracted considerable attention in the literature. Given this level of interest, Chapter 4 provides a useful case study of these estimation problems, in a time-series cross-section of typically small dimensions.

A further complication with such a study is the question of the stationarity or otherwise of the series in the empirical model. With such a short time dimension it is difficult to establish the presence or otherwise of unit roots even with panel tests. I argue therefore that the model should be estimated first assuming stationary series and second dropping this assumption. The first part of this suggestion is carried out, assuming stationarity for all the series. I also estimate the model in first differences as a quick test of robustness to the possible presence of unit roots. More sophisticated estimation of the model assuming non-stationary series is left for further research.

Given the difficulties involved in estimating the model in Chapter 4 the results are unsurprisingly mixed. There is clear evidence of a positive R&D/openness to trade relationship in the standard least squares dummy variables (LSDV) model assuming stationarity. There is a danger that the significant effects found are spurious if some of the series are not truly stationary. However, there is still some evidence of such a positive relationship when the model is estimated in first differences. Overall, the Chapter makes a substantial contribution to the literature both by providing tentative evidence of the existence of a direct R&D/openness relationship and as a study of the practical difficulties involved in the hunt for this evidence.

Chapter 5 moves away from direct consideration of the R&D/openness to trade relationship. Instead, the Chapter focuses on knowledge spillovers, which, as I argue in Chapter 2, are of critical importance for models of trade and growth at both micro and macro levels. I argue that a fuller understanding of the nature of knowledge spillovers is needed if we are to rely on them to such an extent in broader models of trade and growth.

Numerous studies have found evidence of knowledge spillovers at various levels of aggregation. Recently, for example, the work of Coe and Helpman has been instrumental in driving the debate forward at a macro level. However, the Coe and Helpman study, and the majority of other studies in the field, assume that technological knowledge is only transmitted through trade in physical goods between countries, industries or firms. Alternative transmission mechanisms, such as FDI, have been widely suggested but little researched.

In Chapter 5, I empirically examine the nature knowledge spillovers in a panel of UK manufacturing sectors allowing for technological transmission through direct investment in addition to trade in goods. I first present a simple modification to a standard 'variety of intermediates' model that allows for spillover effects through FDI. There are considerable difficulties involved in estimating this model, perhaps the most important of which is the high degree of correlation between variables constructed to proxy for FDI-transmitted knowledge and those constructed to proxy for trade-transmitted knowledge. In these circumstances, it is difficult to be confident that the effects identified by the empirical model are assigned in the correct way to the alternative transmission mechanisms.

However, considerable evidence is found that suggests the effect of knowledge spillovers on productivity in UK manufacturing is not always positive. In fact, for some sectors, both FDI-transmitted spillovers and international trade-transmitted spillovers appear to have a significant *negative* effect on domestic productivity. Equally, there are some sectors for which positive international spillover effects are found. These tend to be the more capital intensive, higher technology engineering sectors. These results, although sensitive in some cases in terms of magnitude, are robust at least in sign to a number of specification changes of the model. An initial conclusion is that FDI appears to play an important role in knowledge spillovers in the UK manufacturing sector.

Two alternative suggestions are made for the apparent negative relationship between foreign R&D and domestic productivity. Firstly, the UK's position as a leading technology source country makes it unlikely to gain from R&D in technologically inferior countries and makes it vulnerable to *outward* spillovers, which might act as a disincentive to invest in technological advancement. Secondly, increasing competition with technologically superior economies, such as the US and Germany, might have a negative effect on incentives to invest in R&D in the UK which is *not* linked to knowledge spillovers. The first explanation has been suggested in existing literature in the context of apparently insignificant or negative spillover effects in other countries. The second explanation has also been suggested elsewhere in the literature, but not in this context.

The results of Chapter 5 provide support for those who argue spillovers may not have consistently positive effects on productivity growth and for those who argue that other relationships may interfere with our attempts to quantify such spillover effects. The Chapter also provides a detailed discussion of the estimation problems associated with such studies, which should be a useful source of guidance for further related research. Although the Chapter was originally intended to be somewhat separate from the preceding Chapters of the thesis, themes from Chapters 3 and 4 are critical in the explanation of the results. In particular, the relationship between intensity of competition and technological progress and the negative incentive effect of spillovers discussed in Chapter 3 both appear as possible explanations of the apparent negative foreign spillover effects found for some UK manufacturing sectors.

Chapter 6 summarises and makes concluding remarks. Suggestions for further research arising directly from the issues discussed in this thesis are collected together. Where it is possible to identify broad policy implications of the research, they are briefly discussed.

Chapter 2

A Survey of the Literature on R&D-Driven Endogenous Growth and International Trade

The field of economic growth has reawakened (Gregory Mankiw, July 1994)[3]

2.1. Background

2.1.1. Introduction

The importance of understanding economic growth is undeniable given the widely different growth rates displayed by different economies across the world. If we can explain why some countries grow faster than others, we may be able to influence policy to improve people's well-being in both advanced and developing countries. In advanced economies in particular, growth has been characterised by the continual introduction of new and better products and processes. These innovations are not manna from heaven, but the result of purposive human endeavour (Aghion & Howitt, 1998). Further, this process of technological advancement has taken place against a background of a world economy that is increasingly open and interdependent. This Chapter reviews the literature on technological progress, growth and their relationship with increasing international trade.

The growth literature is extensive, both on the theory side and the empirical side. This literature also goes back a long way. Indeed, many of the elements of modern growth theory have their roots in the work of the classical economists of the 18th and 19th Centuries. The critical roles played by capital accumulation and technical progress, for example, were recognised very early on:

[3] pp xv, Barro & Sala-I-Martin (1995).

The annual produce of the land and labour of any nation can be increased in its value by no other means but by increasing either the number of its productive labourers, or the productive powers of those labourers who had before been employed....The productive powers of the same number of labourers cannot be increased, but on consequence either of some addition and improvement to those machines and instruments which facilitate and abridge labour; or of a more proper division and distribution of employment (Adam Smith, 1776).[4]

From these early foundations, growth theory progressed through the Harrod-Domar model to the neoclassical model of Solow and others and more recently to models of endogenous growth. The most exciting of these recent endogenous growth models are those that explicitly model technological progress as the result of deliberate actions by profit-seeking agents. The model of Romer (1990) provides the benchmark for this modern growth theory. The empirical growth literature has developed alongside this theoretical literature, although it is only recently, with the availability of detailed cross-country data sets that this literature has really taken off.

The inter-relations between the recent models of R&D-driven endogenous growth and international trade have proved a particularly vibrant branch of research in the 1990s. Grossman and Helpman (1991a) notes that the market incentives to which innovators respond are invariably related to aspects of the international trade environment. A number of models, building on the endogenous growth literature, have been presented to capture these relationships. Empirically, openness to trade has been found to be a significant explanatory factor for growth rates in a number of studies. Considerable attention has also been paid to the possibilities for knowledge spillovers and imitation across countries.

The growth and trade literature continues to develop at a very fast rate, which makes a detailed assessment of 'where we are now' inherently difficult to construct. However, it is useful to highlight just a few aspects of the recent literature in order to get a sense of the progress made in the last few years. First, models have recently been presented in which both capital accumulation and endogenous technological progress can drive growth. Such models begin to capture the idea that endogenous growth models and neoclassical growth models can be complements and not substitutes. Second, recent

[4] Smith, A. (1776). 'The Wealth of Nations,' pp306-307.

advances in empirical growth research have allowed us to assess the relevance of the various theoretical approaches to the real world in some detail. This empirical work has had a strong feedback effect on the development of new models in the theoretical literature. Third, both theoretical models and empirical evidence have been presented suggesting a strong relationship between growth and openness to trade. Fourth, evidence suggests the existence of significant knowledge spillovers across countries. These spillovers can be used in imitation models to explain convergence in the context of endogenous growth models. Overall, the state of the literature is one of rapid development, but also perhaps of movement towards a partial re-integration of the endogenous growth and neoclassical growth paradigms, which in many ways is dependent on the consideration of international trade-related issues.

As well as presenting a review of the existing literature, the purpose of this Chapter is to identify where research in the growth and trade field might go next. In practice, research is likely to continue down many of the current roads for some time, given the infancy of much of this research. For example, the knowledge spillovers literature has only just begun to try to explain and assess how technology spreads from one country to another, a topic I return to in Chapter 5. From a grander strategic point of view, the literature may be moving in the direction of drawing together aspects of the neoclassical and endogenous growth models, in an international trade context, into models endogenously explaining long run growth, allowing for sophisticated dynamics around steady states (Aghion and Howitt, 1998). Recent research that suggests spillovers have heterogeneous effects on different countries, coupled with evidence of 'convergence clubs' presents an exciting avenue for future research. Most importantly, however, I argue in this thesis that while carrying the growth and trade literature forward, it is important to step back a little from these developments and identify where the foundations might need shoring up. These gaps are likely to be inevitable when the development of the literature is so rapid. I identify two such gaps, which are discussed in Chapters 3 and 4 in more detail. The first is a need for alternative models of integration and growth that do not rely on the usual scale effects, for which evidence is inconclusive. The second is the need for direct micro evidence of the relationship between R&D incentives and openness to trade.

The remainder of this Chapter is set out as follows. The next section presents a brief outline of empirical regularities, or 'stylized facts', that are useful to illustrate the growth and trade literature. Section 2.2 outlines the neoclassical growth model and traces the development of the endogenous growth literature through from early AK models to the Romer (1990) model of R&D-driven endogenous growth. Section 2.3 reviews the macro-level empirical evidence in support or otherwise of the various models of economic growth. Staying at a macro level, international trade is introduced in Section 2.4. Section 2.5 reviews the microeconomic literature concerning R&D and market structure and international trade. Section 2.6 considers knowledge spillovers and imitation. Finally, Section 2.7 concludes with a discussion of areas for further research, and a perspective of how the following Chapters fit into this research agenda.

Finally, it is worth taking a moment to discuss what this Chapter does not consider and what it is not intended as. First of all, it is not intended as a detailed review of the growth literature to date. The volume of research is too great for such a task to be undertaken lightly, especially as the primary purpose of this Chapter is to provide background for the following Chapters. In any case, excellent reviews of different aspects of the literature are provided elsewhere.[5] Neither is the purpose to provide a detailed review of the narrower literature on growth and trade, such as that provided by Grossman and Helpman (1991), although this would be a worthwhile exercise given the considerable developments over the last eight years. Finally, this Chapter is not intended as a review of micro-level models of technological progress or micro-level evidence of technological progress in general. I do not consider issues of internal financing, patents, public sector research, or any aspects of the labour market that might affect R&D, for example.[6] Rather, in the context of the wider literature, the Chapter is intended to highlight certain areas where the micro-foundations of the trade and growth literature would benefit from further attention.

2.1.2. Some Empirical Regularities

[5] Barro and Sala-I-Martin (1995) and Aghion and Howitt, (1998), for example.

[6] Neither is there any explicit consideration of the service sector in the analysis. This does not reflect any perceived lack of importance or interest, but is adopted in order to narrow the focus of the analysis to those papers which have had the greatest impact in the growth literature, which have been largely manufacturing based.

In order to illustrate some of the issues that will be discussed in the following sections, it is useful to present some stylized facts, with a few examples, that any analysis of technological progress, growth and trade should aim to explain. In this, I am following the approach of many recent surveys of the growth literature, or parts thereof (see, for example, Barro and Sala-I-Martin, 1995; Aghion and Howitt, 1998; Temple, 1999). In order of discussion, the main empirical regularities considered are firstly that economies tend to grow over time, secondly that growth rates differ across countries, thirdly that capital per worker grows and the return to capital is almost constant. In addition, I briefly discuss the evidence of convergence between countries, the patterns of R&D expenditure and productivity over time and how this all fits with the pattern of international trade over time. These issues are discussed only in terms of broad regularities in the raw data. Discussion of more detailed empirical analyses is left to later sections.

The starting point for this analysis is Kaldor (1963), who presented a list of empirical regularities relating to growth that he called stylized facts. Aghion and Howitt (1998) note that these stylized facts are often easily falsified with various time-series techniques, yet they are still widely used in the literature. The reason is that Kaldor's stylized facts characterise the steady state, or the *long run* path of the economy. This steady state, which is the standard analytical tool with which to examine growth in the long run, is characterised by output, capital and wages all growing at the same constant rate.[7]

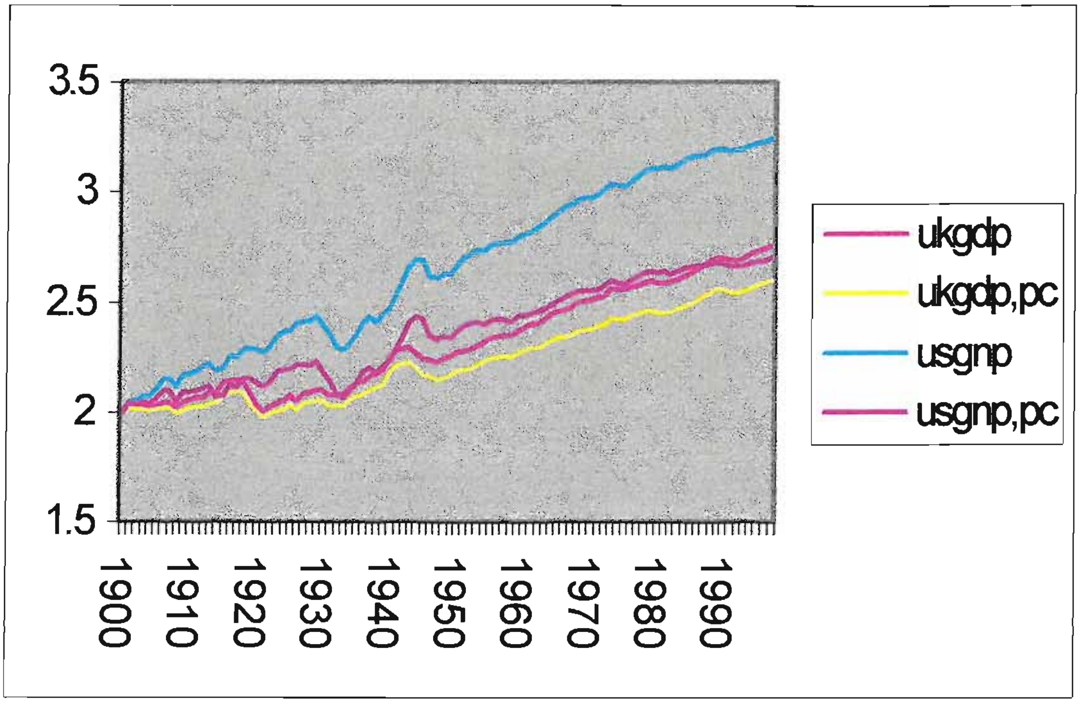
1. Output Grows Over Time

Kaldor's first observation was that output tends to grow over time and its growth rate does not tend to diminish. In other words, economies are characterised by continual long run growth, and are not just fluctuating around a stationary mean level of output, or converging towards such a static level. This is true both in absolute terms and in per capita terms. This simple fact causes serious difficulties with the neoclassical growth

[7] Aghion and Howitt (1998) add knowledge, appropriately measured, to this list.

model, as discussed in Section 2.2. Figure 2.1 below illustrates this long run output growth for the UK and the US from 1900-1997.

Figure 2.1: Log Output and Log Output Per Capita, Constant Prices, UK & US



Notes: UK figures are log of GDP at factor cost, constant prices, 1900 GDP=100. US figures are log of GNP at constant prices, 1900 GNP=100. Per capita series are divided through by mid-year population figures. Sources: The Economist: Economic Statistics 1900-1983, London; ONS Annual Abstract of Statistics, various years, HMSO/TSO, London; Statistical Abstract of the United States, various years, US Dept. of Commerce.

Both the US and the UK display clear growth over the last hundred years (and earlier) with no sign of any long-term reduction in the growth rate. Over the period, UK GDP per capita has grown at an average of just over 1.4 per cent per annum. US GNP per capita has grown at an average of just under 1.7 per cent per annum. The importance of even small differences in long run growth can be illustrated by comparing these two growth rates. UK GDP per capita has grown fourfold since 1900 whereas US GNP per capita is now six times what it was in 1900. Fluctuations around the rising long run trend in output are clear, particularly the contrast between the depression of the 1930s and the rapid growth associated with the 2nd World War.

2. Growth Rates Differ Across Countries

It is clear that income levels vary substantially across countries (hence the division into developed and developing countries). However, it is also the case that growth rates differ substantially across countries.[8] For example, Singapore experienced an average growth rate of 6.3 per cent between 1960 and 1990, compared to an average growth rate of –2.1 per cent in Iraq (Barro & Sala-I-Martin, 1995). Such big differences in growth rates will have huge accumulated effects on income levels over time. At these growth rates, if Iraq and Singapore had the same initial income in 1960, Singapore would be nearly 12 times richer in 1990 than Iraq.

Table 2.1 shows average growth rates of output per worker for a selection of countries over the period 1960-1990 taken from Temple (1999). These countries have been selected to illustrate the extremes of the growth rates distribution, or 'growth miracles' and 'growth disasters'. The Table illustrates the well known pattern of countries in Sub-Saharan Africa performing badly over the period and countries in East Asia performing well.

Table 2.1: Average Annual Growth Rates of Output per Worker, 1960-1990

Miracles	**Growth Rate**	**Disasters**	**Growth Rate**
Korea	6.1	Ghana	-0.3
Botswana	5.9	Venezuela	-0.5
Hong Kong	5.8	Mozambique	-0.7
Taiwan	5.8	Nicaragua	-0.7
Singapore	5.4	Mauritania	-0.8
Japan	5.2	Zambia	-0.8
Malta	4.8	Mali	-1.0
Cyprus	4.4	Madagascar	-1.3
Seychelles	4.4	Chad	-1.7
Lesotho	4.4	Guyana	-2.1

Source: Temple (1999).

[8] These two facts are clearly not entirely unrelated.

One of the main challenges of the study of growth is to understand why different countries grow at such different rates. At the start of this Chapter, I argue that by improving our understanding of this question, we may be able to influence policy in such a way as to make a significant difference to the well being of people in many different countries. A great deal of progress has already been made towards answering this question, and Section 2.3 reviews some of the evidence.

3. Convergence?

Whether countries are converging in terms of income levels is a frequently asked question in the growth literature. In other words, do poorer countries grow faster than richer ones? I leave aside conditional convergence (to different steady states) for now and just consider this question of absolute convergence, although both are discussed in more detail in Section 2.3. Generally speaking, the evidence for convergence in income levels is mixed. Table 2.1 suggests that some of the initially poorest countries also display the lowest growth rates. For example, Ghana and Chad had GDP per capita of just 9 per cent and 7 per cent of US GDP per capita in 1960 but display some of the lowest average growth rates between 1960 and 1990. In contrast, many more developed countries appear to be converging to a common level of income. Table 2.2 shows GDP per capita in 1950 and 1990 for five such countries.

Table 2.2: Converging Advanced Economies

	1950 GDP pc	1990 GDP pc	Rank 1950	Growth Rank 1950-1990
France	3828	13931	3	3
Germany	3094	14487	4	2
Japan	1278	14827	5	1
UK	4959	13066	2	4
US	7091	18399	1	5

Notes: Figures in 1990 $US, based on purchasing power parities (PPPs). Source: The Economist 'Economic Statistics 1900-1983', Barro & Sala-I-Martin (1995).

It is generally accepted that there is a pattern of convergence among already industrialised countries and among regions within countries, such as the US or Japan (see, for example, Barro & Sala-I-Martin, 1995; Temple, 1999). This pattern of convergence disappears when African countries are included in the sample, however. Quah (1997) suggests there is a pattern of convergence 'clubs', with countries becoming stratified into two distinct income classes over time. In other words, already rich countries display convergence to one income level and already poor countries display convergence to another lower level. Some countries in the middle display convergence to the high-income group and others display convergence to the low-income group. For example, Thailand and Malaysia would be in the former category and Senegal and Mozambique in the latter category, as shown by Table 2.3.

Table 2.3: Convergence Clubs?

	GDP per capita/US GDP per capita, 1960	GDP per capita/US GDP per capita, 1990
France	.61	.76
Thailand	.09	.13
Malaysia	.14	.17
Senegal	.10	.04
Mozambique	.12	.03
Mali	.05	.03

Source: Barro & Sala-I-Martin, 1995.

4. Capital per Worker Grows Over Time and its Rate of Return is Constant

Four of Kaldor's stylized facts are all related to physical capital. Firstly, physical capital per worker grows over time. Secondly, this growth rate is similar to the growth rate of output per worker so that the ratio of capital to output is nearly constant. Thirdly, the rate of return to capital is nearly constant and lastly, the shares of labour and capital in national income are nearly constant. Numerous references for detailed discussion of these points are given in Barro and Sala-I-Martin (1995).

The first two facts are very important observations for growth theory as they suggest the accumulation of capital and a correlation between capital accumulation and output growth. Information on capital stocks is difficult to obtain back a long way, and is likely to be far from accurate where it does exist. However, the OECD publishes annual estimates of capital stocks from 1970 onwards for its member countries. Table 2.4 below shows the growth in real capital stocks between 1970 and 1995 for the UK, US, France and Germany.

Table 2.4: Growth of Physical Capital

	Real Net Stock of Capital 1995, 1970=100	Real GDP 1995, 1970=100
USA	199 (1993)	179 (1993)
France	238	183
Germany	208 (1994)	192 (1994)
UK	175 (1991)	152 (1991)

Notes: Capital stock figures for USA, Germany and UK are for 1993, 1994 and 1991 respectively.
Sources: OECD Flows and Stocks of Fixed Capital, 1996; OECD National Accounts Main Aggregates.

The above Table clearly shows the growth in capital stock at constant prices over time, which has been of roughly equal magnitude in the four countries although France has perhaps seen more rapid capital accumulation. The right hand side column compares output growth over the same period, which in all four cases is reasonably close to the growth of capital, although there is an apparent pattern of slightly faster capital accumulation than output growth.[9]

Is the rate of return to capital stable over time? The evidence for this is mixed, with the UK displaying a fairly stable pattern but the US and other countries displaying a declining trend in real interest rates over time. Barro and Sala-I-Martin (1995) argue that the long term constancy pattern should be replaced by a tendency for returns to capital to decrease over time, which suggests transition towards a steady state that has not yet been reached.

Finally, Kaldor (1963) argued that the shares of labour and capital in national income were roughly constant over time. In other words, as capital accumulates, output per worker grows and wages rise roughly in proportion. Barro and Sala-I-Martin (1995)

[9] Barro and Sala-I-Martin are happy that the capital/output ratio remains reasonably constant over time, at least for developed countries (see Barro & Sala-I-Martin, 1995).

believe the weight of evidence supports this observation. Productivity is discussed in more depth below.

5. Technological Progress

Aghion and Howitt (1998) add knowledge, appropriately measured, to the list of series that grow in steady state. This concept of knowledge differs from human capital because it is partially non-appropriable and non-excludable.[10] The stock of knowledge, or technical know-how, is accumulated over time from all the various discoveries of new and better ways to do things or make things. Of course, such an abstraction is not immediately quantifiable. However, a commonly used proxy for this stock of knowledge is the stock of research and development (R&D) expenditure over time.[11] Figure 5.3 in Chapter 5 shows manufacturing R&D stocks for 14 OECD countries between 1973 and 1992. For all countries, this R&D stock is clearly rising over the period. This is one way in which technological progress can be measured. It is also interesting to note that, even for these advanced economies, there are significant differences in accumulated R&D across countries.

Behind the accumulation of technical knowledge, net of knowledge depreciation, is a steady increase in expenditure on R&D through time. This is discussed in more detail in Chapter 4, and Figure 4.1 shows R&D expenditure in manufacturing for 14 OECD countries between 1973 and 1992. In all cases, the trend is rising. An example typical of this sample is Germany, which has seen a twofold increase in real R&D expenditure over this period.

Another commonly used measure of technical progress is total factor productivity (TFP). This is a residual measure left over from decomposing output growth into the component parts driven by the accumulation of the various factors behind growth and is also known as the Solow residual.[12] As such it is really a measure of the efficiency with which inputs (factors) are used to make output. Of course, it is somewhat

[10] This distinction is discussed in more depth in Sections 2.2 and 2.6.
[11] See Chapter 5 for more details.
[12] These vary across studies. A simple model would account for physical and human capital accumulation, say. More sophisticated studies might allow for improving quality of capital inputs. This growth accounting literature is discussed in more detail in Section 2.3.

disingenuous to try to *explain* output growth by the bit of growth that is left over after having accounted for everything else. Nevertheless, TFP growth is often interpreted as a measure of technological progress lying behind output growth. Table 2.5 below shows TFP growth between 1970 and 1990 for a sample of OECD countries.

Table 2.5: TFP Growth 1970-1990

	GDP Growth	Labour	Capital	TFP Growth
US (1972-1990)	2.76	1.45	0.96	0.35
Japan	4.45	0.58	2.11	1.76
France (1972-1990)	2.67	-0.24	0.71	2.20
Italy	3.09	0.37	0.91	1.81
UK (1968-1990)	2.16	0.18	0.68	1.30
Canada (1971-1990)	3.29	1.66	1.04	0.59
Australia (1968-1989)	3.76	1.42	1.11	1.22
Denmark (1972-1990)	1.91	-0.17	0.66	1.43

Notes: Figures are for 1970-1990 unless otherwise stated. Figures are average annual percentage growth rates of GDP, labour contribution to output, capital contribution to output and the residual TFP. Source: OECD (1996).

The clear pattern emerging from Table 2.5 is that factor accumulation cannot account for all the growth in output over the period, but that factors are being used more efficiently over time, or productivity is rising. In other words, we are able to produce better output with the same inputs or the same output with less inputs, and this can only be driven by technological progress (so long as the contribution of the factors has been accounted for correctly). This then is another stylized fact: That output growth cannot be wholly accounted for by factor accumulation. We are continually making new or better products or learning how to do things in new and better ways, and this technological progress appears to explain a significant fraction of output growth.

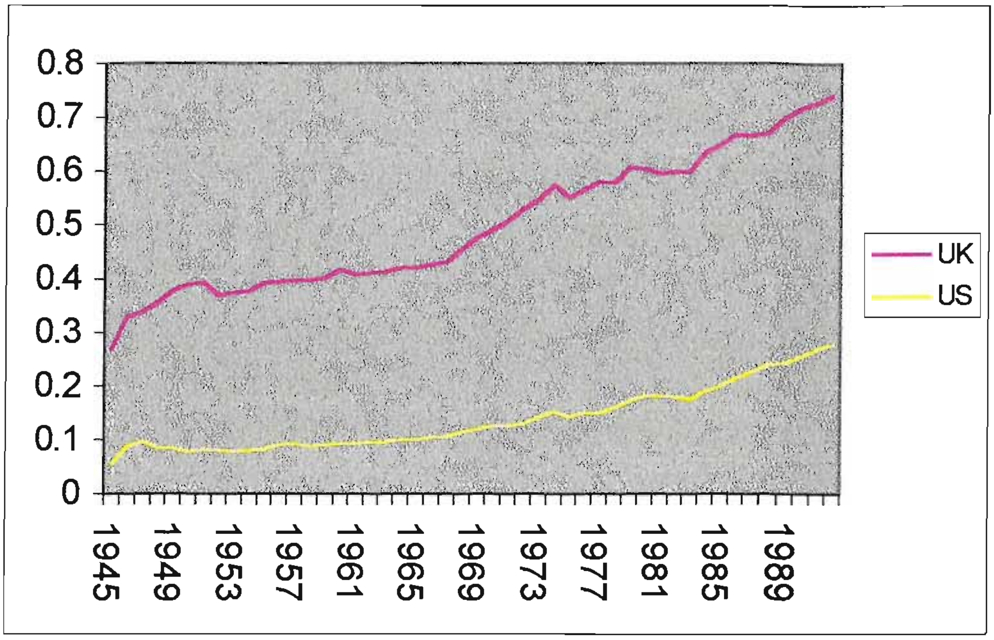
6. Increasing Openness to Trade

It is not surprising that given the growth in output over time trade flows have also increased. However, it is also the case that trade intensity has grown over time. In other words, the proportion of their output that countries trade with one another has increased over time. Related to this is the widespread reduction in barriers to trade (both trade in physical goods and other international transactions such as movements of capital) across the world, falling transport costs and the resulting increase in specialisation according to comparative advantage. Also, economies have become increasingly close in other ways, such as through improved communication networks (Grossman and Helpman, 1991a). This trend has been given extra momentum by the arrival of the internet. All these aspects of increasing internationalisation have huge implications for economic growth and it is this that is the primary focus of this thesis.

Figure 2.2 below shows trade intensity, measured by the ratio of imports plus exports to GDP, for the UK and the US between 1945 and 1992. The rising trend is clear for both countries, and is typical of the majority of countries over the period. Foreign direct investment has also increased dramatically over time, and this is discussed in Chapter 5. Integration of economies through reductions in barriers to trade has been a continual aspect of the world economy over the last 50 years. Examples are the introduction of the European Economic Community and then the Single European Market (SEM), the implementation of the General Agreement on Tariffs and Trade (GATT) and other regional trade blocs such as the North American Free Trade Association (NAFTA) and the Association of South East Asian Nations (ASEAN).

The empirical regularities discussed above form the background to the review of the literature on growth and trade which follows. There is a clear rising trend in output, productivity, capital stocks, stocks of knowledge, technology and trade intensity across countries (at least across developed countries). However, there has been little discussion of how these series might be related. In the sections that follow, I review the theories and evidence that have been presented in the literature to argue that these series are not merely rising independently of one another, but that they are all inter-related, in many cases in a causal way.

Figure 2.2: Trade Intensity in the UK and US, 1945-1992



Notes: Figures are imports plus exports of goods and services over GDP for UK and GNP for US (1945-1983). Figures for 1984-1992 are imports plus exports of manufactures over production of manufactures, scaled by the ratio of the 1983 figure over the 1983 figure for goods and services. Sources: The Economist: Economic Statistics 1900-1983 (1945-1983); OECD STAN Database (1983-1992).

2.2: Growth Theory

2.2.1: The Neoclassical Growth Model

The neoclassical growth model, first presented in the 1950s, represented a huge step forward in the modelling of economic growth. It built on existing work, notably by Ramsey (1928), Harrod (1939) and Domar (1946). Ramsey's contribution was to provide the environment in which to build growth models, with his treatment of household optimisation over time. It was the contributions of Harrod and Domar, however, where growth was driven by the accumulation of capital, that were the immediate predecessors of the neoclassical growth model.

The Harrod-Domar growth model had at its core the prediction that the balanced growth path of an economy was inherently unstable and that any small variation in the magnitude of key parameters, such as the savings ratio, for example, would result in rising unemployment or prolonged inflation. This 'knife-edge' result was criticised by Solow (1956) as being the result of the crucial, but unrealistic assumption of fixed proportions, ie: that capital and labour could not be substituted in production. Solow introduced a standard neoclassical production function to the basic Harod-Domar model.[13] With this basic, but fundamental change, the model gives rise to a stable, balanced growth path.

The standard neoclassical production function combines capital and labour to produce an output good under the following assumptions:

1. Labour and capital are substitutable with a smooth elasticity of substitution,
2. There are constant returns to scale in production, and
3. There are diminishing returns to each input.

More precisely, the first assumption, combined with the third assumption, implies that the marginal product of one input approaches infinity as the amount of the other input approaches zero and approaches zero as the amount of the given input approaches infinity. In other words, if you don't have much capital but have lots of labour, the marginal product of adding an extra unit of capital will be relatively high, so that swapping a unit of capital for a unit of labour will result in a net gain in production. This is true until the marginal products of each are balanced. The assumption of constant returns to scale implies that doubling the combined inputs to production doubles the output. Solow (1994) argues that this is not a critical assumption of the model, but just a simplifying one.

The following Cobb-Douglas production function is a simple example of a production function satisfying the neoclassical conditions:

$$Y = AK^{\alpha}L^{1-\alpha},$$

where A is an exogenously given constant (the technology level) and α is a constant with $0<\alpha<1$, thus displaying diminishing returns to each input but constant returns overall.

[13] I will use the terms neoclassical model and Solow model interchangeably from now on.

The model was developed further, notably by Cass (1965) and by Koopmans (1965), by introducing Ramsey-type consumer optimisation and therefore an endogenous savings rate in contrast to Solow's exogenous savings rate. This development of the model gives richer transitional dynamics as an economy grows towards its steady state level, but does not alter the basic characteristics of the model (Barro & Sala-I-Martin, 1995).

Two such characteristics of the neoclassical growth model are commonly held to be the most important in terms of stimulating subsequent research in the field. They are the prediction of convergence and the failure to explain long run growth within the model (Romer, 1994; Grossman and Helpman, 1994). Both are discussed below. The empirical evidence in support or otherwise of these characteristics of the neoclassical growth model is discussed in Section 2.3.

Firstly, the assumption of diminishing returns leads the Solow model to predict convergence. In other words, economies with less capital per worker will have higher rates of return to capital and therefore higher growth rates than economies with more capital per worker. In the light of later developments, this prediction has become known as *absolute convergence*. The idea is attractive from a policy point of view in that it implies poor countries should automatically catch up with richer countries. There is cross-country evidence for absolute convergence in some cases (see Section 2.1.2 and Section 2.3), and also for cross-region convergence within countries. There is also strong evidence against this prediction of convergence.

The attractiveness of the convergence prediction, coupled with the mixed empirical evidence, makes it both one of the main strengths and weaknesses of the neoclassical model. On the positive side, the existence of a great deal of empirical evidence contradicting the prediction of convergence led to further developments of the basic Solow model in which countries might have different steady state levels of technology, for example, and would therefore not all converge to the same steady state (Romer , 1994). These differences in technology (or a host of other factors that might determine an economy's steady state) lead to the prediction of *conditional* *convergence*, or *β convergence.*[14] Conditional convergence implies that countries with a lower level of capital per worker, relative to their steady state, will grow faster than countries with a higher level of capital per worker, relative to their steady state. This more sophisticated notion of convergence is not so easily rejected by the empirical evidence (see Section 2.3), while remaining consistent with the neoclassical model.

The second crucial implication of the neoclassical growth model is that because of diminishing returns, per capita growth will not go on forever. Growth is driven in the model by capital accumulation, but eventually the return to adding more capital will fall to zero and growth will cease. In this respect, the Solow model is more a model of growth dynamics until a steady state is reached, rather than a model of long run growth. This prediction is contrary to empirical evidence that shows growth persisting over the last 200 years or so (see Section 2.1.2 for a discussion) and is therefore a serious weakness of the model.

Neoclassical growth theorists side stepped this problem by introducing exogenous technological progress.[15] In steady state, it is continuing improvements in technology that drives economic growth. Solow (1957) estimated that technical progress accounted for almost all of economic growth between 1909 and 1949 in the US. This has become known as the Solow residual (see Section 2.3 for a more detailed discussion). Later work by researchers in this growth accounting tradition, such as Denison (1962) and Jorgenson and Griliches (1967), with more sophisticated models, reduced the Solow residual to around one third of economic growth.

Although the neoclassical model can be made more consistent with the empirical evidence by consideration of technical progress in this way, the fact remains that long run growth is not driven from *within* the model, and therefore that the model does not really explain long run growth at all. It was primarily this fundamental weakness of the neoclassical model that eventually inspired the beginnings of the endogenous growth literature of the late 1980s (Romer, 1994). Nonetheless, the neoclassical growth model introduced by Solow remains a landmark of crucial importance in the

[14] In the Cobb-Douglas production function presented above, for example, the 'A'-term might differ across countries in steady state.

growth literature. Indeed, many authors argue that it is still a more realistic description of growth than more recent endogenous growth models (see Section 2.3).

2.2.2: Early Models of Endogenous Growth

The neoclassical model could only explain long run growth by appealing to exogenous technological progress, because of the assumption of diminishing returns to capital. As such, it could not explain the persistence of the huge differences in income levels observed across the world. There are two immediately apparent possible solutions to this problem. First, get rid of diminishing returns. Second, endogenise technological progress in some way. In practice, the two approaches essentially amount to the same thing, as can be seen from the discussion below.

It is straightforward to write down a model without diminishing returns to capital. A simple production function with constant returns would look like this:

(2.1): $y = Ak$,

where y is output per capita, A is constant and k is capital per capita.[16] In this case, capital accumulation would lead to long run growth. This is a simple example of an 'AK' production function and such production functions form the basis of models of endogenous growth that have become known as AK models.

The problem with such models is that constant returns to physical capital are just not realistic, so we cannot just drop diminishing returns. The next step is therefore to re-interpret capital as a broader concept, including other factors as well as physical capital, to which there might be constant returns. For example, broad capital might include human capital and there might be constant returns to the accumulation of this human capital (this leads us to the Uzawa-Lucas model, discussed in the next section). Equally, broad capital might include government provision of services (such as

[15] In the Cobb-Douglas production function presented above, for example, exogenous technical progress would increase the 'A'-term and therefore increase output.

[16] Compare this to the per-capita representation of the neoclassical production function in 2.2.1.

infrastructure) which leads us to the Barro model (see Barro, 1990) with the following production function:

(2.2): $Y_i = AL_i^{1-\alpha}K_i^{\alpha}G^{1-\alpha}$,

where L_i is firm i's labour input, K_i its capital input and G the level of government provision of services. For fixed labour, returns to broad capital (including the public goods provided by government) are constant and therefore endogenous growth is possible.

The learning-by-doing (LBD) model is an alternative method of motivating constant or increasing returns to accumulable inputs. In the LBD model, first suggested by Arrow (1962) and built upon by Sheshinski (1967), technical knowledge is accumulated as an unintended by-product of capital accumulation. In other words, as a firm accumulates physical capital, it also learns how to use this capital more efficiently (Barro & Sala-I-Martin, 1995). Therefore, as above, there can be constant or increasing returns to broad capital (including knowledge) even when the labour force is held constant.

Romer (1986) *replaced* physical capital in the production function with private knowledge capital, which firms could accumulate with diminishing returns. Where this knowledge differs from human capital is in its non-rivalness and non-excludability. In other words, by accumulating knowledge capital, you contribute to the general stock of scientific and technical knowledge that is available for everybody to use. This general or aggregate stock of knowledge is also treated as an input in the production process, as it is assumed to affect the productivity of labour and private knowledge. The production function for this model in its simplest form looks like that in Equation 2.2, with G replaced by K (for aggregate knowledge). However, the constant returns to broad capital are obtained through the actions of profit maximising firms rather than through the imposition of government tax and spend regime. In competitive equilibrium, firms accumulate knowledge in the long run as long as the diminishing returns to the accumulation of this knowledge are outweighed by the increasing returns to factors (labour and broad knowledge) in the production function (Grossman & Helpman, 1991a).

The models of endogenous growth with government provision of services, with human capital and with learning-by-doing, all share the same basic characteristic of constant returns to a broad notion of capital, even though the models are motivated in very different ways. As such, they can all predict long run growth without resorting to exogenous technological progress. However, we have to question the realism of the mechanism of knowledge creation in the LBD model and the assumptions made about the wholly non-rival and non-excludable nature of government-provided goods in the Barro (1990) model.[17] In the latter model, government funded academic research is the best candidate for such a public good (Romer, 1990; Barro & Sala-I-Martin, 1995).

In these models, constant returns are only achieved by resorting to an externality in the production process. Aghion and Howitt (1998) note that this is the only way to remain in a competitive framework and support increasing returns to factors of production overall (including labour).[18] This is because perfect competition cannot support paying all factors of production their marginal product in an increasing-returns environment. In other words, the externality is necessary to keep the private marginal product of broad capital above the discount rate where individual investments would face diminishing returns in the absence of the external boosts to productivity (Grossman & Helpman, 1994). This is a crucial point, so I work through an example below.

The following example is taken from Romer (1990) to describe how increasing returns to production (including a non-rival input) are inconsistent with perfect competition in the absence of external effects.

Suppose a firm faces the production function $Y=F(A,X)$ with X being aggregate rival inputs (eg: labour and physical capital) and A being a non-rival input (eg: technology level). Assume constant returns to rival inputs (as in the neoclassical model and the AK models discussed above), so that $F(A,\lambda X) = \lambda F(A,X)$. If A is productive as well, it

[17] The human capital model is more believable, but discussion of this model is left to Section 2.2.3.
[18] An alternative method is to simply posit a lower bound on the private return to capital to insure investment remains profitable (Grossman & Helpman, 1994). This is adopted by Rebelo (1991).

follows that F cannot be a concave production function because $F(\lambda A, \lambda X) > \lambda F(A,X)$. The sum of the marginal products of the inputs is therefore: $A\{\partial F(A,X)/\partial A\} + X\{\partial F(A,X)/\partial X\} > X\{\partial F(A,X)/\partial X$.

In perfect competition, disk drives would sell at marginal cost and this would exactly equal wage and interest payments to the rival inputs to production (capital and labour), so that $F(A,X) = X\{\partial F(A,X)/\partial X$. Clearly, the marginal product of *all* inputs is greater than the marginal cost of increasing the rival inputs, so that if all inputs were paid the value of their marginal product the firm would suffer losses. Therefore either firms cannot be price-takers or factors cannot all be paid their marginal value product. In other words, perfect competition is only compatible with increasing returns if inputs are not all paid their marginal value product, or are *external* to the firm.

The somewhat stylised nature of these models is not their main weakness, however. Solow (1994) points out that all these models need *exactly* constant returns to broad capital to work. Diminishing returns leads eventually to a stable steady sate with no growth. Even small degrees of increasing returns, on the other hand, lead to explosive growth. Thus, these early models of endogenous growth are criticised by Solow for their 'knife-edge' nature, in a similar manner to his earlier criticism of the Harrod-Domar model.

In terms of fitting with the empirical evidence at a macro level, these models had the advantage of being able to explain long run differences in income levels. Different levels of technology in different countries were possible, leading to different growth rates.[19] Because of the positive link between capital and technology levels in the LBD model and the Barro (1990) model, countries could *diverge* in income rather than converge. An analysis of the consistency of this prediction of divergence with the empirical evidence is presented in Section 2.3. Section 2.4 introduces international trade and discusses how the LBD model, for example, can be made consistent with convergence.

[19] Another important aspect of these models, to which I return in Section 2.4, is that they exhibit scale effects. In other words, the larger the country, the higher the income level.

Despite the fundamental weaknesses of these models in their basic forms, their arrival allowed growth theory to begin to escape the dead end into which it had evolved in the 1970s and 1980s. They also suggested a significant role for trade, which has helped bring the fields of growth and trade together (see Section 2.4). However, introducing more realistic assumptions, such as deliberate R&D leading to ex-post monopoly rents, moves us away from a competitive framework and therefore further from the neoclassical model. It wasn't until this step of allowing imperfect competition was taken that endogenous growth theory really kicked off.

2.2.3: Endogenous Growth with Human Capital

The preceding Section showed various ways of motivating constant returns to a broad notion of capital leading to long run growth in the absence of exogenous technological progress. One such method is to include human capital in the production function, in addition to physical capital. Such a model was presented by Uzawa (1965) and built upon by Lucas (1988) and subsequently by others, notably Rebelo (1991). These models share many of the characteristics of the models discussed in Section 2.2.2: They are all essentially AK models. However, I discuss the human capital models separately to show an alternative method of overcoming diminishing returns without resorting to externalities or technological change, and because they seem plausible in a real-world sense. The discussion will be very brief.

Human capital differs from the technical knowledge of the LBD model because it is not assumed to be non-excludable and non-rival (Lucas, 1988). In other words, an individual is able to accumulate private human capital through education without necessarily adding to social human capital and when he or she uses that human capital in production, it cannot be used elsewhere at the same time. The lack of a non-rival and non-excludable input in the Lucas (1988) model implies that it doesn't need any external effects for long run growth. This provides the first point of departure from the LBD model and the Barro (1990) model.

The basic tenet of Lucas's model is that infinitely lived workers can acquire skills that increase their productivity by forgoing a fraction of potential work-time to invest in education. This process of human capital accumulation is assumed to display constant returns, so that doubling the amount of time spent studying doubles an individual's human capital level. Production of output per worker displays constant returns to broad capital (human and physical). Output growth coupled with constant returns to the accumulation of human capital implies that a smaller proportion of foregone production leads to the same growth in the stock of human capital. Therefore, in steady state, more and more output is invested in broad capital accumulation (in absolute terms) so capital grows faster and faster and this offsets its diminishing returns in production. So the ratio of the return to capital to the cost of capital is constant. This leads to long run growth in steady state. How does this model avoid the problem of paying marginal value products to inputs? It does so, not by assuming external (or spillover) effects across *agents* at a given point in time, but essentially by assuming spillovers across *time* (Aghion & Howitt, 1998). In other words, the previous human capital accumulation of an economy makes current human capital accumulation easier. This inter-temporal external input in the production process is not rewarded. Therefore, the Lucas model of growth with human capital is compatible with perfect competition.

The basic model has been extended, notably by Rebelo (1991), by introducing physical capital into the education sector. Another significant extension is to replace infinitely lived workers with overlapping generations (OLG) of workers, which answers the criticism arising from empirical evidence of diminishing returns to education over a lifetime. Aghion & Howitt (1998) provide a brief but dynamic discussion of extensions and alternatives to the Lucas model in the human capital field, along with micro-foundations and evidence.

The Lucas model shares many of the empirical strengths and weaknesses of the LBD and government services models discussed in the previous section. In particular, it can explain long run differences in income levels and does not predict convergence. It also requires exactly constant returns to scale in the accumulation of human capital otherwise growth will cease or explode. However, even given these similar strengths and weaknesses, the Lucas model does provide a slightly more plausible alternative to the technical progress-driven models of endogenous growth discussed in the next Section than the AK models of Section 2.2.2.

2.2.4: R&D-Driven Endogenous Growth

The weakness of the LBD model (in particular the need for exactly constant returns, shared with the other early endogenous growth models), coupled with the weight of evidence suggesting a link between technological progress and growth meant that an alternative modelling strategy was needed to relate the two. The first stirrings of this strategy can be found in Romer (1987). The benchmark paper that really revolutionised growth theory was Romer (1990), which I discuss below. The key to the revolution was the abandoning of perfect competition.

The common idea behind models of R&D driven growth is that economic agents invest in R&D in order to gain some ex-post monopoly power from innovation.[20] Otherwise, why bother? This brings R&D into the model in a much more explicit sense than the external effects of the LBD model. The idea that incentives to innovate are dependent on ex-post monopoly power has been around for a long time, even in the growth literature. Solow (1956) remarked in a footnote on the desirability of extending the neoclassical model to allow for monopolistic competition (Romer, 1994). Schumpeter had identified monopoly power as necessary for innovation decades before (Schumpeter, 1934). But it wasn't until the late 1970s that the industrial organisation (IO) literature produced the models of monopolistic competition that could be applied to R&D driven growth (Romer, 1994). These models allowed for many firms, all of which could have some market power and thus earn monopoly rents on innovation. The model applied by Romer in his 1990 paper, and still widely used, is the Dixit and Stiglitz (1977) model of product variety. Section 2.5 reviews the IO literature on market structure and innovation in more detail.

Once the model can allow for innovation that earns the innovators rents, the next step is to make this innovation feed back into production and R&D in some way. In other words, innovation improves the technology level of the economy and increases efficiency in production. This feeds back positively into R&D, making it easier

[20] Although until now the terms R&D and innovation have been used largely interchangeably, it is now useful to define terms. Innovation is the realisation of technical capability in the introduction of new or better products or processes. R&D is the necessary process of knowledge creation that enables innovation.

(cheaper relative to production), which drives further innovation and so on.[21] In this way, the profit-maximising actions of firms drive long run technological progress and long run growth. This *inter-temporal spillover* mechanism for long run growth is very similar to that in the Lucas model.

The Romer (1990) Model

Because of its seminal nature, I present the main equations from the Romer (1990) model below. It is presented more as a quick and easy reference than as a critical discussion. Such discussions can be found in Romer's paper as well as elsewhere (eg: Barro & Sala-I-Martin, 1995; Aghion & Howitt, 1998). Some details are skipped over in the following discussion, in the interests of brevity.

Inputs to production are labour (assumed constant from now on), capital, human capital (also assumed fixed in aggregate) and the index of technology. Human capital is distinct from labour, and distinct from the level of technology, which is assumed to be non-rival whereas human capital is assumed to be rival. Each new unit of knowledge (technology) corresponds to a design for a new good. There are 3 sectors in the economy. The research sector uses human capital and the existing stock of knowledge to produce new knowledge (designs for new goods). The intermediate goods sector uses the designs of the research sector together with foregone output to produce producer durables that are available for use in final production. The final goods sector uses labour human capital and the set of producer durables (capital) that are available to produce final output. Final output is given by equation (2.3), where H_Y is the fraction of human capital devoted to final output, and x_i is the set of intermediate goods that are available for use at a given time. This set can be of discrete varieties, or continuous. Romer begins with the discrete case, although the analysis follows for either case. α and β lie between 0 and 1 as usual.

[21] It may be that the stock of possible opportunities for improving technology is finite. If this was the case, R&D would eventually become more difficult. This would correspond to eventually decreasing returns in the equation governing the evolution of the technology level. Such decreasing returns would not be consistent with long run growth. Romer (1990) argues that there is no evidence for this argument. The stylized facts outlined in Section 2.1.2 would seem to support his view.

$$(2.3):\ Y = H_Y^{\alpha} L^{\beta} \sum_{i=1}^{\infty} x_i^{1-\alpha-\beta}.$$

The difference between this production function and standard production functions is that in the standard case, capital goods are all perfect substitutes treated together as an aggregate capital good. In this additively separable case, adding more of one type of capital good does not effect the marginal product of another type. The final output sector is competitive (output is homogeneous of degree 1, so this assumption is not problematical) but the intermediates sector is monopolistically competitive, with a distinct firm i for each type of intermediate good. Once one of these firms has produced or purchased a design, it is just modelled as a black box that turns η units of output into one durable unit of good i, with an infinitely lived patent on the design so that it is the only seller of that good. Assuming away depreciation, the value of the design is therefore the (maximised) present discounted value (PDV) of the flow of monopoly rental income that good i generates.

Total capital (the sum of all the durables) accumulates according to the following rule:

$$(2.4):\ \partial K(t)/\partial t = Y(t) - C(t),$$

or capital accumulated is just foregone consumption. Given that η units of consumption are foregone to create one unit of any type of durable, K is related to the durable goods used in production in the following way:

$$(2.5):\ K = \eta \sum_{i=1}^{\infty} x_i = \eta \sum_{i=1}^{A} x_i.$$

What of the growth of new designs (the growth of A(t))? Romer now moves from discrete to continuous intermediates (this just simplifies the analysis). The sum in equation (2.3) is replaced by an integral. The output of new designs produced by researcher j can be written as a continuous deterministic function of the inputs applied. If δ is a productivity parameter, H^j and A^j are the researcher's amounts of human capital and knowledge (s)he has access to, new designs will be discovered at the rate

$\delta H^j A^j$. Assuming all knowledge is freely available to all, the aggregate stock of designs evolves according to (2.6).

(2.6): $\partial A/\partial t = \delta H_A A$,

where H_A is the amount of human capital employed in research. As technology is accumulated, production of new designs gets easier. This then is the key assumption of the inter-temporal spillover that drives long run growth.

Let spot prices at any point be measured in units of current output and let r denote the interest rate on loans denominated in goods. P_A denotes the price of new designs and w_H denotes the rental rate per unit of human capital. It follows from equation (2.6) that $w_H = P_A \delta A$. Given prices p(i) for all durables, a profit maximising firm chooses quantity x(i) for each durable. The inverse demand for the durables is given by (2.7).

(2.7): $p(i) = (1-\alpha-\beta) H_Y^{\alpha} L^{\beta} x(i)^{-\alpha-\beta}$.

Each producer takes (2.7) as given in setting the price of durables. A firm that has already incurred the fixed cost of creating a new design will choose the level of output x to solve the following problem:

(2.8): $\max_x \; p(x)x - r\eta x = \max_x \; (1-\alpha-\beta) H_Y^{\alpha} L^{\beta} x^{1-\alpha-\beta} - r\eta x$.

This is the monopoly pricing problem of a firm with constant marginal cost ($r\eta x$) that faces a constant elasticity demand curve (2.7). The resulting monopoly price is a simple mark-up ($\bar{p}$) over marginal costs, where $\bar{p} = r\eta/(1-\alpha-\beta)$. The flow of monopoly profit is $\pi = (\alpha+\beta)\bar{p}\,\bar{x}$, where $\bar{x}$ is the quantity corresponding to $\bar{p}$ from (2.7). Given a competitive market for new designs, the price for new designs will be equal to the present value of the net revenue that a monopolist can extract from such a design. Therefore, at every date t,

(2.9): $\int_t^{\infty} e^{-\int_t^{\tau} r(s)ds}\, \pi(\tau)d\tau = P_A(t)$.

In equilibrium, P_A will be constant, and differentiating with respect to t gives us:

(2.10): $\pi(t) = r(t)P_A$

In other words, at every point in time, the excess of revenue over marginal costs is just enough to cover the interest on the cost of the initial investment in a design.

The model is closed by introducing Ramsey consumers with discounted constant elasticity preferences:

(2.11): $\int_0^\infty U(C)e^{-\rho t}dt$, where $U(C) = (C^{1-\sigma}-1)/(1-\sigma)$ for $\sigma \in [0, \infty)$.

The intertemporal optimization condition for a consumer facing a fixed interest rate r is that the growth rate of consumption is equal to $(r-\rho)/\sigma$. These consumers are endowed with fixed quantities of labour L and human capital H that are supplied inelastically. At time 0, consumers own the existing durable-goods producing firms and receive dividends from their net revenues.

The solution of the model for a balanced growth path drops out from the above equations and looks as follows. From (2.7) and (2.10) the price of a new design can be rewritten as

(2.12): $P_A = \pi/r = (\alpha+\beta)\bar{p}\,\bar{x}/r = (\alpha+\beta)(1-\alpha-\beta)\, H_Y^{\alpha}L^{\beta}x^{1-\alpha-\beta}/r$.

Wages paid to human capital in the research sector must be equal to income from the research sector, so the research wage is $P_A\delta A$. Human capital allocation to each sector must equalize the returns to human capital so that

(2.13): $w_H = P_A\delta A. = \alpha H_Y^{\alpha-1}L^{\beta} \int_0^\infty \bar{x}^{1-\alpha-\beta}di = \alpha H_Y^{\alpha-1}L^{\beta}A\,\bar{x}^{1-\alpha-\beta}$.

From (2.12) and (2.13) the allocation of human capital to production of final output is

(2.14): $H_Y = \alpha r/\delta(1-\alpha-\beta)(\alpha+\beta)$.

With fixed $H_A = H-H_Y$, the exponential growth rate of A is (from (2.6)) δH_A. With all else fixed, examination of the expression for final output (2.3) shows that output grows at the same rate as A. With $\bar{x}$ (the optimum quantity of each intermediate) fixed, K also grows at the same rate as A (from (2.5)). Given (2.11), and the steady state condition $g = (r-\rho)/\sigma$, the common growth rate for all these variables is therefore:

(2.15): $g = \delta H_A. = \delta H - \alpha r/(1-\alpha-\beta)(\alpha+\beta) = (\delta H - \Lambda\rho)/(\sigma\Lambda + 1)$,

where $\Lambda = \alpha/(1-\alpha-\beta)(\alpha+\beta)$.

So the Romer (1990) model can give us balanced long run growth driven by profit maximizing entrepreneurs investing in deliberate R&D in return for monopoly profits. As such, it represents a substantial leap forward from the more abstract LBD and public good models of growth. It also provides the basic modelling framework for the models of R&D-based growth and international trade due to Grossman and Helpman (1991a) and Rivera-Batiz and Romer (1991a, 1991b). Like earlier endogenous growth models, there is a constant growth rate which is sub-optimal.[22] More importantly in the trade context, the model also shares the lack of convergence and scale effects properties of the earlier endogenous growth models. Equation (2.15) shows how a larger country with a bigger endowment of human capital H will grow faster than a smaller country with smaller H. In short, the growth rate of a country is proportional to its size. The empirical evidence for such scale effects is at best thin, as discussed in Section 2.3.

A further major criticism of the earlier endogenous growth models is that they need exactly constant returns to broad capital accumulation to sustain finite growth. Romer (1994) points out that his 1990 model suffers from a similar weakness, where the equation governing the evolution of technology needs constant returns (see equation

[22] The model has room for government to subsidise research to achieve the optimal growth rate. Barro and Sala-I-Martin (1995), Chapter 6, provides a good discussion of this.

(2.6)). He argues, however, that this is just a convenient simplification and that decreasing returns can sustain growth for a long time. He also argues that increasing returns can be made consistent with non-explosive growth, by modifying equations elsewhere in the model.

The Romer model presented above assumes deterministic innovation. By introducing uncertainty into R&D, the model can lead to cyclical fluctuations, and a considerable literature on growth and cycles has recently appeared. Aghion and Howitt (1998) provide a detailed discussion of this literature. Indeed, the Aghion and Howitt model of growth through creative destruction (discussed below) incorporates uncertainty, which can lead to fluctuations.

Quality Enhancing R&D

The Romer model discussed above assumes R&D results in the introduction of new varieties of products, horizontally differentiated from existing varieties. This is not the only kind of innovation that has been considered in the literature. Aghion and Howitt (1992) present a model in which innovations are improved versions of existing products. In other words, R&D leads to the introduction of new products that are vertically differentiated, or differentiated by quality, with existing products. This model is discussed briefly below. A further alternative is that R&D could lead to the discovery of new and cheaper ways of producing existing products. This kind of innovation is known as process innovation, and is discussed in Chapter 3.

Aghion & Howitt's 1992 model of growth can also give rise to a constant growth rate generated by deliberate profit-driven R&D. The equilibrium of the model is determined by a forward-looking difference equation, with current research expenditure dependent on the expected amount of research expenditure in the next period. The source of this dynamic relationship is the loss of monopoly power suffered by a firm should its product be improved by another firm. This process is described as *creative destruction.* Despite the differences in the formulation of innovations between this model and the Romer model, many predictions, including scale effects and non-convergence, are common to both models.

2.3: Evidence and Growth Theory

The models discussed in the previous section, no matter how elegant in theory, must in the end be tested against empirical evidence to establish their applicability. The discussion of these models has so far been short on specific details of the empirical literature, although not entirely divorced from such evidence. This section does not repeat the discussion of general empirical regularities found in Section 2.1.2. Rather, it outlines specific empirical contributions to the growth literature on the questions of the role of technological progress in growth, convergence and scale effects.

A number of recent reviews of the empirical growth literature exist that provide valuable sources of information and discussion (eg: Pack, 1994; Barro & Sala-I-Martin, 1995; Aghion & Howitt, 1998; Temple, 1999). However, these reviews are not all entirely neutral. Barro and Sala-I-Martin (1995) tend to favour evidence that supports the neoclassical model, and contains little detailed criticism of cross-country growth regressions. Aghion and Howitt (1998) present the evidence in such a way so as to support their support for suitably augmented endogenous growth models. Temple (1999) provides a clear and concise critique of techniques and perhaps the most neutral review of the empirical literature. My brief critique of these empirical methods is largely based on Temple's points. His general conclusion is that much of the empirical evidence can be made consistent with both neoclassical and endogenous growth models, a view generally shared by Pack (1994). This variety of interpretations of the overall empirical literature confirms the need for a topic-by-topic approach.

Before going on to discuss specific predictions and properties of the various models of growth outlined in the previous section, it is helpful to clarify the main alternative sources of empirical evidence. These are historical analysis, microeconomic analysis, growth accounting and cross-country growth regressions. Solow (1994) argues that given the weaknesses of the cross-country growth regression approach, historical analysis is of crucial importance in assessing growth models. Temple (1999) argues that although historical analysis is useful in terms of pointing to potentially important factors behind growth, in general, quantification of these factors requires some kind of cross-country regression, despite their weaknesses. A wealth of micro level evidence exists that can inform the growth literature, but I stick to macro level evidence here. The exception is a brief consideration of the micro evidence that attempts to evaluate the returns to R&D, which is useful to assess the role of technological progress in growth. The primary sources of macro level empirical evidence on growth are growth accounting and cross-country growth regressions and it is these literatures that I discuss in most detail in this section.

Growth accounting traditionally involves the study of a given country through time.[23] Its purpose is to decompose output growth into its component parts due to factor accumulation and a residual (TFP growth) which represents efficiency growth (see the discussion of TFP in Section 2.1.2). As such, it is a useful technique for assessing the role of technological progress in overall growth, and I refer to many such studies in the following section. These studies start from the simple analysis of Solow (1957) through to more complex recent studies including consideration of quality of inputs and also sometimes knowledge stock inputs (eg: Jorgenson & Griliches, 1967; Griliches, 1988). However, Aghion and Howitt (1998) argues that growth accounting cannot be interpreted as providing firm evidence of the causes of growth because of the endogeneity of capital accumulation. In other words, technological progress drives capital accumulation as well as output growth, so that the overall contribution of technological progress will always be underestimated (Grossman & Helpman, 1994). Temple (1999) argues that growth accounting should be used rather to assess how various shocks act (eg: through factor accumulation or through TFP) and not to assess their ultimate effect. In addition, the interpretation of the residual as technological progress is dangerous if all other possible contributions to output growth are not correctly specified.

Cross-country growth regressions provide the meat of the macro empirical growth literature, and have been possible since the arrival of cross-country data sets such as the Maddison and the Summers-Heston data sets (see Romer, 1994; Barro & Sala-I-Martin, 1995). In their simplest form growth rates for the sampled countries are regressed against initial income, the investment rate and a number of other variables to

[23] A recent development is cross-country growth accounting where suitable capital stock data is available (see Benhabib & Spieger, 1994, for example). This extension allows researchers to explicitly capture cross-country variation in TFP growth (Temple, 1999).

capture human capital, trade barriers and so on. In principle, such regressions allow us to explain the causes of growth and differences in growth outcomes across countries. They can also be used to test model predictions such as convergence (see Section 2.3.2). Panel cross-country regressions have the added advantage of being able to control for fixed unobserved effects, such as initial efficiency level (Temple, 1999).

Temple (1999) summarises the various criticisms of cross-country growth regressions in a clear and concise way. The main problem is that very different countries are unlikely to be drawn from the same surface. In other words, given that countries are very different we would expect parameter heterogeneity across countries instead of the common parameters imposed by the multiple regression approach (see Durlauf & Johnson, 1995; Durlauf & Quah, 1998, for a discussion). For example, the parameter on the investment ratio will be very different in an unstable war-torn country than in a stable, peaceful country (Temple, 1999). Of course, if the heterogeneity is entirely random, then cross-sectional regressions will still estimate unbiased average parameters. The presence of outliers in the sample can also be problematical. Rousseeuw and Leroy (1987) suggest a robust estimation method that should stand up to the presence of such outliers. However, the majority of the empirical literature is still open to both criticisms.

Another fundamental weakness of growth regressions is their extreme sensitivity. Levine and Renelt (1992) find almost all results from growth regressions to be fragile to model specification changes, with the possible exceptions of the parameter on the investment ratio and the correlation of the ratio of trade to GDP with the share of investment in GDP. On top of this, there are problems of measurement error, error correlation (eg: due to regional spillovers) and the endogeneity of many right-hand-side variables (Pack, 1994; Temple, 1999).[24] Pack (1994) also notes the importance of changes in sectoral composition of output; largely ignored by cross-country growth regressions. In short, the results of growth regressions should be viewed with caution. Nevertheless, Temple (1999) argues that they represent the best method for

[24] The endogeneity problem is best addressed in a panel framework, where lagged values can act as instruments. Panels also have the advantage of controlling for omitted country-specific variables that will bias results in cross-sectional regressions.

quantifying the causes of growth and the observed differences in growth rates across countries.

2.3.1: The Role of Technological Progress in Growth

The neoclassical growth model explains long run growth by appealing to exogenous technological progress. The R&D-based endogenous growth models of Romer (1990), Aghion and Howitt (1992) and others also rely on technological progress as the driving factor behind long run output growth, albeit generated within the model. If we are to accept these explanations of growth as valid, the data should show the importance of technological progress in driving growth.[25] What is the evidence for this? Section 2.1.2 shows R&D, knowledge levels and output all growing over time, but is there any link between these trends?

Evidence for the role of technological progress in output growth comes from a variety of different sources. Growth accounting allows us to decompose output growth into that part accounted for by factor accumulation and a residual that is interpreted as efficiency growth (TFP growth). The first paper to attempt this decomposition was Solow (1957). Solow's basic finding was that, after controlling for capital accumulation and labour force growth, technological progress accounted for almost all of economic growth in the US between 1909 and 1949. Later work in this growth accounting tradition, such as Jorgenson and Griliches (1967) adjusted for changes in the quality of inputs and reduced this residual (the Solow residual) to around one third of output growth (Cameron, 1996). Recently, the Jorgenson and Griliches (1967) approach has been applied by Dougherty (1991), Elias (1990) and Young (1995a), for a variety of different countries. Some growth accounting papers also include measures of R&D stocks to try to directly capture the effect of technical knowledge inputs (eg: Denison, 1985; Maddison, 1987; Griliches, 1988). These papers tend to have difficulty in explaining the productivity slowdown in the 1970s, which Pack (1994) interprets as a lack of support for R&D-based endogenous growth theory.[26] Barro and Sala-I-

[25] This is not an analysis aimed at comparing these models. They are both consistent with a significant role for technological progress.

[26] Such R&D data is generally accepted to be poor, which is another possible explanation for the lack of measurable impact on 1970s productivity.

Martin (1995) also criticise this approach on the grounds of the potential for reverse causality.

If we take the results of, say, Jorgenson and Griliches (1967) at face value, then we can put down a large fraction of output growth to technological progress. This provides support for both the neoclassical model and the R&D-based endogenous growth models. However, the large fraction of output growth explained by capital accumulation is not necessarily compatible with simple models of endogenous growth (see, for example, Young (1995b); Jorgenson, 1995). Endogenous growth enthusiasts argue that models can be extended to incorporate a substantial role for capital accumulation (see, for example, Aghion & Howitt, 1998). Temple (1998) argues that the unexpected importance of capital might reflect appropriate technology factors (see Section 2.3.2).

Growth accounting is vulnerable to the criticism that technological progress might act in large part through capital accumulation, which would result in the underestimation of its role in output growth (Grossman & Helpman, 1994; Barro & Sala-I-Martin, 1995; Temple, 1999). The conclusion suggested by the growth accounting literature is therefore that technological progress plays a substantial role in output growth alongside factor accumulation.[27]

Jones (1995) argues that the weight put on R&D in recent endogenous growth models is inconsistent with the time-series evidence from industrialised countries such as the US. I return to Jones's critique in Section 2.3.3 in the context of scale effects. However, in terms of the role of technological progress in growth, he argues that it cannot be all-important since we have observed fairly constant growth rates despite a five-fold increase in the inputs to research in the US. Temple (1999) argues that this could reflect a fall in research productivity whilst remaining consistent with a key role for technology in output growth.

A wealth of micro evidence suggests that R&D plays a major role in driving output growth. For example, private rates of return to R&D in the region of 30-50% have

[27] Another general finding of the growth accounting literature is the importance of human capital accumulation, which lends support to the Uzawa-Lucas model discussed in Section 2.2.3.

been estimated for the US in the 50s and 60s, with social returns (via spillovers) being even higher (Griliches, 1979, 1992).[28] Grossman and Helpman (1994) argue that although business R&D accounts for only around 2% of industry GDP in the OECD, these social returns coupled with the lack of depreciation of knowledge capital imply a huge effect on output growth. Coe and Helpman (1995) find even higher private rates of return to R&D and social rates of return that are higher still, in their study of international knowledge spillovers. Further discussion of this literature is left to Section 2.6 and Chapter 5.

Finally, Grossman and Helpman (1994) argue historical analysis points to the all-important role of technological progress in driving output growth. They ask what would the last century's growth performance have been without the invention and refinement of methods for generating electricity and numerous other key developments.

2.3.2: Convergence?

Convergence is briefly discussed in Section 2.1.2. The conclusion from a quick look at the cross-country data is that convergence is not observed in the sense that poor countries grow faster than rich ones. In fact, Barro and Sala-I-Martin (1995) find a weak positive correlation between initial income and growth rates, suggesting the opposite. Given that convergence is a property of the neoclassical model and not of simple endogenous growth models, its presence or otherwise provides an interesting test between these models.[29] However, this lack of evidence of *absolute* convergence is only part of the story. If we keep all other factors equal (eg: investment rates, human capital etc.) we do observe convergence. This is known as *conditional* convergence and its presence supports the predictions of the augmented neoclassical growth model of Cass (1965) and Koopmans (1965).

[28] In other words, a 1% increase in R&D expenditure leads to a .3-.5% increase in output growth. Doubling R&D could increase output growth by as much as half.

[29] Introducing trade and technology transfers into endogenous growth models can make them fully consistent with convergence, however. Aghion and Howitt (1998) therefore argue that convergence cannot be used to test between neoclassical and endogenous growth models.

Evidence for convergence comes largely from cross-country growth regressions. After controlling for differences in the investment ratio and population growth (and other possible factors such as human capital, for example), a negative coefficient on the initial level of income implies conditional convergence. This is usually interpreted as supporting diminishing returns to capital and therefore the neoclassical model. Examples of growth regressions that find such a negative coefficient on initial income are Baumol (1986), Barro (1991) and Mankiw et al (1992), although there are numerous others.[30]

However, such regressions are widely criticised, as discussed above. For example, a major criticism of growth regressions is their fragility to changes in model specification and sample. Barro (1991) and De Long (1988) show how the particular sample of countries chosen determines whether convergence is found or not.[31] However, despite this sensitivity to sample and specification, Levine and Renelt (1992) conclude that overall there is at least some support for conditional convergence.

Lucas (1988) and Romer (1986, 1989) argue that evidence suggesting absolute divergence (eg: the positive correlation between initial income and growth rates) is difficult to reconcile with the neoclassical model's prediction of convergence. However, Mankiw et al (1992) show how such absolute divergence is perfectly compatible with conditional convergence in an augmented neoclassical model including human capital. Given this evidence, Romer concludes that the evidence of income divergence cannot be used to reject the neoclassical model (Romer, 1994).

Temple (1998) points out that absolute divergence implies that incomes must have been closer in the past and that something must have acted to drive income levels apart. Based on previous work from Atkinson and Stiglitz (1969) and others, he suggests that 'appropriate technology' might provide an explanation. In other words, he argues that efficiency growth rates are positively correlated with capital/labour ratios and that poorer countries need to accumulate capital before they can take

[30] Bernard & Jones (1996) argue that it is the service sector that has driven convergence for 14 OECD countries between 1970 and 1987 rather than the manufacturing sector.
[31] It is generally accepted that including sub-Saharan Africa tends to obscure convergence.

advantage of technological progress taking place in richer, capital-intensive countries.[32] He presents some evidence in support of this explanation.

Temple then highlights a widespread inconsistency of standard explanations of convergence (including Mankiw et al's) when set against evidence of absolute income divergence. If poorer countries are growing slower than richer countries, he argues, they must be converging to their steady states from *above* and not below. This is counter-intuitive so he suggests that to reconcile this fact with the neoclassical model, rates of efficiency growth must vary across countries. This step is also suggested by Pack (1994). Grossman and Helpman (1994) also argue that the assumption of a common rate of technological progress is a fundamental weakness in the Mankiw et al specification. The assumption, they argue, doesn't fit the data and is counter-intuitive, and will lead to possible biases in regression results, if, for example, a country's growth rate of technological progress is correlated with its investment ratio.[33]

As mentioned in Section 2.1.2, Quah (1997) suggests there is a pattern of convergence 'clubs', with countries becoming stratified into (at least) two distinct income classes over time. In other words, already rich countries display convergence to one income level and already poor countries display convergence to another lower level. Some countries in the middle display convergence to the high-income group and others display convergence to the low-income group. The evidence for Quah's assertion is quite strong. The existence of such convergence clubs suggests a possible compromise between the assumptions of common or varying technology growth rates as represented by the Mankiw et al (1992) paper and the Temple (1998) paper.

What does all this tell us in the end? The weight of evidence on convergence seems to favour neoclassical models over simple endogenous growth models. However, Aghion and Howitt (1998) and Temple (1999) point out that just as income divergence cannot be used to reject the neoclassical model of growth, evidence of convergence cannot be used to reject endogenous growth models. In particular, the presence of international trade and technological spillovers can reconcile such models with the prediction of

[32] This is an idea that is also found in places in the spillovers literature, and is discussed at more length in Chatper 5 of this thesis.

convergence (see Section 2.6). One possible conclusion is therefore that the empirical convergence literature offers support for both neoclassical and endogenous approaches to modelling growth.

2.3.3: Scale Effects?

The Romer (1990) and Aghion and Howitt (1992) models of R&D-based endogenous growth share the prediction of scale effects with earlier models of endogenous growth. This property implies that growth rates are proportional to the size of the economy. For example, equation (2.6) in the earlier description of the Romer (1990) model (see Section 2.2.4) contains a scale effect from the size of the labour force. In other words, an economy twice the size of another should have twice the growth rate, other things being equal. The scale effects property of these models is very interesting in the context of international trade and economic integration, as discussed in Section 2.4. However, general evidence is not supportive of this prediction. We don't observe the US growing much faster than Luxembourg, for example. The neoclassical model has no such property.

Jones (1995) argues that the lack of evidence of scale effects is a fundamental blow to standard models of endogenous growth. He uses the example of the US where a five-fold increase in the number of scientists and engineers working in the research sector has gone hand in hand with a fairly constant growth rate.[34] The Romer (1990) model would predict a five-fold increase in growth rates. Backus et al (1992) provides the most systematic search for scale effects at macro (country) level to date. They allow for scale effects consistent with the LBD model (from GDP or from total output of manufacturing), from human capital inputs (total student and teacher numbers) and from R&D inputs (number of scientists and engineers in research sector and R&D expenditures). They find some (weak) evidence of a positive relationship between total GDP and total manufacturing output and growth. However, there is no evidence of

[33] Temple (1999) argues that panel data studies might go some way towards addressing this problem by incorporating initial differences in technology levels as fixed effects.
[34] Jones presents a modified version of the Romer model that is consistent with this lack of scale effect evidence. However, although growth is still endogenously generated by R&D in this model, the long run growth rate is determined only by exogenous parameters, such as the rate of population growth. He calls this semi-endogenous growth. Recently, other models of endogenous growth without scale effects have been suggested (see Eicher & Turnovsky, 1999, for a review).

scale effects from human capital or R&D inputs, which supports Jones's critique of R&D-based endogenous growth models.

However, in defence of these models, Backus et al (1992) note that the quality of R&D data is poor, certainly in terms of international comparability. They also argue that it is the output of the research process (ie: successful innovation) rather than inputs that we should measure. Barro and Sala-I-Martin (1995) argue that in the context of increasing international trade, where ideas may flow freely across borders, countries may not be the relevant economic unit in which to think of these scale effects. Rather, we should think in terms of world inputs. In a cross-country growth regression setting, Barro and Sala-I-Martin (1995) find some evidence of a weak positive scale effect (of population on growth), but this is not significant at standard levels. In conclusion, predictions of scale effects can only be viewed with extreme caution, and the endogenous growth models displaying this property must be questioned.

2.4: International Trade and Growth

So far, the discussion of the growth literature has largely ignored the presence of international trade. In Section 2.1.2, I presented general evidence of the growing volume and intensity of world international trade. Costs of trade, such as transport costs, are falling. There has been a significant reduction in barriers to trade, most notably in the setting up of new regional trade blocs alongside closer integration within existing trade blocs, but also world-wide with the implementation of the General Agreement on Tariffs and Trade (GATT). A great deal of attention has been paid in the literature to the implications of observed international trade for growth.

Section 2.4.1 considers the neoclassical growth model in an open economy framework. Section 2.4.2 then considers R&D-based endogenous growth models in an open economy framework. The implication of open-economy R&D-based endogenous growth models is that trade and growth are not merely important but unrelated economic phenomena, but that they are linked in a causal sense. Section 2.4.3 reviews the extensive empirical literature on the effects of trade on growth. A substantial literature exists discussing the effects of international trade, not only in goods, but also in ideas. Discussion of this spillovers literature is left until Section 2.6. Here, I concentrate on physical trade in goods.[35]

Trade is only considered here in the context of growth models, so I do not provide any detailed discussion of the trade literature itself.[36] However, for background purposes, it is useful to briefly mention the Factor Proportions Theory, the problems reconciling this theory with observed trade patterns and the subsequent emergence of models of trade based on imperfect competition and specialisation. The development of new models of trade based on imperfect competition also played a significant part in the development of the endogenous growth literature. Consider Romer's use of the Dixit-Stiglitz model of monopolistic competition with differentiated products, for example.

Traditionally, trade has been explained by the concept of comparative advantage. For example, assume in a two-country world that country A is land-rich and country B is capital-rich. Country A produces agricultural products more efficiently and country B produces manufactures more efficiently. In the absence of international capital flows and trade in goods, both countries must produce both agricultural and manufactured products. However, the relative inefficiency of A and B in manufacturing and agriculture respectively implies that welfare in both countries would be improved if each specialised in its efficient sector and then (costlessly) traded with one another.[37] The comparative advantages of the two countries therefore drive trade. This is a simple example of the Heckscher-Ohlin model of factor proportions.

Despite its intuitive appeal, the traditional trade theory fails to account for much of what we see empirically. In particular, differences in factor endowments cannot explain the volume and composition of trade between similarly endowed countries and especially the substantial observed amount of intra-industry trade (Helpman and Krugman, 1985). These empirical shortfalls have led to the consideration of

[35] Including capital goods.
[36] See, for example, Krugman (1995) for such a discussion.
[37] Section 2.1.1 presents a brief passage from Smith (1776) in which specialisation is described as one of only three basic ways of improving a country's standard of living. The others are technical progress and factor accumulation. Specialisation as described in the above trade example however, would not have long run growth effects; only a one-off level effect on output. This separation of level effects from long run growth effects is a recurrent theme in the economic integration literature.

economies of scale and imperfect competition in models of trade, amongst other things. The development of this new trade literature mirrors the development of the growth literature in many respects. In particular, relaxing the related assumptions of constant returns to scale and perfect competition was the foundation for much of the new trade theory. Scale economies and imperfect competition can explain intra-industry trade, for example, where factor proportions cannot.

2.4.1: Trade and the Neoclassical Model

In the long run, the neoclassical model assumes growth is driven by *exogenous* technological progress. Therefore, long run growth is unaffected by opening up the model to international trade as it is still exogenously determined. However, trade does have interesting implications for growth in transition to steady state.

Barro and Sala-I-Martin (1995) present a small open economy version of the neoclassical model and work it through to its counterfactual conclusions. In short, if goods and capital are mobile across borders (international trade and international borrowing and lending), production in any one country can diverge from consumption and investment in that country. Because capital is freely mobile there is a world interest rate. Because the domestic economy is small, it does not effect this interest rate, which we can therefore treat as constant for simplicity. If the domestic economy is anywhere below its steady state level of capital, capital will flow from abroad until the marginal product of capital is equal to the world interest rate. Given that this inflow is not constrained by domestic production and investment in any way, it is instantaneous and the country converges to its steady state infinitely fast. In addition, if countries have different levels of time preference, all but the most patient display consumption tending to zero, as they mortgage all their capital and labour income to the patient country in return for early consumption. Asymptotically, this patient country owns everything.

A number of extensions to the basic open-economy Solow model can improve its transitional predictions. Barro and Sala-I-Martin (1995) discuss a number of such extensions. They present a model with human and physical capital and credit constraints. Such extensions can slow down the rate of convergence to more realistic levels and prevent consumption falling to zero.[38] Convergence can also be slowed down by the introduction of adjustment costs to capital investment. Uzawa (1968) suggests a model where rates of time preference can vary over time. Specifically, countries become more patient as they become poorer. This avoids the zero consumption prediction of the basic open economy model, but is unfortunately somewhat counter-intuitive (Barro and Sala-I-Martin, 1995). Blanchard (1985) obtains similar results but with a much less controversial assumption of heterogeneously aged population with finite horizons. Mountford (1996) shows how gradual convergence can be obtained with perfect capital mobility when (exogenous) land is included as a factor of production.

Aghion and Howitt (1998) note the significant contribution to the open-economy neoclassical growth literature made by Ventura (1997). By incorporating factor price equalisation into the model, diminishing returns set in at the world level as opposed to at individual country level because an economy can shift into more capital-intensive export sectors as it accumulates capital.[39] As long as investment rates don't increase with increases in capital stock, then Ventura's model does not predict convergence (Aghion and Howitt, 1998).[40] Although Ventura's model does not replace the need for exogenous technological progress in the neoclassical model, it does extend the transitional growth period, at least until world diminishing returns set in. It also provides a significant role for trade policy, consistent with much of the empirical evidence reviewed in Section 2.4.3. Increasing openness leads to increased growth in the medium run.

2.4.2: Trade and Endogenous Growth

In contrast to the neoclassical case, the introduction of international trade to endogenous growth models has considerable implications for long run growth. Grossman and Helpman (1991) provide a full discussion of R&D-based models of

---

[38] Instead, all but the most patient country become credit constrained in this model.

[39] In the absence of trade barriers, trade from a labour-intensive to a less labour intensive country, for example, acts (through the demand for labour) to raise the relative price of labour in the labour abundant country and vice versa until factor prices are equalised (Brown and Hogendorn, 1994).

[40] Just as introducing spillovers to R&D-based endogenous growth models can make them consistent with convergence, so introducing factor price equalisation to the neoclassical model can make it inconsistent with convergence.

endogenous growth and international trade. River-Batiz and Romer (1991) extend the Romer (1990) model to consider economic integration between two countries. Aghion and Howitt (1998, Chapter 11) provides a more recent survey of this field. I will concentrate on the effects of extending the Romer (1990) model to the open economy case, following the Rivera-Batiz and Romer (1991) approach.

Introducing trade to the Romer model can have effects at both macro and micro level, as outlined below. However, this section concentrates on the macro level effects. Discussion of the microeconomics literature on R&D and the effects of trade at the micro level is left to Section 2.5.

In Chapter 4, I identify four channels through which trade might affect productivity growth.[41] To these I add a fifth 'allocation effect' as described by Grossman and Helpman (1991a) and River-Batiz and Romer (1991a).[42] They are:

1. Trade might affect growth through its effect on the diffusion of ideas across countries.
2. Trade might affect growth by reducing the incentives for duplication in research across countries and by affecting the incentives to co-operate in R&D across national boundaries.
3. Trade affects the intensity of competition facing firms in a given industry, which may affect R&D incentives.
4. Trade increases the size of the market available to successful innovators, which increases incentives to invest in R&D and therefore stimulates growth, ceteris paribus.
5. Trade may increase demand for certain manufactured goods and act to shift labour from the research sector into the production sector for these goods, slowing down long run growth.

Discussion of the second and third points is left to Section 2.5. The present discussion focuses on the first point and on the last two points.[43] The first and fourth points both

[41] See Chapter 4, Section 4.2 for a more detailed discussion focussing on R&D incentives.
[42] Discussion of dynamic comparative advantage is left to Chapter 3 and is briefly touched upon in Section 2.6.
[43] Discussion of the empirical spillovers literature is left largely until Chapter 5, although Section 2.6 considers the implications of spillovers for the growth literature as a whole.

describe effects of trade and growth that work through the mechanism of *scale effects*. Scale effects basically imply that doubling the size of an economy doubles its growth rate, and they are a critical characteristic of the Romer (1990) model, as discussed in Section 2.2.4. International trade stimulates growth by acting to increase the scale of an economy by blurring national boundaries.[44]

Consider the first point. Technological progress in Romer's model is linked to the technology level of the economy. In other words, the higher the level of technological know-how in an economy, the more effective is its R&D. Opening an economy up to international trade allows access to a greater stock of technical knowledge, making R&D easier.[45] In other words, for a given level of R&D, a country can increase its growth rate by tapping the international stock of scientific and technical knowledge. Therefore, the research sector expands and the growth rate increases. Phillips (1966) described this effect as the *technology-push* effect.

The larger the market that an innovator finds herself or himself in, the greater the potential profit from a given innovation. Other things being equal, doubling the size of the market doubles the potential reward from innovation. Incentives for R&D are therefore increased, stimulating growth in the research sector and a higher rate of long run growth. Opening up a Romer economy to international trade has this market enlargement effect and so stimulates growth. Schmookler (1966) called this the *demand-pull* effect.

Rivera-Batiz and Romer (1991) and Grossman and Helpman (1991) show how trade can lead to a negative allocation effect when inputs (eg: human capital) are shared between the manufacturing and research sectors. Increasing openness (or moving from autarky to free trade) can stimulate demand for exported goods which will raise the price of human capital[46] in the export goods sector and therefore cause a re-allocation of human capital from the research sector to the export goods sector. This will slow down the rate of technology growth. The effect would work in the opposite direction for countries that were net importers of human-capital intensive goods. International

[44] Chapter 4, Section 4.4.1 presents a simple comparison of autarky with a fully integrated two-country model that shows scale effects at work.
[45] The relationship between spillovers and trade is discussed in greater detail in Chapters 4 and 5.

competition forces the import goods sector to contract, reducing the price of human capital and therefore causing a re-allocation effect to the research sector (Grossman and Helpman, 1991).

In their model that captures the spillover effect, the market enlargement effect and the possibly conflicting allocation effect, Rivera-Batiz and Romer (1991) conclude that, at least for reasonable levels of openness to trade, the positive effects outweigh the negative. Therefore, there is a positive overall relationship between openness to trade and long run growth. Grossman and Helpman (1991) are more guarded in their conclusions. Aghion and Howitt (1998) conclude that trade and growth are most likely to be positively related, though not without reservations. The empirical evidence on support of these conclusions is discussed in the following Section.

So far the discussion has concentrated on the expanding product-variety model of R&D. What are the effects of opening up the increasing-quality model of R&D of Aghion and Howitt (1992)? The spillovers, market enlargement and allocation effects act broadly in the same way in both models (Aghion and Howitt, 1998). However, the consequences of increasing intensity of competition may be different (see Aghion and Howitt, 1998, Chapter 2, for example). This issue is discussed briefly in Sections 2.5 and 2.6 and in more detail in Chapter 3.

2.4.3: Does Trade Drive Growth?

Both the neoclassical model (suitably extended) and R&D-based endogenous growth models predict that trade will affect growth, either in transition to steady state or in the long run. The neoclassical frameworks of Lee (1993) and Ventura (1997), for example, suggest growth should be positively related to openness to trade. The discussion of R&D-based endogenous growth models in Section 2.4.2 suggested growth could be either positively or negatively related to openness to trade. The empirical growth and trade literature provides a great deal of evidence on the question of the existence and nature of these predicted effects, which can help clarify the nature of the overall growth/trade relationship.

[46] The Stolper-Samuelson effect.

Discussion of micro-level evidence of the effects of trade on growth is left until Section 2.5. Here, I concentrate on macro level evidence (ie: at cross-country level). Section 2.3 discusses the empirical growth literature at some length, but abstracts from consideration of trade. Nevertheless, the empirical trade/growth literature is closely related to the literature described in this earlier section, and in many cases, the papers discussed previously also play a key role in explaining the trade/growth relationship. In particular, much of this literature is based on cross-country growth regressions.[47]

A large literature exists in which some form of cross-country growth regression is carried out, including variables to capture openness to trade on the RHS (see, for example, Edwards, 1992; Barro and Sala-I-Martin, 1995; Lee, 1993; Henrekson et al, 1996; Frankel and Romer, 1999). Lee (1993) finds a negative relationship between growth and tariff rates and growth and black market premia on exchange (used to proxy for trade distortions/barriers) that is consistent with the predictions of an extended open-economy neoclassical growth model. Barro and Sala-I-Martin (1995) provide evidence of similar negative relationships between growth and black market exchange premia and growth and tariffs. Non-tariff barriers (essentially a measure of 'red-tape') are found to be insignificant. Edwards (1991) reports similar results, employing a large number of different (but closely related) measures of trade orientation. Henrekson et al (1996) find evidence of a significant positive effect on growth of an EC/EFTA dummy for highly integrated economies, which seems to stand up to a fairly intensive sensitivity analysis.

Although all these papers suggest the presence of a significant positive growth/trade relationship, they tend to suffer from a number of potentially serious problems, as discussed in Section 2.3. The results should therefore be viewed with an element of caution. On top of the more general criticisms of the growth regression literature, there are inherent difficulties in measuring openness. Results are often not robust to alternative measures of openness (see Levine and Renelt, 1992) and many of these alternative measures themselves are often uncorrelated with each other (Pritchett,

[47] Aghion and Howitt (1998) note that comparative case studies also provide a rich source of evidence on the effects of openness to trade on growth.

1996). Harrison (1996) argues that despite this problem, the evidence of a significant positive growth/openness relationship is still strong.

Another difficulty with these studies is the possible endogeneity of some of the openness measures (Frankel and Romer, 1999). A positive relationship between trade volumes, say, and incomes or growth may be reflecting the fact that richer countries tend to trade more than poorer countries rather than the reverse causality we are looking for.[48] To control for this, Frankel and Romer (1999) estimate the trade/growth relationship using geographical proximity measures to instrument for trade. They find a strong positive relationship.

In summary, even given the problems associated with cross-country growth regressions, the weight of evidence from this literature supports the existence of a significant positive relationship between trade openness and growth. In terms of the Rivera-Batiz and Romer (1991b) model, this evidence suggests that the positive scale effects discussed in the previous section outweigh any negative allocation effects, as they predict. However, the empirical evidence does not necessarily allow us to discriminate between alternative explanations of growth. The positive trade/growth relationship suggested by the bulk of the empirical literature is consistent with an extended neoclassical model as well as R&D-based endogenous growth models.

A substantial literature also exists on measuring the growth effects of knowledge spillovers across countries, where trade is taken as the primary mechanism by which knowledge spreads from one country to another (see, for example, Coe and Helpman, 1995; Coe et al, 1997; Engelbrecht, 1997). This literature provides some strong evidence specifically for the technology-push effect, and it is discussed briefly in Section 2.6 and at some length in Chapter 5.

[48] A substantial literature exists that suggests a positive relationship between export volumes and non-price competitiveness (technology level), for example. See Fagerberg (1988) and Piermartini and Meliciani (1998) for further discussion of this literature.

2.5: Micro R&D Analysis

So far the discussion has concentrated on the macro R&D, growth and trade literature. However, a substantial literature on these topics also exists at the micro-economic level. This literature is an essential component of our overall understanding, both as a source of micro foundations for macro models of R&D-based growth, and also because some of the relationships between R&D and trade can only be observed at a more disaggregated level. In this section I concentrate on just a few parts of this literature which are of particular relevance in the context of international trade. Section 2.5.1 discusses the innovation and market structure literature and the implications of opening up markets to international competition. Section 2.5.2 reviews the literature on R&D co-operation across firms and discusses possible effects of international trade. The reduced research-redundancy effect of economic integration is also briefly discussed. Section 2.5.3 briefly touches upon some other important considerations in the R&D literature, such as the effects of financial constraints, subsidisation and the inter-relationships with skilled labour.

2.5.1: R&D, Market Structure and Integration

The starting point for a discussion of market structure and innovation is a comparison of the incentives for innovation between monopoly and competitive markets. Tirole (1988) and Aghion and Howitt (1998) discuss this comparison in some detail. It is widely accepted that innovation requires some form of *ex-post* monopoly power (otherwise there is no profit incentive), and this is a crucial assumption of the models of R&D-based growth discussed in Section 2.2. The usual mechanism through which this ex-post monopoly power is enforced is patents, which act to protect property rights over research findings and innovations. A literature exists examining the precise nature of patents and their effects in some detail. For example, Takalo and Kannianen (1997) challenge the usual argument that tighter (ie: wider and/or longer) patents speed up technical progress.

Is it also the case that technological progress needs *ex ante* monopoly power, as suggested by Schumpeter (1934)?[49] Monopolies may be less susceptible to the uncertainties involved in R&D and better able to finance R&D internally (see, for discussion, Kamien and Schwartz, 1982; Tirole, 1988).[50] On the other hand, monopolies may be less 'hungry' for higher profits from innovation than competitive firms earning normal profits. One factor behind this is the 'Arrow effect' that monopolies have to gain extra profits over and above their current monopoly profits to make innovation worthwhile (Arrow, 1962). Also, monopolies may not need to innovate because they are in such a good position to quickly imitate any innovations in their markets (Baldwin and Childs, 1969). Gilbert and Newberry (1982) argue that monopolies may engage in patent-shelving, where they buy patents to new developments to protect themselves from competitors, but do not implement them (see Tirole, 1988, for a discussion of this point).

Aghion and Howitt (1998) identify three recent developments in the literature that give further support for a positive competition/R&D relationship. Firstly, building on a suggestion made by Leibenstein (1966), the corporate finance literature often assumes managers of large companies are mainly concerned with preserving their private benefits of control over the company while minimising effort. This is made possible by the 'slackness' in monopoly profit margins, but is not possible in more competitive markets.[51] Secondly, introducing 'tacit' knowledge as a requirement for innovation prevents leap-frogging and therefore gives rise to more neck-and-neck competition which increases incentives to engage in R&D.[52] Thirdly, re-allocation of workers from old lines to new lines may be stimulated by increasing competition, raising the rate of technological progress.

Even if we accept that the weight of argument supports a positive relationship between competition and R&D incentives, it is still the case that different models predict different relationships depending on the nature of competitive behaviour and the type of innovation under consideration. For example, the relationship between process

[49] The discussion is focussed on the product market. The research market is considered in the following section. Schumpeter argued that monopoly in the research market was detrimental to technological progress (Aghion and Howitt, 1998).
[50] Section 2.5.3 discusses the importance of internal financing of R&D.
[51] Empirical evidence presented by Nickell et al (1997) supports this argument.

R&D and market concentration may be of opposite sign to that between new-product R&D and concentration (Cohen and Levin, 1989). Sutton (1996) suggests there may be a similar discrepancy between new-product R&D and quality-enhancing R&D. The relationship for a given type of R&D may also be different under Bertrand and Cournot competition (see Chapter 3 for a discussion of this point).

Given the overall ambiguity of the theoretical relationship between market structure and innovation, the empirical literature takes on added importance. However, the empirical literature generally reflects this ambiguity. Cohen and Levin (1989) provide a comprehensive discussion of this literature up to the end of the 1980s, pointing out the added complication of the likelihood that the causality of the competition/R&D relationship runs in both directions. This literature is discussed in more detail in Chapter 3. A more recent review of the empirical literature can be found in Aghion and Howitt (1998)[53]. They argue that the weight of evidence supports a positive relationship between competition and innovation (see, for example, Blundell et al, 1995; Nickell, 1996).

Section 2.4.2 identified a possible effect of integration on R&D-based growth through changes in the intensity of competition. This 'increased competition' effect has not traditionally attracted as much attention in the trade/R&D literature as the more widely discussed demand-pull and technology-push scale effects, although recently interest has been growing. Nonetheless, at micro level, intensity of competition is one of the key mechanisms through which integration can affect R&D incentives. If a positive net relationship exists between the intensity of competition and R&D incentives, then integration, which acts to intensify competition between firms in different markets, should lead to increased R&D incentives. However, there are situations where this might not be the case.[54] Chapter 3 examines this issue in detail.

---

[52] This point is discussed in more detail in Chapter 3.
[53] Chapter 7.

2.5.2: Duplication, Co-operation and Integration

In Section 2.4.2, two of the suggested ways in which the introduction of trade might affect growth was by reducing the incentives for duplication in research across countries and by affecting the incentives to co-operate in R&D across national boundaries.

Grossman and Helpman (1991) present a clear and detailed discussion of the duplication point.[55] The following briefly summarises their discussion. As touched upon in the previous section, if firms can differentiate their products from their competitors they will face less severe product market competition. Increasing the intensity of competition in a given market will therefore increase the incentives to differentiate. Opening up a market to international competition will act in the same way by encouraging entrepreneurs to pursue new and distinctive ideas and technologies. As a result, trade alleviates the duplication of research effort in different countries. In other words, introducing international trade reduces research redundancy in the world economy and increases the aggregate productivity of resources devoted to R&D. So trade can act through this mechanism to increase the long run growth rate.

What of R&D co-operation across firms? The Cecchini Report claims that one of the likely effects of the implementation of the Single European Market (SEM) will be the 'rapid development of cross-frontier business co-operation for R&D' (Cecchini, 1988). This is one way in which European firms may be able to regain technological leadership in the world economy and therefore lead to increased growth. What is the evidence for this prediction? It seems reasonable to assume that economic integration will change the nature of competition and therefore will change the effects of co-operation in R&D (henceforth Research Joint Ventures) as well as changing the incentives to set up or join Research Joint Ventures (RJVs). It is less clear which direction these effects will take.[56]

---

[54] Aghion and Howitt (1998) present a model in which integration can reduce R&D incentives in industries that are not neck-and-neck.

[55] Chapter 9, Grossman and Helpman (1991).

[56] Tirole (1988) suggests the need for further research on these and related questions.

A sizeable literature exists that examines the effects of RJVs (see, for example, Katz, 1985; Kamien et al, 1992; Choi, 1993). It is suggested that RJVs may stimulate innovation for a number of reasons. Firstly, they prevent wasteful duplication of research across firms. Secondly, they may help knowledge to diffuse more effectively between the venturers. Thirdly, given that possible competing firms are co-operating the market failure of incomplete appropriation (spillovers) may be reduced. Lastly, high fixed costs of R&D can be shared and economies of scale realised. On the other hand, RJVs may be anti-competitive in the research market and in the product market, leading to possibly both dynamic and static inefficiency. Ordover and Willig (1985) conclude that, generally, the benefits of greater co-operation can be expected to outweigh the costs, but this may not always be the case. An RJV is socially desirable when the primary R&D competition facing the co-operating firms is from others and undesirable when the primary competition would have come the venturers themselves. Katz (1985) concludes that the social benefits of RJVs fall as the degree of product market rivalry increases.[57]

There is no such consensus on the question of how integration will affect the incentives to set up or join RJVs. Integration will change the nature of competition in markets (as discussed in the previous section) and therefore affect the social costs and benefits of RJVs. However, it is less clear what effect this might have on the private costs and benefits. An increased number of competitors is likely to make it more difficult for a small number of firms to hold back the pace of innovation. The RJV incentive effects of increased intensity of competition could feasibly go the other way. In Chapter 3 I present a simple model in which increased intensity of competition between similar firms increases the incentives for co-operation in R&D across boundaries. It is likely that this would not be the case when firms have different technology levels. Hagendoorn and Schakenraad (1993) catalogue a growth in absolute numbers of RJVs in Europe alongside integration, but due to overall growth, argue that their relative share has not increased. Overall, the question of how

[57] Grossman and Shapiro (1986) argue that US Antitrust legislation is too harsh and that in many cases the benefits of RJVs outweigh any loss of static efficiency. The European Commission explicitly recognises this in paragraph 3, Article 85 of the Treaty of Rome, which allows for the exemption of some collusive behaviour in cases where there are substantial benefits from R&D co-operation (Jacquemin, 1989).

economic integration affects the incentives to set up or join RJVs seems to merit further research.

2.5.3: Some Other R&D Considerations

A great deal of the micro R&D literature has been skipped over by focussing only on market structure and RJVs. For example, I have not discussed the pros and cons of R&D subsidies or taxes, the relationship between R&D and the demand for skilled labour or the problems of financing R&D. These are issues worthy of more than the brief discussion here. Time and space constraints, however, prevent a more detailed discussion.

When it is optimal for the government to intervene in R&D by providing subsidies or by imposing taxes? In simple aggregate terms, a subsidy is optimal when positive externalities dominate negative externalities. In the Romer (1990) model, the positive inter-temporal externality suggests an R&D subsidy would be welfare enhancing. In the Aghion and Howitt (1992) model it depends on the relative strengths of this positive externality and the negative business stealing effect. However, the basic picture is complicated by a number of micro-level factors. Aghion and Howitt (1998) argue that the possibility of 'slack' in potential innovating firms calls for an R&D tax.[58] Leahy and Neary (1999) argue that the presence of RJVs may also justify the imposition of an R&D tax rather than a subsidy, as the externality may be over-internalised. It may be that targeted subsidies and taxes are optimal, at least where the government has good information (Aghion and Howitt, 1998).[59]

In the models of R&D-based growth discussed in Section 2.2, there was a clear role for skilled labour (or human capital) in both the research and the production sectors. The links between skilled labour and technological progress are the subject of a great deal of micro literature also. Nickell and Nicolitsas (1997) argue that firms may hold back expenditure on R&D when they face a shortage of skilled labour, which is supported by empirical evidence for the UK. This is part of the overall picture of technological progress increasing the relative demand for skilled labour (see, for

[58] Aghion and Howitt (1998), Chapter 7.
[59] For example, the defence industry.

example, Machin and Van Reenen, 1998). An interesting issue in the recent literature is whether the observed increasing wage differentials between skilled and semi-skilled or unskilled workers (particularly in the US and UK) is primarily driven by skill-biased technological progress or the globalisation of trade. Wood (1995) argues that growth of trade with developing countries explains most of the reduction in the relative demand for unskilled labour in industrial countries. Berman et al (1994), for example, argues the opposite.

How do firms finance R&D? It was argued in Section 2.5.1 that monopolies may encourage R&D because they are more able to finance R&D internally. This highlights an important point. R&D is faced by a capital market failure because of the informational asymmetries involved. R&D outcomes being uncertain, a capital provider cannot necessarily identify whether failure of an R&D project is due to genuine 'accident' or lack of effort.[60] Consequently, the majority of R&D is financed by firms internally from retained profits. Hall (1992) and Nickell and Nicolitsas (1997) find a negative effect of financial constraints on firms' R&D expenditure in the UK.

2.6: Spillovers, Imitation and Growth

A common theme running through the literature discussed so far has been the potential for knowledge spillovers between firms and countries and the growth implications of such spillovers in the presence of increasing international trade. In this section, I examine the spillovers literature and discuss the implications of the related imitation literature for growth. Since Chapter 5 presents a detailed discussion of the spillovers literature itself, it is the growth implications that are the main focus here.

2.6.1: Knowledge Spillovers, Growth and Integration

The simplest interpretation of 'knowledge' is that it is a by-product of R&D. A firm that invests in R&D may not only innovate, but also add to the current stock of general

[60] In other words, there is a moral hazard problem. This is discussed in more detail in Chapter 4, Section 4.5.1.

scientific know-how by its research.[61] The characteristics that make this general stock of scientific knowledge interesting are its non-rivalry and non-excludability.

A number of mechanisms have been suggested to explain how such knowledge diffuses from one organisation or country to another. For example, Cameron (1996) suggests information leaks, imperfect patenting and the movement of skilled labour as possible mechanisms. Grossman and Helpman (1991) argue that knowledge diffuses across national boundaries through commercial interactions between foreign and domestic agents. For example, know-how can be gained by reverse engineering imported goods or through the information given by importing firms to exporting suppliers. Hejazi and Safarian (1998) argue that FDI is another important conduit for spillovers. Grossman and Helpman argue that integration, by increasing commercial interaction between foreign and domestic agents, will lead to increased knowledge spillovers.

A closed economy is limited in its technical capabilities to what it can learn through its own research. An open economy can gain know-how from foreign stocks of knowledge, whilst enabling foreign firms to learn from the domestic stock of knowledge (Grossman and Helpman, 1991b). The R&D-based models of endogenous growth discussed in Section 2.2.4 assume a country's innovative capacity depends on its stock of knowledge. If knowledge can be accumulated from abroad in addition to from domestic research, then a country can accumulate more knowledge in less time, driving a higher rate of long run growth. So, knowledge spillovers between countries can lead to higher growth. By making international knowledge spillovers easier, economic integration therefore stimulates growth. In Section 2.4.2 this is referred to as the technology-push effect.

However, whilst it is entirely possible that countries gain from foreign spillovers, they may also lose out because they cannot appropriate all of the benefits from their own research (Grossman and Helpman, 1991b). The disincentive effect of knowledge spillovers is further discussed in Chapter 3 and in Chapter 5 of this thesis. Given the existence of such a disincentive effect together with the positive technology-push

[61] This is the basis of the LBD model of Romer (1986).

effect, the question as to whether spillovers are net contributors to world growth is best answered empirically. This empirical literature is discussed in some detail in Chapter 5. What follows is a brief summary.

Positive country-level domestic TFP-growth effects from foreign R&D are found by Coe and Helpman (1995), Nadiri and Kim (1996), Bayoumi et al (1999) and Coe et al (1997). Nadiri and Kim (1996) provide evidence that suggests significant heterogeneity in the importance of foreign spillovers across countries. Italy, for example, is found to benefit from foreign R&D to a far greater extent than the US. Engelbrecht (1997) also finds evidence of such diversity, but to an even greater extent. In fact, technology-source countries (eg: US, Canada, West Germany) display a negative relationship between domestic productivity and foreign spillovers. This may reflect the negative incentive effect of spiilovers. Engelbrecht (1997) also finds some low-technology countries to have weak and sometimes negative spillover effects, which he speculates could reflect the need to have reached a certain threshold level in domestic capability to be able to benefit from technological advances abroad. This is similar to the appropriate technology idea put forward by Temple (1998), for example, as discussed in Section 2.3. Such considerations allow us to reconcile conflicting evidence of convergence and divergence and can provide explanations for the convergence clubs observed by Quah (1997). In their simplest (positive) form, the existence of international knowledge spillovers allows us to reconcile basic models of endogenous growth and evidence of convergence.

Of course, spillovers are not just a country level phenomenon, but can occur at a more disaggregated level. Bernstein and Nadiri (1988) find evidence of significant spillovers at firm level. A substantial literature exists based on the premise that knowledge spillovers at firm or industry level may be localised, leading to clustering. This can have significant implications for growth in the context of economic integration (see, for example, Krugman and Venables, 1995).

2.6.2: Imitation and Growth

Models of growth through imitation follow on directly from the knowledge spillovers literature. However, there is a clear distinction to be made between general scientific knowledge capital, which is the stuff of spillovers so far discussed, and specific blueprints for innovation, which is the stuff of imitation. Both types of diffusion take place. Entrepreneurs closely copy the designs and processes that rivals have developed as well as applying abstract concepts in new and different contexts (Grossman and Helpman, 1991a). Empirical evidence supports this assertion. For example, Mansfield et al (1981) find about 60% of innovations are imitated within 4 years. It is not difficult to think of examples of successful imitation. One such example is the extent to which rival PC producers were able to produce IBM-compatible PCs. Secondly, manufactured exports from Newly Industrialised Countries (NICs) are dominated by goods once produced in countries such as the US or the UK (Grossman and Helpman, 1991a). Young (1992) points to the active support of the Singapore government for foreign investment and expertise in several leading edge industries in which indigenous Singapore industry is now very successful, for example.[62]

Barro and Sala-I-Martin (1995) present a simple model of growth through imitation.[63] The following is a brief summary of the structure and implications of this model. There are two countries, one of which is the leader (innovator) and the other the follower (imitator). This imitation can either take place through indigenous firms in the follower country stealing designs for new goods for which the innovator receives no payment, or through the innovator investing in imitating its own goods in the follower market.[64] Assume the leader country develops new differentiated intermediates as in the Romer model, and that the current set of intermediates is N. If the follower is able to imitate some of these goods, it gets monopoly power in its market, thus providing the incentive. Imitation is assumed to be cheaper than innovation.[65]

[62] Foreign Direct Investment and its implications for indigenous industry is discussed in more detail in Chapter 5 of this thesis.

[63] Chapter 8.

[64] The model assumes intermediates are not traded. This is a rather extreme assumption, which in some respects is critical to the conclusions of the model. Empirical evidence suggests that intermediates are traded across countries and that the most appropriate formulation for such a model would be to allow for both trade and non-traded intermediates (Coe and Helpman, 1995). The analysis of the model would still apply for non-trading sectors.

[65] Mansfield et al (1981) find evidence to suggest that imitation costs average just 65% of innovation costs.

What are the growth implications of this basic model? The follower grows faster than the leader in the short run until it catches up with all N designs. Barro and Sala-I-Martin (1995) argue that Japan may have reached such a point in its technological development. The follower will grow faster the larger the variety, N, of intermediates there are, because it can cherry-pick those goods it is best able to imitate. Also, the lower the cost of imitation, the faster the follower grows. In the steady state, there is a common, constant long run growth rate.

The model displays both scale effects and convergence. As such, it provides an example of how the diffusion of technology across countries can reconcile the assumption of constant returns to capital common to the endogenous growth literature, with the empirical evidence of convergence often used to support the neoclassical model. This point is discussed in Section 2.3.2 and returned to again in Chapter 5.

2.7: Summary and Concluding Remarks

In this last section, I draw together and summarise the preceding discussion of the R&D, growth and trade literature and identify the main fields within the literature and the main themes running through it. I provide a brief overview of the most recent research in the field and a perspective on where the literature might be going in the next few years. Lastly, I suggest possible areas for further research arising from the discussion of the literature, and show how the following three Chapters can be placed in this context.

2.7.1: Summary

The growth literature with all its related fields is of considerable size and highly dynamic. Only a very small part of this literature has been discussed here, but hopefully the discussion has been broad enough to touch on many of the most interesting topics within this literature. A number of key themes run through these topics and these have been considered throughout the discussion.

Modern day growth literature has its roots in early classical economics of the 18^{th} and 19^{th} Centuries. These roots have developed into the two paradigms of neoclassical and endogenous growth theory outlined in Section 2.2. Endogenous growth theory itself is a broad church including models were growth is driven by accumulation of knowledge through learning-by-doing, human capital, government provision of public goods that enter the production function and by deliberate R&D carried out by profit-maximising entrepreneurs.

R&D-based endogenous growth models were made possible by the crucial step of dropping perfect competition. This allows them a considerable realism when compared to the more abstract externality-driven models of growth such as LBD. However, they do suffer from the same reliance on constant returns, which, it has been argued, are not necessarily supported by the available evidence. Nonetheless, R&D-based models of endogenous growth have proved to be the most vibrant of the new models and certainly the most widely used basis for recent research into international trade and growth issues. In this trade and growth field, the most common models of R&D-based growth are new-product models, first introduced by Romer (1990). Significant research also exists that assumes R&D adds to the quality of existing products.

At the same time as the theoretical literature has developed, there has been a considerable explosion in empirical growth research, which it has been argued was set off by the availability of detailed cross-country data sets in the last ten years. This research has motivated many of the theoretical developments and has been used widely to try to support various models and criticise others, as discussed in Section 2.3. It is unclear whether the empirical growth research body has come down in favour of either neoclassical growth models or endogenous growth models. Many researchers argue that the two strands of the literature should be considered complements rather than substitutes in the light of this empirical literature. Indeed this reconciliation between the two 'camps' of neoclassical and endogenous growth is one of the main characteristics of recent growth research, as discussed in the following section.

Trade has proved to be a significant factor in the overall growth picture in much of the empirical research. The most development on the theory side of trade and growth has been using models of growth through new-product variety R&D (eg: Rivera-Batiz and Romer, 1991a, 1991b; Grossman and Helpman, 1991a). This is partly driven by the lack of any long term trade effects in the neoclassical model, as discussed in Section 2.4. In contrast, R&D-based growth models display a number of significant trade-related effects on long run growth. The most commonly discussed effects are the scale effects of market enlargement (demand-pull) and access to wider international knowledge stocks (technology-push). Unfortunately the empirical evidence for such scale effects is at best ambiguous.

The whole area of stocks of knowledge and the possibility of international spillovers of knowledge has spawned a considerable theoretical and empirical literature. This is reviewed briefly in Section 2.6 and returned to in Chapter 5 of this thesis. The implications of this literature for the main body of growth literature are enormous, particularly with respect to predictions of convergence and divergence that distinguish the more basic forms of alternative growth models.

In addition to the macro-level literature on R&D, growth and trade, there is a considerable micro-level literature that both provides detailed foundations and allows us to examine issues obscured at aggregate levels. Section 2.5 considers two such topics. One, the relationship between market structure and R&D incentives, is the subject of a considerable literature primarily in the IO field. Only recently have the implications of economic integration been considered in the context of this literature. The second micro literature considered is that concerning co-operation in R&D and incentives. As yet, the implications of economic integration are largely unexplored by this literature.

It is possible to identify a number of key themes that run throughout this review. One way of interpreting these themes is as the fault lines that run through the growth literature. These are summarised below.

- Constant, diminishing or increasing returns?
- Perfect or imperfect competition?
- How much is explained by technological progress?
- Convergence, divergence or both?

- If trade is important for long run growth, why is there not stronger evidence of scale effects?
- Are there significant micro-level effects of integration on R&D-based growth?

The fact that many of these questions are as yet unresolved suggests that the study of economic growth and international trade will be a vibrant field for many years to come.

2.7.2: Where Are We Now and Where Are We Going?

The most general answer to the first question is that we appear to be at a stage where the neoclassical and endogenous growth paradigms are treated by many as being complements rather than being in conflict (see, for example, Aghion and Howitt, 1998). Such a position can include those reluctant to abandon the neoclassical model and those as yet unconvinced by the newer growth models (eg: Solow, 1994) and those who firmly believe endogenous growth models are the only way forward (eg: Romer, 1994). In the context of international trade, the vast majority of the literature falls into the latter camp, given the lack of long run trade effects in the neoclassical model.

Some of the conflicting predictions of endogenous and neoclassical growth models can be reconciled by the introduction of international trade. In particular, consideration of knowledge spillovers and imitation can allow convergence in models without diminishing returns. The potential importance of knowledge spillovers has stimulated a great deal of interest in the literature. Recently, attempts have been made to look beyond goods trade for possible mechanisms through which spillovers might take place. The fact that spillovers might not have the same effects in different countries is also a topic that has attracted some recent interest (see Chapter 5 of this thesis for a discussion of these points).

Models can also be built where growth is driven by both the accumulation of physical capital and profit-driven R&D, and these too provide a mechanism to reconcile some of the predictions of neoclassical and endogenous growth (see, for example, Aghion and Howitt, 1998, Chapter 3). Such models are clear examples of the potential complementarity of the two approaches. The relationships between capital investment and R&D investment at firm level are the subject of a number of recent empirical papers (see Chapter 4 of this thesis for a discussion).

A perceived barrier to further development of the growth and trade literature is the lack of any consistent evidence for scale effects. This is the basis for a powerful critique of endogenous growth models that predict such scale effects. A number of recent models that retain some of the aspects of endogenous growth but reject such scale effects are surveyed in Eicher and Turnovsky (1999). The 'growth through R&D and capital accumulation' model of Aghion and Howitt (1998) is an example of such a model. Scale effects are avoided in this model by linking the *complexity* of technology or the *spread of research effort* to the size of the economy. These 'Non-Scale' models of growth are one direction in which the literature is going.

Another exciting recent development is the interest in market structure considerations in the growth and trade literature (see, for example, Aghion and Howitt, 1988, Chapter 7). This avenue of research offers the potential for micro long run trade effects on growth that do not act through scale effects. This is an important motivation behind Chapter 3 of this thesis.

In general, the empirical growth literature is also in a state of rapid development. A good recent survey is Temple (1999). One of the areas of development that Temple highlights is how some of the widely perceived weaknesses of cross-country growth regressions are being addressed. Another interesting empirical development is the apparent emergence of convergence 'clubs' (see Quah, 1997). Given that overall empirical evidence on convergence is still mixed, this offers an exciting new avenue for research.

Even from such a brief discussion of recent research, which by no means covers the scope of the developments in the growth and trade literature, it is clear that the field is a very exciting one, with a great deal of promise for real progress in the next few years. Many of the developments in recent years have opened up new avenues for research that offer the potential to address issues such as the convergence question or the existence of non-scale integration effects on growth. The theoretical literature is now developing hand-in-hand with the empirical literature (Barro and Sala-I-Martin, 1995; Aghion and Howitt, 1998; Temple, 1999) and the quality of data is likely to improve considerably as time goes on.

Specifically, I identify three areas where there is a clear need for further research in the light of the existing literature. In some sense, these can be interpreted as going back and filling in some of the gaps left by the rapid development of the field. Firstly, the micro-level effects of international trade on R&D incentives and hence growth offer a very promising avenue for research.[66] In particular, the effects of changing market structure on technological progress have not been fully worked out, although a considerable literature exists on the relationship between market structure and R&D in the absence of trade. Chapter 3 of this thesis addresses this issue. I also briefly consider the incentives to co-operate in R&D under integration in Chapter 3, which also have the potential to be an interesting avenue for further research. This issue has not yet attracted much attention in the literature.

The ambiguity of much of the existing empirical evidence on growth and openness to trade implies a real need for further research. Developments of the techniques for cross-country growth regressions offer one way forward. Improved data on openness, in particular, offer another. Chapter 4 takes a slightly different approach by examining the empirical relationship between openness to trade and R&D *incentives*. I argue that this is an important gap in the literature, given the critical role played by R&D incentives in models of trade and growth. Also, given the potential for significant integration effects at a disaggregated level, as suggested by Chapter 3, for example, it seems appropriate to study the openness/R&D relationship at this level.

The spillovers literature is a particularly vibrant field, with huge potential to improve our understanding of growth. As discussed above, spillovers allow convergence to be reconciled with constant returns, for example. However, despite a considerable empirical literature measuring the growth effects of spillovers across countries, industries and firms, there is still a significant degree of ignorance as to exactly how knowledge spills across boundaries and what effects such spillovers have in different

[66] That integration might affect the intensity of competition and therefore incentives to innovate and that integration might affect incentives to co-operate in R&D. See Sections 2.4 and 2.5 for discussion.

situations. Many recent papers suggest that FDI should be studied as an alternative mechanism to goods trade through which spillovers might occur across countries. Chapter 5 follows this suggestion for the UK at industry level. This disaggregate approach allows a great deal of heterogeneity to be picked up across sectors.

Chapter 3

Strategic R&D and Economic Integration

3.1: Introduction

If there are dynamic efficiency gains to be made as a result of economic integration then it is possible that integrating a set of economies may lead to a higher long run rate of growth. This possibility has attracted some attention in recent years as regional and global integration has progressed, A variety of ways in which the R&D and subsequent innovation decisions of profit maximising entrepreneurs might be affected by integration have been suggested (see Section 2.4 for a general discussion). The most generally used arguments are that integrating a number of economies enlarges the market and knowledge base available to firms and may therefore enable them to exploit scale economies in R&D. These scale effects are outlined below.

The technology-push hypothesis discussed by Phillips (1966) argues that there is a relationship between the level of underlying scientific knowledge in an economy and the rate of technological progress. Rivera-Batiz and Romer (1991a) argue that if the rate of innovation is positively related to the current stock of knowledge capital in an economy, economic integration will allow access to an international stock of knowledge and therefore increase the rate of technological progress. Rivera-Batiz and Romer (1991a) also discuss the importance of what Schmookler (1966) referred to as the demand-pull hypothesis, where integration presents a bigger market and therefore a greater reward to a successful innovator. This market enlargement effect will have an unambiguously positive effect on incentives to innovate, other things being equal. The problem with these predictions is that empirical evidence for such scale effects is, at best, mixed.

At the same time as these scale effects, integration will change the nature of the competition facing firms. Markets will be characterised by increased intensity of competition resulting in extra pressure on home market shares and added opportunities to steal shares in foreign markets. It is likely that such changes in the intensity of competition will have an effect on R&D incentives, as discussed in Section 2.5. These 'increased competition' effects of integration have not, so far, received as much attention in the literature as the demand-pull and technology-push scale effects. However, a considerable wealth of knowledge on these matters can be found in the IO literature. Drawing on this literature, this Chapter takes a closer look at the competition effects of integration and their implications for R&D and long run growth.

The remainder of this Chapter is set out as follows. The following Section discusses the IO literature on competition and innovation and some of the particular characteristics of these models that may have critical implications when they are extended to consider integration. Section 3.3 presents a simple two-firm two-country model where integration is modelled as an increase in the substitutability between differentiated goods. Two interesting special cases of this model are also considered: Firstly, when firms are allowed to co-operate in research and secondly when the extent of appropriability of R&D is related to the degree of integration. Section 3.4 extends the model to allow for asymmetric initial costs of production. Section 3.5 further extends the basic symmetric model to an n-sector general equilibrium model in which national growth rates are shown to be positively related to the intensity of competition and therefore the degree of integration. Section 3.6 discusses the likely implications of relaxing some of the key assumptions of the model. Finally, Section 3.7 summarises and concludes.

3.2: Intensity of Competition, Integration and R&D

How might increased intensity of competition affect R&D? This is discussed in some detail in Section 2.5.1. The loose conclusion made there is that the arguments for positive and negative competition effects on R&D balance out in favour of an overall positive relationship. More intense competition will make a firm's share of its market more vulnerable to successful innovation from its competitors. This will give added incentives to keep up with R&D. Equally, successful innovation will allow the firm to more easily expand its share of a given market at the expense of its competitors by stealing their shares (Sutton, 1996). In other words, there is strategic complementarity in R&D between firms (Miyagiwa & Ohno, 1997). These strategic considerations add to R&D incentives when competition is intensified.

However, as ever, the story is not that simple. The relationship between concentration and R&D incentives may be very different for different types of innovation and different types of competitive behaviour. For example, the relationships between R&D aimed at reducing production costs and market concentration, and between R&D aimed at introducing new products and concentration may be of opposite sign (Cohen & Levin, 1989). Sutton (1996) finds a similar discrepancy between quality improving R&D and new product R&D.

Many authors have discussed the different implications for R&D incentives of modelling competition by quantity or by price (Cournot or Bertrand), mainly for process innovation. Brander and Spencer (1983), for example, argue that Cournot competition favours innovation in that one firm's cost reduction will lower the output of its competitors, whereas under Bertrand competition, a cost reduction will lower its competitors' price. In general, no real consensus has emerged as to the relationship between concentration and R&D under either Cournot or Bertrand competition. For example, for a homogenous good and process innovation, Dasgupta and Stiglitz (1980) find a negative relationship between R&D and competition. Qui (1994) finds a negative relationship under Bertrand and a U-shaped relationship under Cournot. Bester and Petrakis (1993) find a positive relationship between the degree of product substitutability and R&D incentives for both Cournot and Bertrand competition, and that incentives to innovate are higher (lower) under Bertrand than Cournot for a high (low) degree of substitutability.[67]

This lack of consensus highlights the sensitivity of results to the particular assumptions on which these models are built. I discuss four such assumptions below, which are of particular interest in the current paper. Firstly, the specification of linear demand functions is shown to have critical implications. Secondly, the degree of appropriability of the benefits of R&D can be important. Thirdly, the degree to which firms can co-operate in research has implications for the competition/R&D

[67] The degree of substitutability is the measure of competition intensity I adopt here.

relationship. Finally, the symmetry or otherwise of firms in terms of technology levels may also have a significant effect on the nature of the competition/R&D relationship.

Given the overall ambiguity of the competition/R&D relationship in the theoretical literature, a critical role for empirical evidence is implied. However, the empirical literature is also lacking consensus and evidence is mixed. Positive relationships between competition and innovation are found by Horowitz (1962) and Mansfield (1968), for example. Williamson (1965) and Bozeman and Link (1983) find negative relationships. Scherer (1967) finds evidence of a non-linear inverted U-Shaped relationship, with an optimum degree of competition. Sutton (1996) shows how apparent relationships are weakened when industry-specific effects, such as technological opportunity, are included. Problems with data, such as distinguishing between process and product innovations and the lack of reliable industry level elasticity estimates have prevented a definitive empirical examination of these matters. There is also the added complication of reverse causality between R&D and concentration. Cohen and Levin (1989) discuss these problems in a detailed review of the empirical market concentration and R&D literature, pre 1989.

A key assumption shared by the models of Brander and Spencer (1983), Bester and Petrakis (1993), Qui (1994) and Sutton (1996) is the specification of linear demand functions.[68] With such a linear demand specification, increasing the substitutability between goods, or increasing the intensity of competition, has the effect of shrinking the size of the market available to each firm. This effect is outlined and contrasted with the effect of increasing substitutability in the non-linear demand specification adopted in the current paper. It is this shrinking market effect that is behind Qui's predicted U-Shaped relationship, with R&D incentives being driven by a similar shaped marginal profits function. More specifically, R&D incentives are responding to the interaction between the negative shrinking market effect and the positive strategic market-share effects discussed above, which eventually dominate. In what follows, I adopt a non-linear demand specification that does not have this shrinking market effect. This is more appropriate to a study of integrating markets, whereas the linear

[68] In fact, Sutton presents both a relatively simple linear model and a less simple, but more general, bounds-approach model. He advises us to take both theoretical results from such simple linear models and the empirical results from related simple regression models with caution.

demand specification may be more appropriate to studies of entry into given markets. This is my first formal point of departure from the IO literature.

The usual linear demand specification (taken from Qui (1994)) is as follows:

(3.1): $p_i = \alpha_i - q_i - \gamma q_j$,

where γ ($0<\gamma<1$) is the degree of substitution between the two goods and i and j denote competing firms. Inverting and summing over i gives us total market demand:

(3.2): $q_i + q_j = \{(\alpha_i + \alpha_j) - (p_i + p_j)\}/(1 + \gamma)$.

This is clearly decreasing with γ.

In contrast, the non-linear demand specification adopted in the model of the following Section displays market-size stability when the degree of substitutability is altered, as shown by (3.3). For simplicity, I assume symmetric prices, $p_i = p_j = p$.

(3.3): $q_i + q_j = 2mp^{-\gamma}/2p^{1-\gamma} = mp^{-1}$,

Clearly, in this case, market size is invariant with γ.

Another common assumption in the IO literature on concentration and innovation discussed above is the full appropriability of the benefits of R&D, or in other words, that there are no knowledge spillovers.[69] The presence of such spillovers has been shown to have important implications for R&D incentives in a related literature on R&D co-operation (see, for example, D'Aspremont & Jacquemin 1989; Kamien et al, 1992). These papers argue that the incentives to invest in R&D will be lower if some of the benefits of that R&D spill over to firms in direct competition, because spillovers will lessen the relative advantage to be gained over these competitors. For firms not in direct competition, this is not such an important consideration. If we accept this argument, then integration should increase this spillover disincentive for R&D. This is

[69] Spillovers are discussed in detail in Chapter 5 and also in Section 2.6 of this thesis.

likely to have important implications for the overall integration/R&D relationship through intensity of competition effects, and is another point of departure from the papers discussed above.

The assumption of incomplete appropriability is given further importance in this context if we assume the level of spillovers is related in some way to the degree of integration between firms or countries. That knowledge spillovers across national boundaries might be positively related to the level of interaction between countries has been suggested by, among others, Grossman and Helpman (1991b). Empirical evidence discussed in Chapter 5 suggests that such spillovers are strongly related to trade flows and also possibly FDI. There is also strong evidence that flows of goods within countries are highly correlated with spillovers between firms and industries. If this is the case, then integration, by strengthening spillovers, may have a further disincentive effect on R&D. This possibility is considered in Section 3.3.3, and is shown to have some very interesting implications. In particular, the interaction of the positive strategic and negative spillover effects can lead to an inverted U-Shaped relationship between R&D incentives and integration, suggesting the possibility of an optimum level of integration.

In section 2.5.2 the literature on R&D co-operation was reviewed in the context of integration. This is an area of research that has been overlooked so far in the literature, although claims have been made that integration might stimulate co-operation and therefore dynamic performance (Cecchini, 1988). To recap, co-operation in research (or RJVs) may have benefits and costs. On the one hand, knowledge may diffuse more quickly between venturers increasing the social rate of return to R&D. Also, because firms explicitly co-operate, the disincentive effects of spillovers may be reduced. Further, high fixed costs in R&D can be shared and economies of scale realised. On the other hand, co-operative research agreements may be anti-competitive in both the research and production markets, which could lead to both dynamic and static inefficiency. According to Ordover and Willig (1985), generally the benefits of co-operation can be expected to outweigh the costs. However, as the degree of product market rivalry increases, a co-operative research agreement is less likely to lead to increased research (Katz, 1985). This is the key point here. By fostering anti-competitive behaviour in RJVs, integration may lead to reduced R&D expenditure and slower technological progress. This is considered in Section 3.3.2.

One of the most important fault lines running through the competition/R&D literature is the assumption of symmetry or the nature of the asymmetry assumed. Bester and Petrakis (1993) find a firm that has gained a cost advantage in the past is more inclined than its competitors to invest in further cost reductions.[70] In a dynamic setting, this would suggest a dominant firm outcome, such as found by Grossman & Shapiro (1987) and Harris & Vickers (1987), for example. In these models, the follower becomes increasingly disillusioned as the technology gap widens. However, the predictions of these models depend on assumptions made about the technological opportunities facing firms. In many cases, it is assumed that the best followers can do is catch up (see Aghion and Howitt, 1998, for an example). Alternative assumptions can lead to the opposite implications for follower incentives. In Section 3.4, I present a model where the low-technology firm has the greater incentive to invest in R&D, resulting in a stable symmetric equilibrium. A detailed review of such symmetry considerations is provided by Beath et al (1994).

In the following Section I set up a simple two-firm, two-country model in which integration is modelled by an increase in the substitutability in preferences between the two goods. Although commonly used in the IO literature to model increased competition, this is a novel approach to modelling economic integration. It is not without precedent, however. Danthine and Hunt (1994) adopt the technique in their study of integrating labour markets. What this approach allows is the isolation of the intensity of competition effect from other integration effects, such as demand-pull. By modelling integration in this way the market enlargement effect of integration disappears. A market may become twice as big, but a firm's share of this market is divided by half because the number of equal competitors doubles.[71] This is perhaps the most crucial assumption of the model presented in the following Section. By concentrating solely on the strategic intensity-of-competition effects of integration on

[70] This has unfortunate implications for the stability of their symmetric equilibrium.
[71] This is not necessarily the case if there is asymmetry in the model. Also, modelling integration as a reduction in tariffs, even in a purely symmetric model, does not allow us to abstract from market enlargement effects so easily.

R&D, this Chapter formalises an alternative mechanism through which integration can affect long run growth to the standard scale effects discussed above.

Before presenting the model, it is useful to consider the legitimacy of modelling integration in this non-standard way. The simplest justification is to think of integration as being a standardization process. For example, European integration may involve the standardization of electrical products to common European specifications. The benefit of this interpretation is that it requires no leap of imagination to swallow, although there is a cost in terms of the narrowness of the integration concept. However, if we are prepared to take a small intuitive step, this way of modelling integration can be thought of in much broader terms. Danthine and Hunt (1994) warn that we should not take it literally that integration manifests itself as a change in consumer preferences. However, from the producer's point of view, it is *as if* integration has caused preferences to shift. Consumers pay more attention to goods sold by foreign producers. Market shares become more sensitive to price differentials as competition is intensified. A simple example is presented below in order to clarify this concept.

Imagine two supermarkets, far enough apart so that consumers living near one cannot walk to the other costlessly and vice versa. This would manifest itself as if the food products offered by the two supermarkets were not close substitutes in preferences, with the supermarkets having monopoly power over the consumers living at their end of town. Overall market share would be insensitive to the prices set in the other supermarket at the other end of town. Now consider a reduction in the barriers to trade, such as the introduction of a free bus service between the supermarkets. Consumers could now shop at either supermarket costlessly, which would manifest itself as if the products offered by the supermarkets were closer substitutes. The market share of each supermarket would now be much more sensitive to the price setting of the other supermarket. Intuitively, we can consider integration in this way.

Appendix A reworks the model of Section 3.3, for the symmetric competitive case, with integration modelled in the more standard way as falling tariffs. The fundamental results of the model are unchanged in this alternative.

3.3: A Two-Country Duopoly Model of Integration with Process Innovation

At the expense of a modicum of generality, the following model allows the intensity of competition effects of integration on R&D to be formalised in a very simple way. The stylized nature of the model does not detract from the potential of its conclusions, however. It is left to further research to relax some of the more extreme assumptions in order to generalise the model's predictions. In the light of this, the model should be seen as a first step in formalising the intensity-of-competition effects of integration on growth through R&D, rather than an end product. This step is made all the more important given the continuing lack of convincing evidence for scale effects, such as those suggested by Rivera-Batiz and Romer (1991a).

Consider an industry consisting of two firms, each producing a single differentiated good. Assume for now that entry into the market is impossible, due to, say, high fixed costs that have been paid in the past by the incumbents. This assumption simplifies the analysis, allowing the isolation of integration effects through intensity of competition. Section 3.6 briefly considers the possible implications of relaxing the no-entry assumption.

Imagine that the first firm is located in the first of the two countries and that the other is located in the second country. As in Rivera-Batiz and Romer (1991a), consider only similar countries, so avoiding any question of comparative advantage.[72] With this approach, the existence of these countries is not explicitly modelled, and there is no explicit modelling of international trade.[73] However, thinking of the market for the two goods as being made up of these two countries allows a significantly different interpretation of an otherwise fairly standard differentiated duopoly model. Integration is modelled as an increase in the elasticity of substitution in consumer preferences between the two goods, which is assumed to be exogenous.

[72] Section 3.4 relaxes this assumption to allow technological superiority in one of the countries. All else is symmetric.

Identical consumers across both countries have Cobb-Douglas preferences over a composite of the R&D goods and a non-R&D competitive good. They therefore spend m, a fixed proportion of their income, on the two goods, i and j, which they do according to the constant elasticity of substitution (CES) utility function given by equation (3.4).

(3.4): $U(q_k) = \{q_{ik}^{(\theta-1)/\theta} + q_{jk}^{(\theta-1)/\theta}\}^{\theta/(\theta-1)}$, for all $\theta>1$, for consumer k.

The parameter θ is the elasticity of substitution in preferences between the two goods. Consumers divide m, their lump of income, between the two goods, giving us a representative consumer's demand for each good, as given by equation (3.5).

(3.5): $q_{ik}(p,m) = mp_i^{-\theta}/(p_i^{1-\theta} + p_j^{1-\theta})$, and similarly for good 2.

The demand curves (3.6) facing each firm can be found by summing over all consumers, where 'a' denotes the aggregate of the m's. As pointed out in Box 3.2, the resulting non-linear specification of the demand curve does not share the shrinking market effect of Qui (1994) and Bester and Petrakis (1993) and is therefore more suitable for a study of economic integration. Increasing the elasticity of substitution makes market share more sensitive to price but does not reduce market size.

(3.6): $q_i = ap_i^{-\theta}/(p_i^{1-\theta} + p_j^{1-\theta})$.

Costs of production are linear and given by:

(3.7): $C_i(q_i) = c_iq_i$.

Firms may invest in deterministic R&D, carried out 'in-house', which results in process innovation, reducing unit costs by $f(X_i)$, where X_i denotes the effective level

[73] All that is required is some way of differentiating between consumers located close to the producer of good 1 and consumers located close to the producer of good 2, whether geographically, or in some other sense.

of research, given by (3.8), with x_i being firm i's investment in research and β being the spillover parameter between firms, $0 \leq \beta \leq 1$, assumed to be exogenous.

(3.8): $X_i = x_i + \beta x_j$,

This set up for R&D is an extension of that of D'Aspremont and Jacquemin (1988) and Kamien et al (1992). Following these earlier papers, I assume $f(X_i)$ to be concave, and more precisely to be defined as:

(3.9): $f(X_i) = c_i X_i^{1/2}$.

Therefore R&D displays diminishing returns. D'Aspremont and Jacquemin (1988) cite Dasgupta (1986) as evidence in support of such a concave R&D function. The extent to which costs can be reduced through R&D is also related to the level of initial costs. In other words, it gets more difficult, in absolute terms, to reduce unit costs through time. This assumption is necessary in order to ensure a constant growth rate in the extension to the model presented in Section 3.5.[74]

Firms play a two-stage game. First they determine R&D expenditure, x, which determines unit costs for stage 2 in which there is Cournot competition.[75] Three alternative scenarios are considered. First, firms have identical initial production costs and compete in both stages of the game. This case is worked through in Section 3.3.1 below. An interesting extension of this case is presented in Section 3.3.3, where spillovers and integration are related. Secondly, firms with identical unit costs compete in the product market but can co-operate in the R&D stage. This is shown in Section 3.3.2. Thirdly, I consider the case where one firm has an initial cost advantage, due, for example, to a higher level of innovation historically. The first two scenarios can be solved analytically. The third is solved numerically in Section 3.4. In all cases, the model is treated simply as a one-shot game.

[74] It is not necessary to obtain the partial equilibrium results, however. An earlier version of this Chapter assumes $f(X_i) = X_i^{1/2}$ and the same pattern of results is obtained.
[75] I am implicitly assuming there is no difficulty in obtaining funding to carry out R&D from retained past profits. Section 3.6 considers the likely implications of replacing Cournot by Bertrand competition.

3.3.1: R&D Competition with Identical Unit Costs

Equation (3.9) assumes both firms have the same R&D opportunities, being in the same industry, and that they face the same R&D function, which follows from the assumption of similar countries. These symmetry assumptions lead us naturally to a symmetric equilibrium.

Consider the stage 2 output decision In the second stage, firm i finds its profit maximizing level of output, for a given level of R&D expenditure in stage 1. Total costs are given by (3.10). Multiplying both sides of (3.6) by p_i, multiplying top and bottom of the RHS by $p_i^{1-\theta}$, and then substituting the expression for relative prices obtained by dividing (3.6) by the equivalent expression for firm j, gives us total revenue for firm i, (3.11). Hence profits are given by (3.11) – (3.10). The subscripts are not dropped from unit costs at this stage for consistency with the asymmetric case outlined in Section 3.4.

(3.10): $TC_i = (c_i - f(X_i))q_i + x_i$.

(3.11): $TR_i = p_i q_i = a / \{(q_j/q_i)^{(\theta-1)/\theta} + 1\}$.

The fist order condition with respect to q_i is given by (3.12), where $\phi = (\theta-1)/\theta$.

(3.12): $c_i - f(X_i) = [a\phi(q_j/q_i)^{\phi}] / [q_i\{(q_j/q_i)^{\phi} + 1\}^2]$.

The Cournot model displays the standard downward sloping reaction curves, so the two goods are strategic substitutes. Manipulation of equation (3.12) and the mirror equation for firm j, gives us:

(3.13): $q_j = q_i [c_i - f(X_i)] / [c_j - f(X_j)]$.

So, relative outputs just depend on initial costs and R&D expenditures. If a firm can adopt a lower cost technology than its competitor, it can increase its market share at the expense of the higher cost firm. Matters are further simplified here because initial unit costs are assumed to be equal, so that output shares just depend on differences in R&D expenditures. This is where the R&D incentive comes from in the model. The incentive is purely strategic. No other R&D incentive exists.[76]

Substituting (3.13) into (3.12) gives us the firm's profit maximizing level of output, as shown in (3.14). Let F(X) denote $[c_i - f(X_i)] / [c_j - f(X_j)]$.

(3.14): $q_i = [a\phi F(X)^{\phi}] / [(c_i - f(X_i)\{F(X)^{\phi} + 1\}^2]$

Substituting (3.14) into the expression for profits (3.11)–(3.10), gives profits in terms of R&D expenditure:

(3.15): $\Pi_i = [a + (a/\theta)F(X)^{\phi}] / [\{F(X)^{\phi} + 1\}^2] - x_i$.

In Stage 1, the firm sets the profit maximizing level of R&D expenditure, x_i. Notice that $\partial X_i/\partial x_i = 1$ and that $\partial X_j/\partial x_i = \beta$. Given that the model has been set up with symmetry maintained throughout, there is nothing to make one firm's R&D decision any different to the other's, so $x_i = x_j$. The first order conditions give us the following expression for optimal effective R&D:

(3.16): $X_i^{1/2} = 1/2 \pm [\sqrt{\ } \{1 - a(1-\beta)(\theta-1)/2\theta\}]/2$

Second order conditions rule out the larger of these two solutions. So, the smaller of the two solutions denotes profit-maximizing effective R&D level.[77]

[76] Because of the cost of R&D, in symmetric equilibrium, profits are decreasing in R&D expenditure. If there was no competition and no threat of entry, the firm wouldn't bother at all.
[77] In order to ensure a non-negative expression in the square root, the parameters have to obey the following condition:
(3.16a): $1 \geq a(1-\beta)(\theta-1)/2\theta$.
Rearranging gives us a condition linking the demand parameter, a, with the elasticity of substitution, θ, as shown in (3.16b):
(3.16b): $\theta \leq a/(a-2)$.
The condition is binding when $\theta=a/(a-2)$. In other words, for the model to be valid, $a \geq 2$. When a is close to 2, θ can be large. When a is large, θ must approach 1.

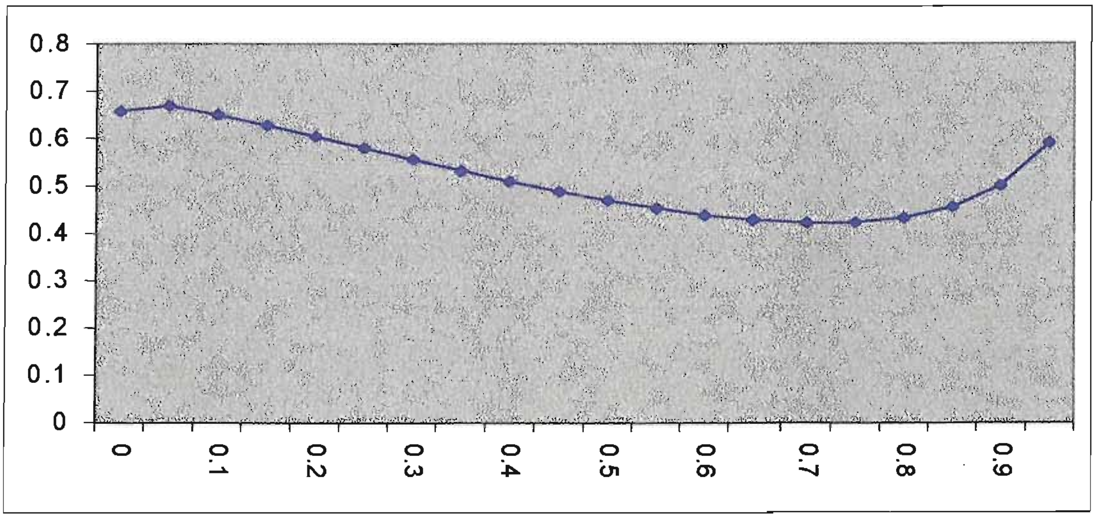
Figure 3.1 shows firm i's profit function for a given level of firm j R&D expenditure. The local maximum and minimum are clear. Note that profits tend to infinity as R&D expenditure tends towards 1. However, this extreme scenario can be ruled out by assuming there is some lower bound on post-R&D costs at all times: $x_i<0.25$, for example.

From (3.16) it can be seen that:

(3.17): $\partial(X_i^{1/2})/\partial\theta = a(1-\beta)/\{8\theta^2\sqrt{[1 - a(1-\beta)(\theta-1)/2\theta]}\}$,

which is positive for all $\beta<1$. So, integration, as measured by an increase in the substitutability parameter, θ, leads monotonically to a higher level of effective R&D for any given level of spillovers. Therefore firms' expenditure in R&D is increasing with integration.

Figure 3.1: Firm i's Profit Function.



Notes: The above function is Equation (3.15) for parameter values $a=2$, $\theta=2$, $\beta=0$ and $x_j=.1$.

R&D increases with integration at a decreasing rate the more integrated the industry becomes. There is no shrinking market effect, so integration has no negative effect at low levels of integration, such as that found by Qui (1994). As expected, a firm closely integrated with its competitor loses a big slice of its market share if its price is undercut. Equally, it can steal a big slice of its competitor's market share if it can undercut its price. These effects are much more muted when substitutability is low.

Although R&D is expensive, and although cost reduction is of no benefit to the firm when matched by its competitor, these strategic incentives to carry out R&D result in a situation where both firms invest more and more the more integrated the industry becomes.[78]

What of the effect of spillovers on R&D? From (3.16) it can be seen that:

(3.18): $\partial(X_i^{1/2})/\partial\beta = -a(\theta-1)/\{8\theta^2\sqrt{}[1 - a(1-\beta)(\theta-1)/2]\}$.

Given $\theta>1$, the expression in (3.18) is negative. In other words, effective R&D and firm's own R&D are decreasing with spillovers. This is a common result for this kind of model (see, for example, D'Aspremont & Jacquemin, 1988). For higher spillovers, investment in R&D lowers the competitor's costs and this increases the competitive pressure on the firm. Also for high spillovers the firm may be tempted to free ride on its rival's R&D expenditure. Both these factors act to lower R&D expenditure. These factors outweigh the positive spillover effect on effective R&D from (3.8) at all parameter values. In the extreme, with perfect spillovers, there is no R&D carried out at all, since no strategic advantage can be gained.

The negative spillover/R&D relationship is stronger the higher the value of θ. In other words, the more intense is the competition between the firms, the stronger the disincentive effect of spillovers on R&D expenditure. As discussed in Section 3.2, it is more of a disincentive to lower a direct competitor's costs than a distant competitor's costs.

3.3.2: R&D Co-operation with Identical Unit Costs

Consider the co-operative research scenario, where firms set R&D expenditure to maximise total industry profits in the first stage, but still compete in the product market in the second stage. This corresponds to D'Aspremont and Jacquemin's second case (D'Aspremont & Jacquemin, 1988) and case C of Kamien et al (1992). This is

[78] The Single European Market (SEM), if modelled in this way, should lead to higher technology levels for European firms and therefore better global competitiveness.

probably the most realistic scenario of co-operative research since antitrust legislation does not generally encourage product market co-operation but, certainly in Europe, does allow for co-operative R&D in many cases (see Jacquemin, 1988, for an informative discussion of EC Antitrust Law).

Profit maximizing output is given by (3.14), since firms still compete in the second stage. In (3.15), firms now maximise $\Pi_{total} = \Pi_1 + \Pi_2$, with respect to research expenditure, x_i. Solving this gives us the corner solution of $\partial\Pi_{total}/\partial x_i = -1$. In contrast to the competitive scenario, there is no longer any strategic incentive to invest in R&D so firms do not conduct any R&D. Firms prefer to co-operate and hold back innovation, since they save themselves wasted R&D costs.[79] For a given level of spillovers, integration increases the profit-gap between the co-operative R&D and competitive R&D scenarios, and therefore increases the incentives to co-operate in R&D.

Katz (1985) makes a similar prediction for firms producing a homogenous good under Bertrand competition. The critical assumptions behind the above result are that there is no entry and that the industry is not competing for demand with other industries that may be advancing technologically. Section 3.6 discusses the likely effects of relaxing the entry assumption in the competitive case. The fact that firms favour co-operation over competition will not change in this case, although we would expect some non-zero level of R&D expenditure.

3.3.3: Integration and Spillovers

Section 3.2 introduced the possibility of a positive relationship between integration and spillovers. Such a possibility is discussed in depth in Section 2.6 and in Chapter 5 of this thesis. For example, Grossman and Helpman (1991a) argue that increased trade between countries will involve increased interaction between agents in different countries and therefore more opportunity for knowledge to pass from agent to agent.

[79] There will, however, be an incentive for firms to cheat on the agreement and lower unit costs unilaterally in order to capture a bigger market share. This may make such agreements unstable in the absence of credible punishment strategies.

Equally, increased trade might lead to more opportunities to reverse engineer technical products. Empirical evidence for such a relationship is strong (see Chapter 5).

In the current model, this spillovers/integration relationship can be captured by a simple function, such as (3.19), where $0 \leq g(\theta) \leq 1$.

(3.19): $\beta = g(\theta)$,

with $g'(\theta)>0$, $g''(\theta)<0$ and $g(\theta) \to 0$ as $\theta \to 1$.

In other words, a reduction in product differentiation makes it easier for a firm to apply any process innovation by its competitor to its own production process.

Substituting (3.19) into (3.16) and differentiating with respect to θ gives us the following expression for the effect on R&D of integration:

(3.20): $\partial(X_i^{1/2})/\partial\theta = \{a/8\sqrt{\ } [1 - a(1\text{-}g(\theta))(\theta\text{-}1)/2\theta]\}.\{(1\text{-}g(\theta))/\theta^2 - g'(\theta)(\theta\text{-}1)/\theta\}$.

The first term is positive, as is the first term in the right-hand bracket. The second term in the right-hand bracket is negative which gives us an ambiguous sign for the whole expression. Sketching the relationship shows an inverted U-shaped relationship between integration and effective R&D. With small θ (ie: a high degree of product differentiation), integration leads to a higher level of R&D expenditure. With large θ (ie: a low degree of product differentiation), further integration leads to a lower level of R&D expenditure.

There is therefore an optimal level of integration at which the negative spillover effect and positive competition effects are balanced.[80] This is given by the stationary point of the expression in (3.20), as shown in (3.21).

(3.21): $\theta^* g'(\theta^*)(1\text{-}\theta^*) - g(\theta^*) + 1 = 0$.

Imposing a simple form on function g that meets the conditions outlined in (3.19) can give us an illustrative optimum integration level. Take the example of $g(\theta) = 1-1/\theta$. In this case (3.21) solves for an optimum integration level of $\theta^*=2$. Of course, this is all somewhat stylized. Nonetheless, the suggestion that the interaction of a negative spillover effect on R&D incentives and a positive strategic effect on incentives might give rise to such a non-linear integration/R&D relationship purely through intensity of competition effects is potentially very interesting. If the prediction is robust to generalization of the model in further research, then there may be significant implications for trade policy.

3.4: R&D with Asymmetric Initial Unit Costs

The results so far have assumed all is symmetric in the model. This means that firms cannot start with a cost advantage, or build such an advantage by their R&D efforts. Also, all the potential benefit in terms of profits from R&D is lost when any cost reduction is matched exactly by the competition. Now consider relaxing the symmetry assumption to allow one firm to have an initial cost advantage, due to, say, historical innovation in cost-reducing processes. Are the relationships between integration, spillovers and R&D robust to this change? I consider only R&D competition in this case, leaving aside any question of co-operation. The model is also only considered explicitly as a one-shot game.

The model is exactly the same as that set out in Section 3.3 except that initial unit costs are not the same for each firm. This means a crucial simplifying assumption is lost at the profit maximizing stage, and the system of non-linear equations, given by maximising (3.15), can no longer be solved analytically. However, solving the model numerically is relatively straightforward.[81]

Results are presented below, in Tables 3.1 and 3.2, for the case where firm 1 has initial unit costs equal to 1.005 and firm 2 has initial unit costs equal to 0.995. In other words, firm 2 has a slight initial cost advantage. The other parameters are set as a=2,

[80] Optimal in a growth maximizing sense.
[81] In this case, Excel is used to obtain approximate numerical solutions.

β=0, 0.1 and 0.5 and θ=2, 5 and 10. These are chosen as plausible values for zero, low and high spillovers and low, medium and high substitutability. Three points in each range is sufficient to see the pattern of solutions emerge. As long as the parameter values chosen fall within the range determined by the condition $\theta \leq a/(a-2)$, the pattern of the numerical results is consistent with alternative sets of parameter values outside these ranges.

In the asymmetric case, the high cost firm engages in significantly more R&D than the low cost firm, for all permitted parameter values. In other words, the initial cost gap gets endogenously smaller through profit maximising R&D. In fact, the higher R&D expenditure of the high cost firm is always sufficient to equalise post-innovation unit costs, as shown in Table 3.2. The symmetric equilibrium presented in Section 3.3.1 is therefore stable, so that a shock to one of the firm's cost levels is absorbed by the model and the symmetric equilibrium reverted to in a single period. This result is driven by the fact that R&D expenditure is more effective in reducing costs for the high cost firm (from Equation (3.9)) until technology levels are exactly equalised.[82] Since the benefits to R&D are largely competed away, the low cost firm prefers to save on R&D costs when it can.

Table 3.1: R&D Expenditure with Asymmetric Initial Unit Costs

θ,β	**0**	**.1**	**.5**
2	.027, .018	.021, .016	.010, .005
5	.080, .075	.060, .055	.028, .020
10	.094, .090	.090, .085	.034, .025

Notes: Values of θ=2,5 and 10 represent low, medium and high elasticity of substitutability respectively, and therefore low, medium and high levels of integration. Values of β=0, .1 and .5 represent zero, low and high spillovers. Figures for firm 1 (the high cost firm) are on the left of each cell, with figures for firm 2 on the right of each cell.

Table 3.2: Stage 2 Unit Costs with Asymmetric Initial Unit Costs

θ,β	0	.1	.5
2	.85, .85	.86, .86	.89, .89
5	.72, .72	.75, .75	.81, .81
10	.70, .70	.70, .70	.79, .79

See notes for Table 3.1.

Bester and Petrakis (1993) find the opposite result for firms with asymmetric costs, ie: that the low cost firm invests more in R&D than the high cost firm, making the cost gap endogenously wider. This is likely to be the result of the fact that the marginal returns to investment in R&D in the Bester and Petrakis model are increasing over time, as cost reduction is absolute rather than proportional to initial unit costs as in the current model. In other words, the marginal return to R&D is higher for the low cost firm. An alternative version of the current model, with a different specification for Equation (3.9), where cost reduction does not depend on the level of unit costs, but where all else is unchanged, leads to the same prediction as the Bester and Petrakis model (see McVicar, 1998).[83] Despite obtaining the opposite prediction concerning the evolution of the technology gap using the two alternative specifications for (3.9), the effects of increasing integration and increasing spillovers are the same in both cases. The current specification for Equation (3.9) is preferred to the alternative for the stability property of the symmetric equilibrium and to give a constant growth rate in the extension to the model outlined in the following section.

How does R&D expenditure relate to the degree of integration and the level of spillovers? Firstly, R&D expenditure is everywhere decreasing with spillovers, as in the symmetric case. This is unsurprising, since the intuition applies equally to both scenarios. More interestingly, the ratio of high cost to low cost firm R&D expenditure is higher when spillovers are larger, in all cases. In other words, the higher the level of spillovers, the more benefit of the low cost firm's R&D expenditure acts to narrow the technology gap, so the less inclined to invest in R&D it is. The relationship between integration and R&D expenditure is similar in the asymmetric case to that in the

[82] The single-period reversion to symmetric equilibrium is a reflection of the one-shot nature of the model and the small size of the initial cost difference. Making the game dynamic, with a larger cost difference and a limited step-size for cost reduction in each period would slow the process down.

symmetric case. In other words, R&D increases at a decreasing rate with integration for both firms. The negative spillover incentives and positive intensity of competition incentives are therefore robust to alternative specifications for initial technology levels.

3.5: A Symmetric N-Sector Extension

A simple extension to the model of Section 3.3.1 allows the economy-wide growth rate at a given point in time to be calculated. This is intended to be illustrative of the potential of the model, rather than being a serious contender for a model of endogenous growth and integration in its own right. This is left for further research. The predictions of the partial equilibrium model carry over to the simple extended model. This has significant implications for the growth/integration literature. Foremost of these is that it is possible for integration to increase the growth rate in a simple R&D-based endogenous growth model, without relying on scale effects, but purely through the effects of the intensity of competition on R&D incentives.

Assume there are a fixed number, N, of sectors, all differentiated duopolies, with one firm in each sector in each country as before. The goods are differentiated both between and within sectors. Cobb-Douglas preferences across sectors ensure a constant proportion of consumer's income is spent on the goods from each sector. Two assumptions allow us to focus on a single representative time period and to apply directly the results of Section 3.3.1. Firstly, discounted utility and profit functions are additively separable for each period. Secondly, the countries, sectors and firms are all symmetric and all are integrated to the same degree at any one time. Relaxation of this additional symmetry assumption is left for further research. By making this assumption, we impose a common θ.

Each individual sector is set up as before. A labour endowment, L, is introduced for each country, being the only factor used in production. Innovation can now be thought of as labour saving. Assuming equal sized sectors, with market clearing wages, each firm uses L/N units of labour. The unit cost of production can now be thought of as the

[83] Cost reduction is defined as $f(X_i) = X_i^{1/2}$ in this alternative specification.

amount of labour needed to produce a single unit of output multiplied by the wage, giving total costs as in (3.22).

(3.22): $TC_i = q_i(c_i\text{-}f(X_i))w_i - x_i$,

where notation carries through from Section 3.3 and w_i is the wage paid to workers in firm i. .

Total revenue for firm i is given, as before, by equation (3.11), where firms i and j are two firms in a given sector. Profit maximizing output in Stage 2, for given R&D, is therefore given by the marginal condition (3.23).

(3.23): $q_i = [a\phi F(X_i)^{\phi}] / [w_i(c_i\text{–}f(X_i))\{F(X_i)^{\phi}+1\}^2]$.

Note that $\phi=(\theta\text{-}1)/\theta$, as before. The only difference in notation is that $F(X_i)$ now includes a term for relative wages, ie: $F(X_i) = w_i(c_i\text{-}f(X_i))/w_j(c_j\text{-}f(X_j))$.

Country output is given by Nq_i. Given the (unchanged) symmetric equilibrium R&D expenditure from equation (3.16), and given all else is symmetric, $w_i=w_j$ and $F(X_i)=1$. Therefore (3.23) simplifies to:

(3.24): $q = a\phi/4w(c\text{-}f(X))$.

Firm output is increasing in the demand parameter and the substitutability parameter, and decreasing with the wage and unit costs. The labour demand equation gives us a second expression for q:

(3.25): $q = L/N(c\text{-}f(X))$.

Equations (3.24) and (3.25) solve for the symmetric equilibrium wage, given by (3.26):

(3.26): $w = Na\phi/4L$.

Equilibrium wages are increasing in the demand parameter, the elasticity of substitution and the total number of sectors, N and decreasing with the size of the labour force. Substituting (3.26) back into (3.24) gives country output, Y:

(3.27): $Y = L/(c\text{-}f(X))$.

Costs are reduced in period t, consisting of the research and production stages, through innovation, by f(X). If initial costs at the start of period t are labelled as c then starting costs in period t+1 are c-f(X). So, from (3.27):

(3.28): $Y_t = L/c$ and $Y_{t+1} = L/(c\text{-}f(X))$.

The growth rate of country output, for given t, is therefore given by (3.29).

(3.29): $(Y_{t+1} - Y_t)/Y_t = \{[L/(c\text{-}f(X))] - [L/c]\} / [L/c]$.

Equation (3.29) simplifies to the following per capita growth rate:

(3.30): $g = [c/(c\text{-}f(X)] - 1$.

Therefore, unsurprisingly, growth is increasing in R&D. It is a small step to substitute the expression for f(X) from (3.16) into (3.30) to give per capita growth as a function of the parameters of the model:

(3.31): $g = 2/[1 + \sqrt{\ } \{1 - a(1\text{-}\beta)(\theta\text{-}1)/2\theta\}] - 1$.

So, per capita growth is a function $g(a,\beta,\theta)$ with $g'(a)$ and $g'(\theta) > 0$, and $g'(\beta) < 0$.

Growth is driven purely by the endogenous R&D decisions of profit maximising firms. The growth rate is positively related to the degree of integration, solely because of the intensity of competition effect. Although there are scale effects in the model, given by the positive relationship between g and a, this integration/growth result is not related to scale, as the parameter a is assumed constant. In other words, the model shows how the intensity of competition effect of economic integration can be used to obtain a non-scale endogenous growth effect of integration. The intuition for this result and the other results given by (3.31) is the same as that discussed in Section 3.3.

Holding all other parameters constant, the productive capability of the whole economy grows over time. In fact, the economy displays a constant growth rate, given by:

(3.32): $g = [2/(1+\eta)] - 1$,

where $\eta \leq 1$.

Of course, the model presented above is very stylized, being built on a large number of fairly extreme assumptions. In the following section, I discuss the likely effects of relaxing some of these assumptions in the partial equilibrium model of 3.3. The additional symmetry assumption imposed on the extended model is unlikely to be critical, at least not with Cobb Douglas preferences between the sectors.[84] Most importantly, however, the model needs to be extended to an explicitly dynamic form, where the additive separability assumption can be relaxed, and asymmetric technology levels can be accommodated. The results above describe a *snapshot* of time, allowing us to carry out comparative statics to deduce the effect of integration on technological progress, but should not be interpreted as a growth model proper. Nonetheless, the model should prove very useful as a foundation for further research.

3.6: Discussion

The model in its various guises makes a number of key predictions relating to R&D, trade and growth. First, effective R&D is increasing with integration and decreasing with spillovers. Second, co-operation in research will hold back the pace of technological progress by removing strategic incentives. Third, there is an optimal level of integration when spillovers are higher the more integrated the economies

[84] Aghion and Howitt (1998) present an n-sector model where some sectors and symmetric and some asymmetric. Increased competition in this model encourages innovation where firms are neck and neck, but discourages innovation in leader/follower sectors. Overall, the effect is ambiguous.

become. Fourth, a firm with a technological disadvantage will invest more in R&D than its rivals, thus bridging the technology gap. Finally, by extending the model to an N-sector two-country model, it is possible to derive a constant growth rate, which can be raised by integration and lowered by spillovers.

The model is built on a number of strong assumptions. Although generalising the model is left for further research, I consider the possible effects of relaxing two of these assumptions here, as a brief thought experiment. Appendix A already shows that the predictions of the model can be reproduced using falling tariffs as an alternative measure of integration. An earlier version of the paper (see McVicar, 1998) uses a slightly different specification for the returns to R&D which has implications for the evolution of the cost gap in the asymmetric case, but otherwise leaves the central predictions of the model unchanged. However, what happens if the assumption of no entry is relaxed? Secondly, what if firms compete on prices instead of outputs?

Given the no entry assumption, everything is driven by the strategic interaction of the two incumbent firms. However, profits are positive in the model, so there is an incentive for new firms to enter the industry until profits are driven down to zero. A number of alternative scenarios exist depending on whether entrants produce one of the two existing differentiated goods, or new varieties, and at what technology level they might enter.

Assume the incumbent has a technological advantage (eg: patents on past innovations), so entrants can only enter at a higher initial unit cost. If new firms produce the existing varieties in the sector, but at a higher unit cost, then the picture might look something like that in the asymmetric scenario in Section 3.4, as long as the cost gap is not so wide that it cannot be bridged.[85] If the predictions of the static model hold true over time then the duopoly will not be sustainable and new firms will enter until profits are driven down to zero. Introducing uncertainty to the R&D process and explicit modelling of R&D financing might be a way to protect the position of incumbents in such a case.

[85] There has to be some limit on the size of the cost reduction in the model, as suggested in Section 3.3.1.

If entrants can produce new differentiated varieties of the good, then entry has the effect of raising the degree of substitutability between varieties, assuming the total market size remains constant as in the Cobb Douglas case. Regardless of entrant technology level, this is likely to increase the incentives of both entrants and incumbents to invest in R&D. In this case, however, an alternative modelling strategy for integration is likely to be needed.

Secondly, consider Bertrand competition instead of Cournot competition. This comparison is the subject of Qui (1994) in a similar model, although with linear demand. Qui finds a monotonic negative relationship between competition and R&D investment in the Bertrand case. The intuition behind this result is that price reduction by one firm forces its rival to also reduce prices, which reduces the profits of the innovating firm. This negative strategic effect is stronger the closer the firms compete, or the more integrated the market. In contrast, innovation in the Cournot case increases output, which reduces competitor's output, thereby benefiting the innovating firm. In other words, Qui (1994) shows how the strategic effects of integration under Bertrand and Cournot competition might act in opposite directions. Bester and Petrakis (1993) also show this contrast in the strategic effects between price and output competition. These contrasting strategic effects, if they are present in the non-linear demand structure of the current model, would have significant implications for the integration/growth relationship of the previous sections.

3.7: Summary and Concluding Remarks

This Chapter considers the effects of economic integration on a firm's incentives to innovate starting from the viewpoint of the IO literature concerning innovation and market structure. A model is set up of an industry consisting of two competing firms producing differentiated goods, with deterministic process innovation and knowledge spillovers. Increased competition resulting from integration makes a firm's market share more sensitive to prices. A symmetric equilibrium results where R&D is increasing with integration. This is because of intensified strategic incentives to both steal the competitor's market share and to protect own market share. In addition, two special cases of the model are considered. If firms are allowed to co-operate in R&D,

the strategic incentives are removed and R&D stops. Secondly, if spillovers are positively related to the degree of integration, then an inverted U-Shaped relationship between integration and R&D is the result, with an optimal level of integration. Finally, the model is extended to an n-sector two-country economy that displays a constant growth rate. This growth rate increases with integration.

Existing models explaining possible mechanisms for a positive link between growth rates and integration are generally driven by improved R&D performance resulting from the scale effects of increasing market size and access to a wider stock of technical knowledge (eg: Rivera-Batiz and Romer, 1991a, 1991b). This Chapter sets up a model where increased strategic incentives to innovate provide an alternative mechanism to these scale effects through which integration can stimulate R&D.

The model is consistent with the weight of the empirical evidence in the IO literature suggesting a positive relationship between competition and innovation.[86] By adopting a simple extension, the model can also be made consistent with the kind of non-linear relationship found by Scherer (1967). In Section 2.4, I review the empirical literature on integration and growth, which broadly supports a positive relationship. The model is consistent with the consensus of this literature. Chapter 4 of this thesis looks more directly for evidence of a relationship between integration and R&D expenditure at sectoral level in a sample of OECD countries. The broad conclusion is that where such a relationship is significant, it is usually positive, again consistent with the predictions of the simple model presented here. Of course, the intensity of competition effects described here may be small compared to the demand-pull and technology-push effects, which is a question best analysed empirically. However, the case for their existence is clear.

Of course, given the somewhat stylised nature of the model, the results should be treated with caution. This Chapter is intended as a starting point for further research, which can move towards generalising the model. Section 3.6 begins this process by briefly discussing the possible effects of dropping some of the key assumptions on which the model is built. Specifically, it would be interesting to see if the results hold

[86] See Section 2.5 for a review.

when the model is extended to an explicit dynamic framework on which a more solid growth model could be built. Clearly this line of research has potential for development.

Chapter 4

R&D Incentives and Openness to Trade

4.1: Introduction

In this Chapter I look for evidence of a relationship between the incentives to invest in R&D and the degree of openness to trade. Although a number of recent papers suggest such a relationship may exist, and that it is most likely to be a positive one, there is a noticeable lack of empirical evidence in support of this theory.[87] I argue that there is sufficient ambiguity as to the sign of any openness/R&D relationship to make such empirical evidence essential.

Using a time-series cross-section data set at sectoral level across 14 OECD countries, and using R&D expenditure to proxy for R&D incentives, I estimate a simple dynamic R&D equation with openness to trade as an explanatory variable. The nature of the data coupled with the dynamic specification of the R&D equation leads to a number of complications with estimation. Despite these complications, I find evidence of a positive relationship between R&D expenditure and openness to trade in a number of manufacturing sectors. One sector, however, displays a weak negative relationship.

The remainder of this Chapter is set out as follows: The following section discusses the possible links between openness and R&D incentives. Section 4.3 reviews some existing empirical evidence. Section 4.4 outlines a theoretical model linking openness to trade with R&D incentives (from Rivera-Batiz and Romer, 1991b). Section 4.5 outlines a more general empirical model that I go on to use to test the openness/R&D relationship. Section 4.6 describes the data and identifies some 'stylized facts'. Section 4.7 discusses the econometric considerations for the estimation of the empirical model. Sections 4.8 and 4.9 present and discuss the results and draw conclusions. Details not otherwise

[87] Existing literature on R&D incentives and openness to trade is reviewed in Sections 2.4 and 2.5.

contained in the text, including details of data construction and sources, some test results and some summary and descriptive statistics, are provided in Appendix B.

4.2: R&D and Openness

Modern growth theory recognises the importance of endogenous technological progress as the driving force behind long run economic growth (see Section 2.2.4 for a review). In these models growth typically arises because of decisions by profit maximizing agents to invest money in the research and development of new and better products or improved processes of production. Such decisions rest on a large number of different economic considerations.[88] In this Chapter, I concentrate on just one such consideration, namely the degree of openness to trade in the agent's industry.

The openness to trade of an industry may affect R&D in a variety of ways. Cameron et al (1998) identifies 5 channels through which trade might affect productivity growth. Four of these are relevant to the more narrowly defined question of how openness to trade might affect the incentives to invest in R&D. Broadly speaking, they are:

1. Trade may affect R&D expenditure through its effect on the spillover of ideas, or scientific knowledge, across countries. Although it is argued that such spillovers should increase growth, their effect on R&D incentives may work in the opposite direction. Therefore the sign of this effect is ambiguous.

2. Trade may affect R&D expenditure by reducing the incentives for duplication of research in different countries. Related to this, the extent of trade may affect the incentives to co-operate in research across national boundaries. Again, these effects could increase growth, but may in fact decrease R&D expenditure, implying an ambiguous sign.

[88] See, for example, Sections 2.4 and 2.5 for a discussion of some of the factors affecting R&D incentives.

3. Trade increases the size of the market available to successful researchers, which has an unambiguously positive effect on the incentives to invest in R&D.

4. The degree of openness to trade affects the intensity of competition in a given industry, which has ambiguous effects on R&D incentives.

Consider the spillover channel. If technological know-how or scientific knowledge is an input in the research process, then a firm's R&D output can be increased by increasing its level of scientific knowledge, other things being equal.[89] At a wider level, if there is such a relationship between technological progress and the level of underlying scientific knowledge in the economy, economic integration or increased trade might stimulate technological progress by allowing agents access to a greater stock of scientific knowledge. A number of different mechanisms by which this might occur have been suggested, such as increased opportunities for reverse engineering, increased communication between skilled human capital across countries and attendance at foreign conferences (see Chapter 3 for a discussion). The existence of such spillovers will lead to more technological progress from a given investment in R&D than would be the case if no such spillovers existed. This would provide a positive incentive effect for R&D expenditure and may lead to increased long run growth (see for example, Rivera-Batiz and Romer, 1991a; Grossman and Helpman, 1991a, 1991b).

However, the overall effects of increased spillovers on R&D expenditure may be more ambiguous. A number of papers show that there may be a negative relationship between spillovers and R&D incentives (see, for example, D'Aspremont and Jacquemin, 1988; Kamien et al, 1992). This negative incentive effect is likely to be stronger the more direct is the competition between firms. Because the extent of these spillovers may be positively related to the degree of openness to trade, increasing openness may therefore strengthen this spillover effect and reduce R&D incentives (see Chapter 3 of this thesis).

Secondly, it has been argued that increased trade will bring about a reduction in the level of duplication in research across countries (Rivera-Batiz and Romer, 1991b), through a firm's desire to be differentiated from its competitors. It may also be that integration fosters co-operation in research across countries, which may lead to more effective R&D and further reduce wasteful duplication.[90] Again, these effects will lead to greater technological progress from a given level of R&D expenditure, which should increase long run growth.

However, as with the spillover channel, the effect on R&D expenditure itself is not so clear, and may even lie in the opposite direction. It is not clear what effect a reduction in duplication would have on overall R&D expenditure. Although less R&D would be carried out in a particular area, the remaining funds may be diverted to R&D on alternative projects. The effect on R&D expenditure of increased co-operation in research is also somewhat ambiguous. There may be positive effects from the reduction of the market failure of incomplete appropriability (or increasing spillovers within the group and reducing spillovers to outside the group). High fixed costs can be shared. These effects may outweigh the negative effects of possible anti-competitive behaviour (Ordover and Willig, 1985). On the other hand, Katz (1985) argues that for a high degree of product market rivalry between firms, a co-operative research agreement is less likely to lead to increased research. Chapter 3 of this thesis argues that integration will therefore lead to less expenditure on R&D in co-operative research agreements. All this leaves us with a somewhat ambiguous effect of openness trade on research expenditure through the duplication/co-operation channel.

The third channel identified through which increased trade could affect R&D expenditure is through the market enlargement effect, also known as the demand-pull hypothesis (Schmookler, 1966). The larger the market that an innovator finds him or herself in, the greater the potential rewards in terms of profits from a given innovation. So, for a given R&D expenditure, you get a higher payoff, other things being equal. This will have an unambiguously positive effect on the incentives to invest in R&D (Rivera-Batiz and Romer, 1991a).

---

[89] This is the technology-push hypothesis discussed in Sections 2.4 and 2.6. It is discussed in the context of economic integration in Chapter 3.
[90] This possibility is discussed briefly in Section 2.5.2 of this thesis.

Finally, trade may affect R&D incentives through changes in the nature and intensity of product market competition. The direction of this effect is not universally agreed upon, and may depend on the particular circumstances under consideration or the particular assumptions made in a particular model. Indeed, models exist where increased competition can either increase or decrease the incentives to conduct R&D (see the discussion of Section 2.5.1). More intense competition may reduce profits from innovation, impacting negatively on R&D incentives (Aghion & Howitt, 1992). Equally, more intense competition leads to an increased need to protect own market share and opportunity to 'steal' competitors' market shares, which might increase R&D incentives (see Chapter 3 of this thesis).

Generally, the integration/R&D-driven growth literature hypothesises a positive relationship between openness to trade and R&D incentives, although Rivera-Batiz and Romer (1991b) and Grossman & Helpman (1991a) describe situations where a negative relationship might exist (or a non-linear relationship overall).[91]. These studies present models where both technology-push and demand-pull effects increase growth. They are equally applicable to the question of how openness to trade affects R&D incentives since their growth rates are simply and directly related to the level of R&D expenditure. These models, along with the consideration of the strategic incentive effects of integration presented in Chapter 3, provide us with theoretical hypotheses linking increased openness to increased R&D. It should be clear from the above discussion, however, that the overall relationship between openness and R&D is somewhat ambiguous. This ambiguity implies a need for empirical research to establish the true nature of the relationship.

The above discussion suggests there is an important distinction between trade and growth relationships and trade and R&D expenditure relationships. Some of the mechanisms through which trade might affect R&D incentives may work in opposite directions to the way trade may affect growth overall. More R&D is not always a good thing in terms of higher economic growth and less R&D does not always reduce growth rates. A full

[91] See Section 2.4 for a discussion of this literature. I concentrate on symmetric models in this Chapter, so abstracting from questions of dynamic comparative advantage, which may further muddy the waters. See

discussion of the intricacies of these inter-relationships is beyond the scope of this Chapter. It is the trade/R&D expenditure relationship that I study here. However, since the literature on this relationship and on the wider trade/growth relationship is in many cases inextricably linked, it is not always straightforward to disentangle the two in the discussion. Care needs to be taken in drawing conclusions from one field to another.

4.3: Existing Empirical Evidence

Although I have been unable to find existing empirical work that directly examines the relationship between openness to trade and R&D incentives in a similar way to the current study, a number of recent papers have examined related empirical issues, and we can draw some useful conclusions from them. The most important group of such papers is that which examines the trade and growth relationship and that which attempts to establish a relationship between openness to trade and productivity growth.

A large literature exists in which some form of cross-country growth regression is carried out (see, for example, Edwards, 1992; Barro and Sala-I-Martin, 1995; Lee, 1993; Henrekson et al, 1996; Frankel and Romer, 1999).[92] Lee (1993) and Barro and Sala-I-Martin (1995) find a negative relationship between growth and tariff rates. Some evidence is also found of a negative relationship when using black market premia for foreign exchange as a measure of trade distortions (Lee, 1993). Edwards (1991) reports similar results, although he employs a larger number of different (but closely related) measures of trade orientation. Henrekson et al (1996) finds evidence of a significant positive effect on growth of an EC/EFTA dummy for highly integrated economies, which seems to stand up to a fairly intensive sensitivity analysis.

Although all these papers suggest the presence of a significant positive growth/trade openness relationship, they tend to suffer from the inherent difficulties of measuring openness. Results are often not robust to alternative measures of openness (see Levine

Feenstra (1996) for an example of how different initial conditions in asymmetric countries can influence the sign of the trade/growth relationship.

and Renelt, 1992, for a discussion) and many of these alternative measures themselves are often uncorrelated with each other (Pritchett, 1996). Harrison (1996) argues that despite this problem, the evidence of a significant positive growth/openness relationship is still strong, and indeed is stronger when a panel data approach is adopted. Another difficulty with these studies is the possible endogeneity of some of the openness measures (Frankel and Romer, 1999). A positive relationship between trade volumes, say, and incomes or growth may be reflecting the fact that richer countries tend to trade more than poorer countries rather than the reverse causality we are looking for. To control for this, Frankel and Romer (1999) estimate the trade/growth relationship using geographical proximity measures to instrument for trade and still find a positive and significant relationship. In summary, even given the problems with these studies, the weight of evidence from this literature supports the existence of a significant positive relationship between trade openness and growth at a cross-country level.

At a more disaggregated level, and closer to the spirit of the research in this paper, Cameron et al (1998) looks for evidence of a relationship between productivity growth and openness to trade at sectoral level for the UK between 1970 and 1992. They use information from a number of alternative openness measures to categorise sectors as 'relatively open' or 'relatively closed', using a variety of different estimation techniques to try to solve the possible endogeneity problem. They find some evidence of a significant positive relationship between openness and productivity growth for a number of sectors, which they argue is more likely to be driven by incentives to innovate rather than reallocation of resources across sectors relating to comparative advantage. However, although this study goes some of the way to support the hypothesis that openness to trade may have a positive effect on R&D incentives, it is still essentially a growth/trade study and does not provide any *direct* evidence of such a relationship.

A growing literature exists on measuring the effects of knowledge spillovers, primarily across countries, on productivity growth.[93] Trade is generally taken as the primary mechanism by which knowledge spreads from one country to another in this literature

[92] Cross-country growth regressions are discussed in more detail in Section 2.3.
[93] Chapter 5 reviews this literature in some detail and presents a new study of spillovers at UK sectoral level.

(see, for example, Coe and Helpman, 1995; Coe et al, 1995; Engelbrecht, 1997). Similar studies have been carried out at a more disaggragated level across countries (see, for example, Branstetter, 1996; Keller, 1997). Although, once again, this literature does not directly address the question I am interested in here, it does provide some evidence of a link between openness to trade and productivity growth specifically through the technology-push channel of knowledge spillovers. In this sense it is more precise than the more general trade/growth literature. Lichtenberg and Pottelsberghe de la Potterie (1996) and Hejazi and Safarian (1998) extend this analysis to include foreign direct investment (FDI) as a possible transmission mechanism for technological knowledge. In as much as FDI is an alternative measure of openness or integration, this again supports the hypothesis that more open economies grow at a faster rate, at least through the spillover channel.

In short, there is substantial evidence for the existence of a positive relationship between openness to trade and growth, both in general and more specifically through the transmission of scientific knowledge. This evidence exists at cross-country level and also at a more disaggregated level. This evidence is consistent with the existence of a positive relationship between openness to trade and R&D incentives, a point which a number of the papers emphasise (see, for example, Henrekson et al, 1996; Cameron et al, 1998). However, because of the differences that may exist between productivity growth and R&D expenditure, this literature cannot be interpreted as firm evidence of such a relationship. In this chapter, I test *directly* for the existence of a relationship between R&D expenditure and openness to trade, using expenditure to proxy for incentives. This study therefore provides a valuable contribution to the empirical growth/trade literature and also informs the use of models of integration and R&D, which are becoming increasingly relevant in today's 'global economy'.

There are, of course, good reasons why such direct empirical evidence has not yet been presented, not least because these models are still quite new. Also, the empirical specification of R&D functions is still a somewhat *ad hoc* process (see Nickell & Nicolitsas, 1997 for a discussion). Lastly, although suitable data is now available for such a study, it is far from ideal for the purpose. This is on top of the more general unsatisfactory nature of R&D expenditure data (see Cohen and Levin, 1989 for a discussion). Nonetheless, the analysis goes some way towards establishing this empirical evidence, despite the difficulties inherent in such a project.

In what follows, I examine whether there is evidence for a positive relationship between R&D expenditure (incentives) and increased international trade in a panel of OECD countries. I look for evidence of an overall relationship and am not yet in a position to quantify the relative importance of the separate possible channels listed above in Section 4.2. This refinement is left for further research.

4.4: A Model of Openness to Trade and R&D

In Section 4.2 I outlined four channels through which openness to trade might affect R&D incentives. Unfortunately, I know of no model that captures all four of these channels simultaneously. Recent work has concentrated rather on providing models that can link openness to trade and R&D incentives (and hence to growth) through one or other of the channels individually (see, for example, Grossman and Helpman, 1990, 1991a, 1991b; Rivera-Batiz and Romer, 1991a, 1991b; Chapter 3 of this thesis).[94]

4.4.1: Scale Effects and Increasing Returns

The most straightforward way to think about the effects of increasing openness to trade on R&D incentives is through the concept of scale effects, and this is reflected in the theoretical literature. Scale effects basically tell us that the larger the economy, the higher the incentives to invest in R&D. Integration, or increased openness to trade, has this enlarging effect. The standard channels through which these scale effects work are the demand-pull (larger market therefore larger reward to successful innovator) and the technology-push (larger economy therefore larger stock of knowledge) channels.

[94] Or models that link openness to trade and growth and hence R&D incentives.

These scale effects can be introduced to the standard Romer (1990) model without difficulty. For example, Rivera-Batiz and Romer (1991a)'s 'knowledge driven' R&D model, in which knowledge creation (R&D) proceeds at a pace relating to the current knowledge level of the economy (stock of past designs etc.) describes the scale effect through the technology-push hypothesis. By setting up trade and communications links with foreign economies, integration allows researchers access to an international body of knowledge and therefore stimulates R&D. In the same paper, Rivera-Batiz and Romer present an alternative 'lab-equipment' model. This describes the demand-pull hypothesis in which the larger the market, the greater the potential reward to a fixed expenditure on R&D. By expanding the size of the market in which a firm operates, increasing openness to trade has the effect of stimulating R&D incentives once again. I describe one of these models (the lab-equipment model) below, but first it is useful to set down the concept of scale effects a little more formally. The discussion below follows that of Rivera-Batiz and Romer (1991b).

Consider two identical countries, which for the time being I treat as totally isolated from each other. Each of these countries has a two-sector economy, ie: the R&D sector which produces designs for capital goods, and the manufacturing sector which produces these goods on licence and also a final consumption good. Inputs to both sectors are human capital (H), labour (L), physical capital goods (K) and knowledge (A). All are rivalrous except knowledge, which is non-rival. We can write down a reduced form equation for output, Y, and for the output of new designs, $\partial A/\partial t$:

(4.1): $Y = C + \Delta K = F(H_Y, L_Y, K_Y, A).$

(4.2) $\partial A/\partial t = R(H_A, L_A, K_A, A).$

The key assumption is to allow for increasing returns in both Equations (4.1) and (4.2). These may be purely external, arising from knowledge spillovers. On the other hand, the basic inputs could be connected to the final sectoral output through a large set of specialised activities and intermediate inputs that are provided by monopolistically competitive firms.[95] Integration, in the first case, is achieved by an international flow of ideas, and in the second case by trade flows of intermediates between countries. In either case, if there are increasing returns in the sector, restrictions to the flow of ideas or the flow of intermediates will reduce the combined output of that sector.

Comparing the autarky situation with that of full integration, we see that the world output of new designs is given by $2R(\mathbf{Z_R})$ and $R(2\mathbf{Z_R})$ respectively, where $\mathbf{Z_R}$ denotes the vector of inputs to the research sector in autarky, and similarly for manufacturing output. This assumes there is no duplication in new designs in the two countries. With duplication, autarky world output of new designs could lie anywhere between R(Z) and 2R(Z), which would make the R&D effects of integration stronger. So, increasing returns in these reduced form equations is all that is needed for economic integration to stimulate R&D output. This is the basic idea behind the concept of scale effects.

But what of R&D incentives and partial integration? To analyse this some structure needs to be imposed on the model. This can be done in two steps, following the approach of Rivera-Batiz and Romer (1991b). First, holding the stock of inputs constant in the two different sectors of each economy, we can look at the effects on output of allowing an existing piece of specialised equipment previously only available in one country to become available in the other. Second, we can compare the rate of return to inputs in the two sectors after the partial integration. These will generally not then be equal, leading to a reallocation of inputs between the two sectors. The model presented is essentially the lab-equipment model of Rivera-Batiz and Romer (1991a).

4.4.2: The Lab-Equipment Model

Many types of capital goods may be used in production that are not perfect substitutes, and technological progress arises from the invention of new types of capital goods. Let i index the different types of capital good and x(i) be the usage of each good. Producers of consumption or capital goods face a function of the following form:

[95] See Section 2.2 for a detailed discussion of this point.

$$(4.3): \; Y = H^{\alpha}L^{\beta}\int_{0}^{A} x(i)^{1-\alpha-\beta}di.$$

Assume one unit of consumption is foregone for each extra unit of a capital good produced. Capital overall is given by:

$$(4.4): \; K = \int_{0}^{A} x(i)\, di.$$

The total output of the manufacturing sector is simply $Y=C+\Delta K$. Given A, symmetry allows us to write a common value for x(i), $x_c = K/A$. Substituting this into the production function above gives us:

$$(4.5) \; Y = (HA)^{\alpha}(LA)^{\beta}K^{1-\alpha-\beta}.$$

Whilst the underlying production function is homogeneous of degree 1, this reduced form expression is homogeneous of degree $1+\alpha+\beta$.

There are no external effects in the manufacturing sector, but increasing returns arise because of the imperfect competition. The producer of good i is the only producer of that good and charges the monopoly price for it. Because of competition between firms to be the producer of i, the price paid for the patent equals the present discounted value of the stream of monopoly profits from the production of good i.

Research, which is human capital intensive, is given by:

$$(4.6) \; \partial A/\partial t = \delta H_A A^{world}, \text{ for the home country, and}$$

(4.7) $\partial A^*/\partial t = \delta H_A^* A^{world}$, for the foreign country. A^{world} denotes the world stock of knowledge, or the total number of varieties of capital goods across both countries.

To close the model preferences are given by the standard Ramsey type:

$$(4.8)\quad U = \int_0^{\infty} e^{-\rho t}.(C^{1-\sigma}-1)/(1-\sigma)dt,$$

This implies the following relationship between interest rates and consumption growth:

$$(4.9)\quad r = \rho + \sigma(\Delta C/C).$$

Assume across the board symmetric tariffs, τ, on intermediates, the revenues of which are redistributed to consumers. With trade, production and usage of intermediates in a country may differ, so x(i) denotes the stock of i produced by home firms and x*(i*) denotes production of a different intermediate in the foreign country. d(i) denotes domestic usage of i and m(i*) denotes domestic usage of foreign good i*. We can therefore write:

$$(4.10)\qquad Y = H_Y^{\alpha} L^{\beta}\{\int_0^{A} d(i)^{1-\alpha-\beta}di + \int_0^{A^*} m(i^*)^{1-\alpha-\beta}di^*\},$$

and symmetrically for the foreign country.

Aggregate demand for each good x(i) is derived by solving the following maximisation problem:

$$(4.11)\quad \max_{d,m} Y(H_Y,L,d,m) - \int_{R+} P_d(i)d(i)\,di + \int_{R+} (1+\tau)_m{}^*(i^*)m(i^*)\,di^*,$$

where P_d is the home price of domestic inputs and P^*_m the price received by foreign producers for their exports. $P^*_m(1+\tau)$ is the price paid by domestic users of foreign imports. This gives:

$$(4.12)\quad d(i) = (1-\alpha-\beta)^{1/(\alpha+\beta)} H_Y^{\alpha/(\alpha+\beta)} L^{\beta/(\alpha+\beta)}\, P_d(i)^{-1/(\alpha+\beta)}$$

and $$m(i^*) = (1-\alpha-\beta)^{1/(\alpha+\beta)} H_Y^{\alpha/(\alpha+\beta)} L^{\beta/(\alpha+\beta)}\,(1+\tau)^{-1/(\alpha+\beta)}\, P_{m^*}(i)^{-1/(\alpha+\beta)}.$$

Foreign demand for $d^*(i^*)$ and $m^*(i)$ is derived in the same way. In equilibrium, the production $x(i)$ of a domestic input is equal to the sum of domestic usage and exports $d(i) + m^*(i)$. Given symmetry, home production of $x(i)$ is identical for all i and can therefore be written as:

$$(4.13)\quad x = d(P_d) + m^*(P_{m^*})$$

Because of tariffs, firms can price discriminate and output levels will be set in each market to maximise total profits. One unit of foregone output is needed to produce one unit of capital, so the flow opportunity cost of these units is $r(d+m^*)$. The instantaneous rate of profit earned by the holder of each patent is therefore:

$$(4.14)\quad \pi = \max_{d,m^*} P_d(d)d + P_{m^*}(m^*)m^* - r(d+m^*)$$

Equating marginal revenue to marginal cost in each market and substituting for prices from (4.12) gives $P_d = P_{m^*} = r/(1-\alpha-\beta)$, which we can label simply p. Because of the free flow of the final consumption good and free capital movements, interest rates and prices of durables are equal in both countries. Because of free entry into the durables sector, the discounted value of revenue minus variable costs equals design costs P_A:

$$(4.15)\quad P_A = \int_{R+} e^{-\int_t^{s'} r(s)\, ds}\, \pi(s')\, ds$$

Differentiating with respect to time gives the following:

$$(4.16)\quad r = \pi(t)/P_A + (dP_A/dt)/P_A\,.$$

For constant P_A this reduces to:

$$(4.17)\quad P_A = \pi(t)/r = (1/r)[p(d+m^*) - r(d+m^*)] = (p/r - 1)(d+m^*).$$

Substituting $p/r = 1\text{-}\alpha\text{-}\beta$ and using (4.12) and the equivalent condition for the foreign country, gives the following relationship between P_A, r, d and m*:

$$(4.18)\quad P_A = (\alpha+\beta)(d+m^*)p/r = ((\alpha+\beta)/r)(1-\alpha-\beta)H_Y^{\alpha}L^{\beta}[d^{1-\alpha-\beta}+(1+\tau)^{-1}m^{*\,1-\alpha-\beta}].$$

Equating the marginal product of labour in research ($\delta P_A A^{world}$) and in manufacturing ($\partial Y/\partial H$) gives the following relation between tariffs and interest rates:

$$(4.19)\quad \delta((\alpha+\beta)/r)(1-\alpha-\beta)H_Y^{\alpha}L^{\beta}A^{world}[d^{1-\alpha-\beta}+(1+\tau)^{-1}m^{*\,1-\alpha-\beta}]$$
$$= \alpha H_Y^{\alpha-1}L^{\beta}[Ad^{1-\alpha-\beta}+A^{*}m^{1-\alpha-\beta}].$$

Solving the model for H_A and r, from (4.9) and (4.19) allows us to write down the balanced growth rate as given in (4.20).

The growth effect of tariffs stems from the reallocation of human capital between the manufacturing and research sectors, and since H is not monotonic in tariffs, neither is the growth rate:

(4.20) $g = [2\delta H - \Lambda\rho f(\tau)]/[\Lambda\sigma f(\tau)+1]$,

where $f(\tau) = [1+(1+\tau)^{-(1-\alpha-\beta)/(\alpha+\beta)}]/[1+(1+\tau)^{-1/(\alpha+\beta)}]$, $\Lambda = \alpha/(\alpha+\beta)(1-\alpha-\beta)$ and τ is the (symmetric) tariff rate.

The highest point of this function is attained where tariffs are equal to zero, ie: free trade. Growth is then decreasing in tariffs until eventually a point is reached where growth gradually increases again as $\tau\rightarrow\infty$. Behind this relationship are two offsetting forces. Firstly, increasing tariffs causes a fall in imported intermediates which reduces the price of new designs and shifts human capital out of the R&D sector. Counteracting this, increasing tariffs also reduces exported intermediates and this shifts human capital out of manufacturing. However, for all reasonable levels of tariffs, the first effect is the stronger, and only when tariffs are very large does the second effect begin to dominate.

Rivera-Batiz and Romer's conclusion is that if two similar countries increase their tariffs on imports produced in a monopolistically competitive sector, the net effect will be to drive resources out of that sector. This will reduce the rate at which goods in these sectors are introduced because of the resulting reduction in rents to compensate the research sector. In short, in this model, trade restrictions reduce the return from R&D and thus the incentives to carry out R&D (until tariffs pass a critical point at unrealistically high levels). Before continuing, let us briefly remind ourselves of the simple intuition for this result, although the formal model takes some time to arrive at it. If R&D unit costs are fixed in some sense, trade liberalization will increase the size of the market open to a potential innovator, thus increasing the potential returns from that innovation, thus increasing the incentives to carry out research.

The results of the above model rest on a number of key assumptions. Firstly, the assumption that intermediate goods are used as inputs in the research sector is crucial. Aghion and Howitt (1998) present a similar extension to the Romer model, but where intermediates are not an input to research, and no such positive trade/R&D incentives relationship exists. It is more difficult to apply a similar criticism to the knowledge-driven model however, although the particular form of research technology specified can have important consequences for the model's predictions (Aghion and Howitt, 1998).

A more fundamental criticism of the Rivera-Batiz and Romer knowledge-driven model relates to its instability. The equilibrium is balanced on a knife-edge, and any deviation in initial conditions from perfect symmetry (such as different initial stocks of knowledge) will lead to wildly different outcomes (Devereux and Lapham, 1994). Although this is undoubtedly important in terms of the applicability of the model to the real world, moving away from the symmetry assumption takes us into the realm of comparative advantage and changing allocations across countries (or across sectors within countries in more disaggregated models). These matters are beyond the scope of the present Chapter, but readers who wish to can refer to Feenstra (1996), for example, which considers integration between initially different economies.

Chapter 3 of this thesis describes the strategic R&D incentives of protecting your own market share and 'stealing' competitors' market shares. Integration acts in this model to increase the intensity of competition between firms located in different countries, therefore increasing these strategic incentives (in all but a few special cases).[96]. This gives us an example of how integration can affect R&D incentives other than the standard scale effects outlined above. The model of Chapter 3 has the advantage of not resting on the symmetry assumption, and therefore not displaying the knife-edge characteristic of the Rivera-Batiz and Romer model.

[96] The relationship between competition and R&D is not universally accepted to be a positive one. Empirical studies on this question in the IO literature have found both positive (see for example, Horowitz, 1962; Mansfield, 1968) and negative (see for example, Williamson, 1965; Bozeman and Link, 1983) relationships with some studies suggesting non-linear relationships such as Scherer's (1967) inverted 'U'-shaped curve. Aghion and Howitt (1998) present a basic Schumpetarian model in which the R&D/intensity

Summing up at this stage, models exist that predict a positive relationship between openness to trade and R&D incentives (at least for reasonable levels of openness) through the channel of scale effects. It is also possible to present models that propose a positive relationship through intensity of competition effects, although this channel may also work in the opposite direction. Consideration of duplication, although not present in the models discussed above, would add to the strength of such a positive relationship. From this discussion, at the very least, we have reason to believe that there may be a significant relationship between openness to trade and R&D expenditure. The likely sign of this relationship is not clear, although most theory and related empirical evidence points to the greater likelihood of a positive relationship. Again, the clear need for an empirical study such as this is apparent.

4.5: An Empirical Model

Moving from theoretical models towards an estimable empirical model of R&D is not straightforward. Such theoretical models tend to be highly specialised, in that R&D is considered in relation to just one other variable of interest. Examples are openness to trade in the model outlined in the previous section, the degree and type of product market competition in the IO literature, or labour market conditions (see, for example, Grout, 1984; Piermartini and Ulph, 1996). While this allows us to concentrate on the matters that interest us most in theoretical models, if we are to estimate the real-world relationship between R&D expenditure and openness to trade, we need a well-specified R&D equation. In other words, we need to have a fairly good idea of what *all* the most important factors entering into an agent's R&D decision are, and include them in the empirical model. Omitting an important factor will result in a poorly specified model and unreliable results. The difficulty lies in identifying this group of most important factors and, of course, obtaining suitable data on each and every one.

The existence of a theoretical link between R&D and openness to trade has been established above, but there are undoubtedly a host of other factors, external to this

of competition relationship turns out to be negative. Ambiguity remains about this relationship both model, which may explain R&D expenditure. In the absence of an all-encompassing theoretical model, such factors must be chosen in an essentially *ad hoc* manner to build an empirical model. However, this process needn't be unguided. Previous empirical R&D studies have established a set of potential explanatory variables, based on theoretical justifications derived from other alternative R&D models, and, I imagine, data availability (see, for example, Lach and Schankerman, 1989; Hall, 1992; Himmelberg and Petersen, 1994; Nickell and Nicolitsas, 1997). The specification of the empirical model follows both from these past studies and my own particular data limitations.

4.5.1: Explanatory Variables

For reasonable levels of openness to trade, following the discussion of the last section, openness is expected to enter the empirical model (the R&D equation) with a positive coefficient. There is sufficient ambiguity in the hypothesized relationships to retain an open mind, however. I cannot distinguish between the various different factors (channels 1-4) that may be captured by this variable. Rather, the main aim is to test the existence of an *overall* relationship. In order to compare knowledge spillover effects, strategic and competition effects, and demand-led effects, a far more detailed structure and far more detailed data would be required. This is left for further research.

A number of previous studies have discussed the likelihood that certain measures of openness to trade (such as trade intensity) might be endogenous in R&D models, with the possibility of a causal relationship running from innovation to trade performance (see, for example, Fagerberg, 1994; Meliciani and Piermartini, 1998). The possible endogeneity of trade has also been discussed in the trade and growth literature (see, for example, Frankel and Romer, 1996; Cameron et al, 1998). However, the difference between R&D inputs (expenditure) and R&D output (innovation and growth) has already been noted. It is in models of R&D outputs and trade where the endogeneity of trade is likely to be more of a problem, and indeed where it has been identified in the literature. Intuitively, contemporaneous R&D expenditure is unlikely to affect contemporaneous trade intensity.

theoretically and empirically.

Previous studies have tended to include some measure of financial constraints facing firms in the R&D equation (see, for example, Nickell and Nicolitsas, 1997; Hall, 1992). Research is faced by a capital market failure because of the informational asymmetries involved. R&D outcomes being uncertain, a capital provider has no way of knowing whether a project will be successful or not. If the project is not successful, the capital provider has no way of knowing whether this is a result of poor performance by the researcher or whether the envisaged innovation is simply not possible, or research has been down a blind alley. In other words, there is a well known problem of moral hazard when it comes to financing R&D. Consequently, the majority of R&D expenditure is financed by firms internally, from retained profits. A profitable firm, facing less extreme financial constraints, can be expected to be able to spend on R&D more easily than an unprofitable, financially constrained one. It requires the aggregating leap of assuming therefore that a profitable industry will be more likely to invest in R&D than an unprofitable one, ceteris paribus, to apply this to a sectoral-level model. Once again, there may be the possibility of endogeneity here. A country-sector that is more technologically advanced (or has a better than average non-price competitive position) may have an edge over competing country-sectors allowing it to return better than average profits for the sector. However, as before, intuitively it is R&D output and profits where the potential endogeneity problem will be more pronounced and not R&D expenditure.

An explanatory variable that can be drawn directly from the model outlined in Section 4.4 is scale. Abstracting from trade momentarily, a common prediction of endogenous growth models (Romer, 1990, for example) is that a larger economy would grow faster than a smaller one. In terms of R&D incentives, in a large economy (the US, say) there is a greater potential reward to successful innovation and a greater stock of scientific knowledge on which to feed than in a smaller economy (Finland, for example). Theoretically this will be reflected in higher R&D incentives in the US than in Finland, other things being equal.[97] The same principal can be applied at sectoral level. On top of this, it should be obvious that in general larger firms will spend more on R&D than smaller firms. By this I do not mean larger firms will necessarily be more R&D *intensive* than smaller firms, which, at sectoral level, is the scale effects argument above, but that R&D expenditure might rise in proportion with size. Aggregation suggests equally that larger industries will spend more on R&D than smaller ones.

Some measure of the size of the market is a regular feature in R&D equations, most commonly sales. The common sense 'proportional' argument would suggest a positive coefficient on size of around unity although there may be complicating factors. For example, a small industry that is nonetheless growing, may invest more in R&D than a similar sized industry in decline.[98] I follow previous studies by including a measure of sectoral size in each country. The scale effects argument above would suggest a positive coefficient on size greater than one, with R&D increasing more than proportionally with scale.

There is a potential problem here with the size variable capturing part of the openness effect, namely the demand-led effect of trade on R&D and the knowledge-driven trade scale effect. However, this is not necessarily as much of a problem as it appears at first glance. It allows some of the different openness effects to be partially disentangled. A size coefficient greater than one would suggest the presence of scale effects, which would support the existence of a positive relationship between trade and R&D through these scale effects. The openness variable, in the presence of this size variable, may be capturing the non-scale effects of openness on R&D, such as competition and duplication effects. In practice, both variables are likely to capture a little of the trade-based scale effects. This may act to reduce the strength and significance of any openness/R&D relationship found. In this sense, I may actually be underestimating the importance of the R&D/trade relationship.

Some wider measure of demand at the country level may also have an effect. Over and above the effects hypothesized above, a country in recession may invest less in R&D than a booming one. I therefore test for the inclusion of national economic growth in the equation.

[97] I am assuming closed economies here.
[98] Also, the presence of the LDV has implications for the size coefficient.

Investment in physical capital has been included in previous studies. Lach and Schankerman (1989) argue that common shocks might affect both R&D and investment in fixed capital, so that once the other factors driving R&D are included, fixed capital investment may be able to capture these remaining common shocks. Firm-specific factors such as risk-premia or management characteristics are possible candidates for such common shocks. This would give an expected positive relationship, as argued originally by Schmookler (1966). In this case I am assuming these characteristics are correlated across firms in a given sector and therefore aggregate to industry level.[99] However, there may be other forces at work here. For example, if a firm faces a choice, after dividends, of spending retained profits on R&D or on fixed capital investment, this might in part counteract this positive relationship. There is a clear and unequivocal endogeneity problem with this variable, with R&D and fixed capital investment being jointly determined. It may even be that any relationship in fact runs in the opposite direction of causality, with capital investment being needed to implement R&D (see Aghion and Howitt, 1998, for such a model, and Lach and Schankerman, 1989; Nickell and Nicolitsas, 1997; Stoneman and Kwon, 1997, for supporting evidence). Where a relationship is found from R&D to investment in these studies, it is with positive sign. Stoneman and Kwon (1997), for example, find an elasticity of investment with respect to R&D expenditure of around 4%, although the relationship often appears insignificant. Against this opposite causality is the timing argument once again.

It is generally accepted that R&D expenditure is somewhat persistent over time. Around 90% of expenditure on R&D goes on skilled labour and the equipment (lab-equipment) they need to carry out research (see Nickell and Nicolitsas, 1997, for a discussion). The cost of adjusting R&D expenditures is therefore based on the cost of adjusting the employment of skilled labour, which is likely to be high (at least over the short term). On top of this, R&D projects are typically undertaken over a number of years, and therefore a given budget may be divided over a number of periods in such a way as to add to this persistence. I therefore also include a lagged dependent variable (LDV) in the R&D equation, to capture this persistence, making the R&D equation a dynamic one. In this I follow Van Reenen (1996), Nickell and Nicolitsas (1997) and others.

The presence of the LDV has some important implications for the interpretation of the results, aside from the well known econometric difficulties it causes (of which more later). Firstly, since the system is now dynamic, a distinction needs to be made between short run and long run effects. For example, an innovation (shock) in the profitability variable at year t will have a short run effect on R&D at year t given by the relevant coefficient. However, because R&D at year t enters into the equation for R&D at year t+1, this innovation in profitability will also affect R&D at year t+1 and so on into the future. In other words, an innovation at time t will have both a short run effect given by the coefficient and persist to give a long run effect (on R&D at time $t+\tau$, $\tau \rightarrow \infty$) given by the coefficient over one minus the coefficient on the LDV.[100] So a change in one of the explanatory variables, although it may well have an immediate effect, will have an overall effect that builds up over time because of the dynamic nature of the relationship. When interpreting the effects of openness to trade and scale, for instance, on R&D expenditure, we must bear this in mind.

Secondly, the likely persistence of R&D expenditure has implications for the endogeneity or otherwise of the explanatory variables in the equation. I have argued briefly above that R&D spending today cannot intuitively affect trade or production today, say, because the results or output of that research will take time to come through, quite possibly a number of years. This suggests that the explanatory variables (with the exception of investment) can be treated as exogenous. However, if R&D expenditure today is closely related to R&D expenditure in the future, and R&D spending today might affect future trade or production, then R&D in the future might look as if it

[99] Technological opportunity is another candidate for a common shock at industry level.

[100] This can be seen from repeated substitution. Take the simple system as below (ignoring error terms etc.):

$y_t = \beta x_t + \gamma y_{t-1}$

The short run effect of x on y is just given by the coefficient beta above. However, repeated substitution gives:

$y_t = \beta x_t + \gamma(\beta x_{t-1} + \gamma y_{t-2})$ and so on until:

$y_t = \beta \Sigma \gamma^i x_{t-i}$, which, as i tends to infinity, is just:

$y = \beta x/(1-\gamma)$. So the long run effect of a change in x on y is given by $\beta/(1-\gamma)$.

explains part of trade or production in the future. This will reduce the chance of testing negatively for the endogeneity of these variables. This, too, must be borne in mind.

4.5.2: An Empirical Model

Although the list of potential regressors could go on forever, I am constrained by data limitations to a select few. Those above capture the factors of interest from the standard models of integration and R&D, such as that outlined in the previous section, along with the main remaining variables consistently included in previous studies. These variables together provide the empirical model, as given below by equation (4.21).

(4.21) $R\&D_{it} = f(R\&D_{it-1}, \textit{Openness, Scale, Investment, Profitability, Country Effects,} \ldots)$

The natural empirical specification for this model is given by equation (4.22), and this provides an estimable equation:

(4.22) $R\&D_{it} = \beta_{0i} + \beta_1 R\&D_{it-1} + \beta_2 OP_{it} + \beta_3 PROF_{it} + \beta_4 SIZE_{it} + \beta_5 INV_{it} + \varepsilon_{it}$.

All variables are in logs with i indexing countries (i=1-14) as outlined in the following section. Equation (4.22) is estimated separately for each sector as a panel of countries over time, allowing for group dummies (country-specific effects) in the formulation. The country dummies should reduce any potential simultaneity problem, where more technologically advanced countries might trade more, by acting to capture initial technology levels. Estimating this equation by industry allows us to control for possible unobserved sector-specific effects, for example, technological opportunity, which have also been suggested as possible explanatory factors in the R&D expenditure decision (see Cohen and Levin, 1989, for a discussion).[101]

Treating each sector as a separate panel also allows us to see whether any relationship varies systematically across sectors, for example, between high and low technology

[101] Further research could set up panels of sectors by countries to test the robustness of the results.

industries. It may be that high technology industries are characterised by more intense R&D competition and more significant spillovers, for example, which might affect the R&D/openness relationship in different ways.

4.6: Some Data Details

I have industry level annual data for 13 manufacturing sectors across 14 OECD countries between 1973 and 1992 taken from the OECD STAN and ANBERD data sets. These sectors are listed in Appendix B, along with more detailed source information. Supplementary data is from the OECD National Accounts, various years. Unless otherwise stated, all variables are at sectoral level in US$1990, where purchasing power parity (PPP) data, taken from STAN, has been used to transform from national currency current values.

R&D expenditure data (ANBERD) exists for all 13 sectors over the time period. Expenditure is used to proxy for R&D incentives, assuming changes in incentives are reflected by changes in expenditure. The data itself gives all business enterprise total intramural expenditure on R&D, as estimated by the OECD. They note that these estimates may be different from official data for a number of countries, a point also made for the STAN data. Set against this is the obvious advantage of the international comparability of the data.

In general, R&D expenditure has been rising, in real terms, across countries and sectors over the sample period. For example, R&D in the US manufacturing sector as a whole increased from $53.6bn in 1973 to $83.7bn in 1992, at 1990 prices. Over the same period, UK R&D in total manufacturing increased from $7.7bn to $10.1bn. Production and income levels have also been rising, but R&D has generally increased over and above this growth, so that R&D intensity (R&D/production), or R&D as a share of GDP, also displays an upward trend. Figure 4.1 shows R&D intensity for the 14 countries (as listed in Appendix B) over the sample period for total manufacturing. There is clearly an upward trend here for all countries, although not always monotonic. For example, R&D

intensity in US manufacturing increases from below 2.5% in 1973 to above 3% in 1992, although it peaks at 3.5% in 1986, and for Japan from 1% to over 2.5%. There is also a wide disparity across countries in R&D intensity, with, for example, Australia's manufacturing R&D intensity at around 1% in 1992 as compared to that of the US, around 3%. In fact the US is the most research-intensive nation throughout the period until the 1990s when it is overtaken by Sweden. Of the sampled countries, Spain remains the least R&D intensive in total manufacturing for the entire sample period. The UK displays a relatively flat curve, with little growth. The differences across countries suggest the need for country specific effects to be allowed for in the empirical model by the inclusion of group dummies.

Figure 4.1: R&D Intensity, By Country, Total Manufacturing

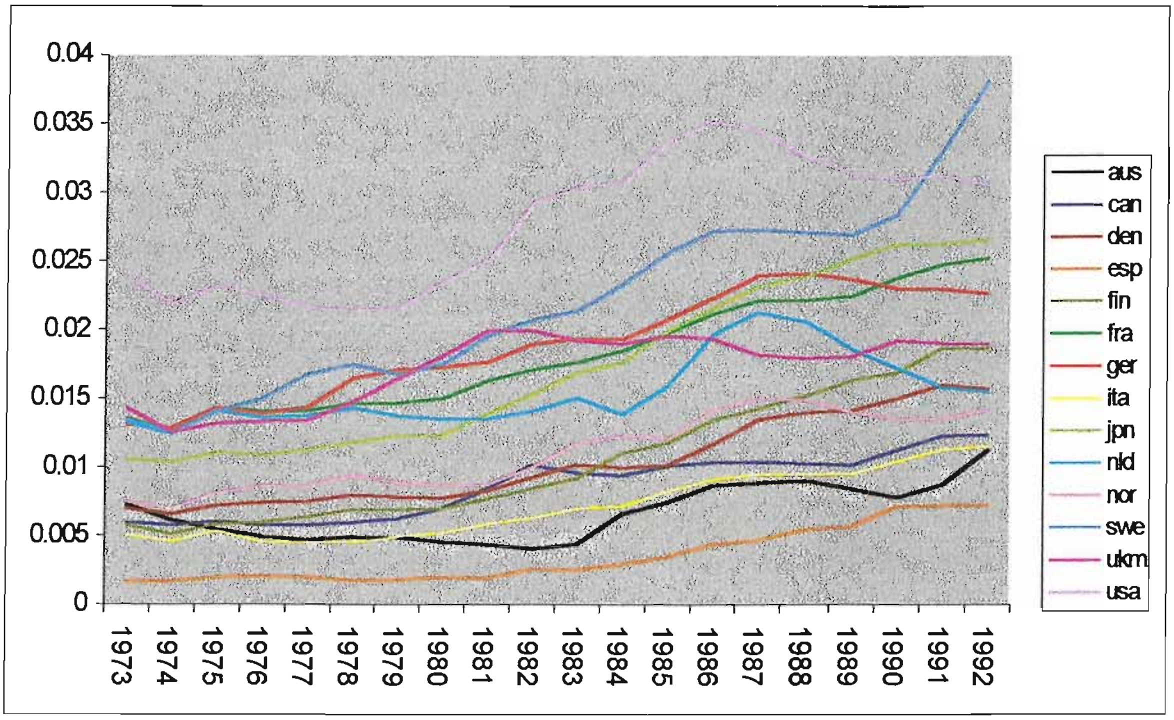
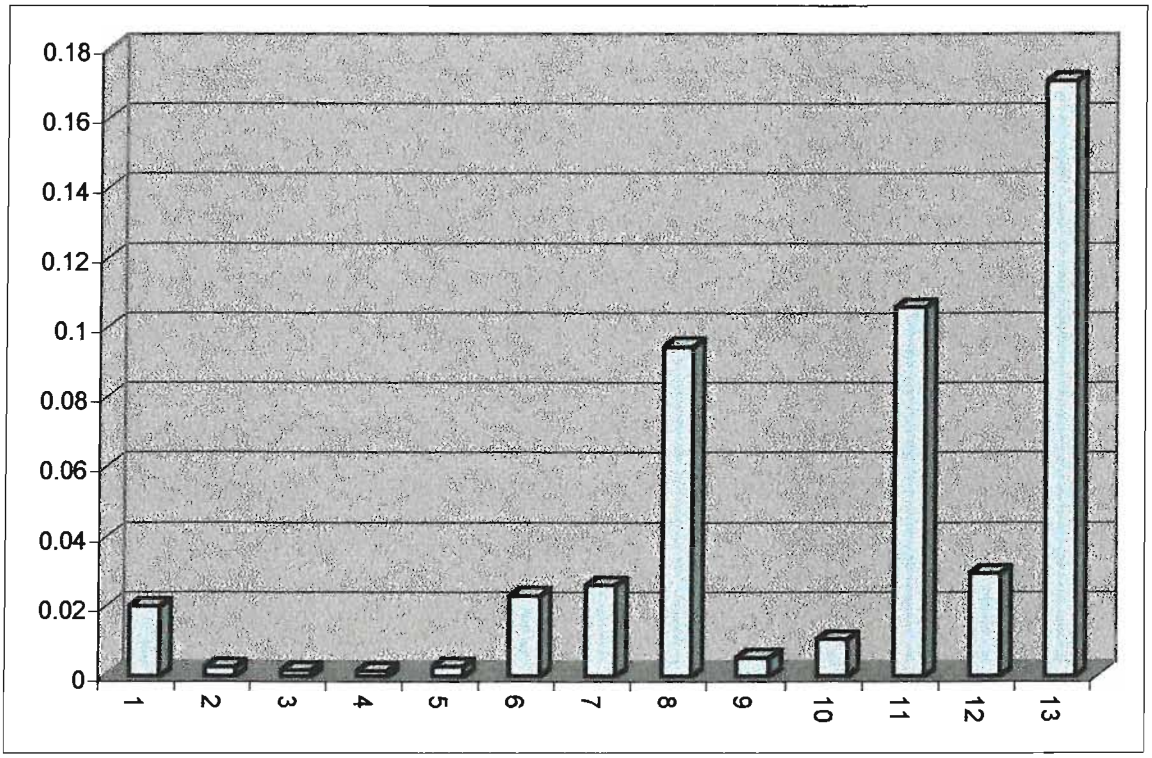


Figure 4.2 gives a sectoral breakdown of average R&D intensities, across all countries for the sample period, for the thirteen manufacturing sectors (listed in Appendix B). Sector 1 is total manufacturing, for which the mean R&D intensity is below 2%. The disparity across sectors is much more dramatic than that across countries for total manufacturing. At the extremes, wood products and furnishings (sector 4) has an intensity considerably less than half of one percent, with an R&D intensity for the

aerospace sector (sector 13) of around 13%. From Figure 4.2 sectors can be classified as high, medium or low technology industries, with these three distinct groups forming. High technology industries in the sample are drugs and medicines, office machinery and computers and aerospace, all with mean R&D intensity above 8%. The middle group of medium technology industries is characterised by R&D intensities between 1% and 8%, including chemicals excluding drugs, industrial chemicals, motor vehicles and total manufacturing. The remaining industries are classified as low technology, with mean R&D intensity below 1%. They are food, textiles, wood, paper, petrol and rubber.

Figure 4.2: R&D Intensity, By Sector, Across Countries



Notes: Sectors 1-13 are defined in Appendix B. Figures given are means across countries and across the sample period.

The primary explanatory variable is openness to trade. There are well-documented problems with trying to establish how open a particular country, or country-sector, is to trade (see, for example, Harrison, 1996; Pritchet, 1996). Pritchet (1996) shows that many commonly used measures of openness are in fact, with few exceptions, pairwise uncorrelated, which helps to explain the lack of robustness displayed by many studies of growth and openness (Levine and Renelt, 1992). Studies generally held to have been most successful (by, for example, Aghion and Howitt, 1998) are those that use a variety of different measures to test this robustness (or lack of it); Harrison (1996) is a good example. However, I have to rely on a single measure of openness to trade, measured as exports plus imports over production (trade intensity) because of data limitations in a broad sectoral level sample such as this. Fortunately, this measure is arguably the most suitable for such a study, with full and accurate time series available from the STAN data set for exports, imports and production at sectoral level.

Possible measures of openness are numerous. Commonly used measures, and examples of studies in which they are applied, include trade intensity or import/export penetration (see, for example, Levine and Renelt, 1992; Cameron et al, 1998); tariffs (Lee, 19931 Barro and Sala-I-Martin, 1995); estimated non-tariff barriers to trade (Nevin and Roller, 1991; Driffill et al, 1997); black market premia on currency exchange (Barro and Sala-I-Martin, 1995; Lee, 1993); foreign direct investment (World Bank, 1996; Cameron et al, 1998) and others. Some measures of openness can be built up from a number of such component parts, such as the World Bank composite index (World Bank, 1996) or the overall index used by Cameron et al (1998).

Ideally, I would like to follow those papers that apply a variety of measures of openness, or a broad composite measure. As mentioned above, this has unfortunately not been possible in this Chapter due mainly to data limitations. For example, in the same country, tariffs may be set at different levels on imports from different source countries. The EC has a non-zero external but zero internal tariff-rate on many goods. To use a measure of openness such as this would require bilateral trade data at sectoral level in order to construct some average tariff level. This was deemed too lengthy a process at this stage, but may be an avenue for further research. Non-tariff barriers suffer from similar construction problems and existing indices (such as Nevin and Roller's NTB Index) are too imprecise and have too narrow a coverage. Black market premia on currency exchange was not considered since it is clearly not suitable for this particular sample of advanced countries.

Once these alternative measures have been ruled out, we are left with measures of trade intensity and possibly some measure of FDI. Cameron et al (1998) builds up a sectoral level composite index of openness for the UK from trade intensity and FDI data. This too is not possible here, since FDI data at sectoral level does not readily exist for all the countries over the sample period. In the end, trade intensity is the only available measure of openness to trade. The natural choice for such a measure is the overall intensity of trade, defined as the ratio of imports plus exports to production.[102] This reliance on a single measure of openness is a clear weakness with the current study, and one that it should be possible to address with alternative data in further research.

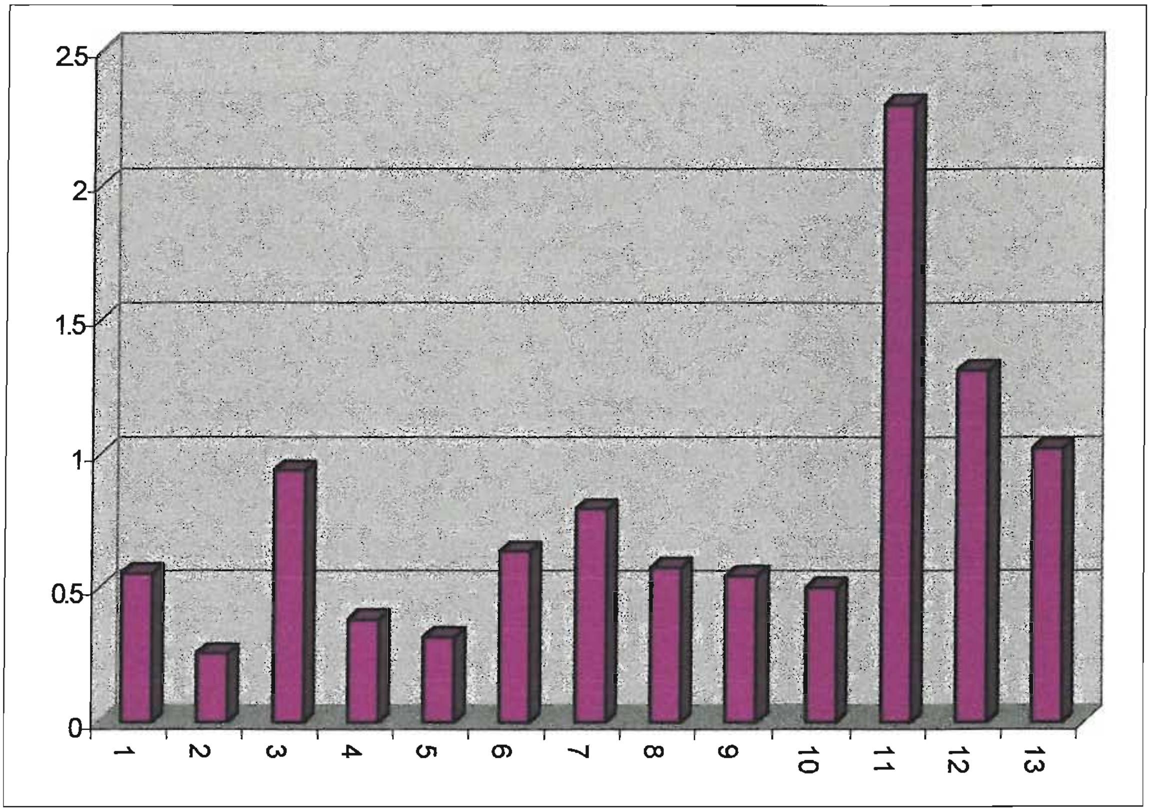
Figure 4.3 shows trade intensity, across countries for the whole sample period, by sector. As shown for R&D in Figure 4.2, there is substantial variation across industries in terms of their openness to trade. Sectors range from means of around 0.25 (food) to around 2.25 (office machinery and computers). Using this graph, sectors can be classified into high or low trade intensity sectors, in a similar way to the classification for high/low technology sectors. For example, we can think of sectors 11, 12 and 13 (office machinery and computers, motors and aircraft) as high trade intensity sectors, with means greater than 1. Sectors with openness means below 0.5 we can classify as low trade intensity (food, paper and wood) and the remainder (0.5-1) we can classify as medium trade intensive.

There is no consistent pattern in the correlation of trade intensity and R&D intensity across sectors in the raw data. Table 4.1 shows the ranges of sectoral pairwise correlation coefficients for the key variables in the model. Correlations for openness to trade and R&D intensity range from -.35 to +.37. This lack of consistency echoes the ambiguity of the trade/R&D relationship discussed in Section 4.2. There is a more consistent pattern when looking at the correlation between openness to trade and the level of real R&D expenditure with correlation coefficients ranging from -.56 to -.30. In other words, the raw data suggest that R&D expenditure is lower in those country-sectors where trade intensity is high. Of course, we need to turn to the econometric analysis outlined in Section 4.8 in order to test the true nature of this relationship in a causal sense. However,

[102] This ratio usually falls between 0 and 1 but is not bounded by 1.

the background to the econometric analysis is one of there being no pattern in the raw data suggesting a positive openness/R&D relationship, and perhaps even some evidence of a negative pattern.[103] This provides a first 'stylized' fact.

Figure 4.3: Openness to Trade, By Sector.



Notes: Figures given are (exports + imports)/production, means across countries and time. The sectors 1-13 are defined in Appendix B.

Consider the other explanatory variables in equation (4.22). Firstly, I use the value of production, as given in the STAN database, as the scale variable. Many previous studies, particularly firm level studies (see, for example, Van Reenen, 1996), have used the volume of sales for their size variable. The advantage that sales figures have over production figures is that they should more closely reflect demand. Unfortunately, I do not have suitable sales data for the sample. However, in theory, if markets are in or close

[103] It would be interesting to look at the correlations for individual countries across sectors. For example, the position of the US as the highest R&D spender and the least trade-intensive nation in the sample may be partly responsible for driving these apparent patterns in the raw data. This alternative panel set up is a possibility for further research.

to equilibrium, production figures (supply) should be very close to sales figures (demand). On top of this, by aggregating to industry level, I should be averaging out any firm-specific discrepancies between supply and demand, which should bring the two measures even closer together. As we should expect, production is strongly correlated with R&D expenditure across all sectors, with correlation coefficients ranging from a low of 0.67 (chemicals excluding drugs) to a high of 0.98 (paper). More informative is the evidence of a positive correlation between production (size) and R&D intensity. This is positive for all but one sector (textiles), although correlation coefficients, ranging from -.10 to 0.69 (motor vehicles), are typically smaller. This suggests that larger sectors tend to be more R&D intensive, or in other words, that there is a positive *scale effect* in R&D expenditure. This provides an interesting 'stylized fact'.

Gross fixed capital formation data is provided in the STAN data set and this is the investment variable. A strong positive correlation with R&D expenditure is expected, which is indeed the case for 12 out of 13 sectors.

Finally, it was necessary to construct a profitability (or financial constraints) variable. Unfortunately, I cannot refer to company accounts as many firm-level studies have done. However, an industry-level returns-to-capital variable can be constructed in the following way. Firstly, capital stock series for each industry/country pair are constructed from the gross fixed capital formation data, using the perpetual inventories method.[104] Given the capital stock estimates, I estimate profitability by value added minus labour costs over capital stock, given in (4.23).

(4.23) $\Pi_{it} = (\textit{Value Added}_{it} - \textit{Labour Costs}_{it})/\textit{Capital Stock}_{it}$.

An industry with high returns to capital is assumed to be more profitable than one with low returns to capital. Section 4.8.1 discusses the legitimacy of this assumption in the light of the empirical results. From Table 4.1, this profitability measure appears to be largely unrelated to R&D expenditure, with correlation coefficients for sectors of between -.23 and +.33. The correlation pattern with R&D intensity, although still weak,

[104] See Chapter 5 and Appendix C for a discussion of this method.

is slightly stronger, with an average correlation coefficient of around -0.3. This contrasts with what we would expect to be a positive relationship because of the financial constraints argument.

The panels are balanced for each separate sector by discarding those series with missing observations. Since the missing observations are not evenly spread across the series, this gives heterogeneous sample sizes for the different industries, with different countries and/or different years dropping out in each. Details on the specific samples used for each industry are presented in Appendix B. The sample sizes range from 280 observations (14 countries over 20 years) to just 108 observations (6 countries over 18 years).

Table 4.1: Pairwise Correlations

	R&D	**R&D/Prod**	**Openness**	**Production**	**Profitability**	**Investment**
R&D		.32→76 (.68)	-.56→-.30 (-.47)	.67→.98 (.96)	-.23→+.33 (-.07)	-.09→+.96 (.88)
R&D/ Prodn	.32→76 (.68)		-.35→+.37 (-.01)	-.10→+.69 (.60)	-.40→+.34 (-.33)	-.13→+.59 (.57)
Open-ness	-.56→-.30 (-.47)	-.35→+.37 (-.01)		-.60→-.37 (-.60)	-.69→+.09 (-.34)	-.59→+.61 (-.59)
Prodn	.67→.98 (.96)	-.10→+.69 (.60)	-.60→-.37 (-.60)		-.20→+.43 (-.05)	-.13→+.97 (.94)
Profit-ability	-.23→+.33 (-.07)	-.40→+.34 (-.33)	-.69→+.09 (-.34)	-.20→+.43 (-.05)		-.22→+.38 (-.10)
Invest ment	-.09→+.96 (.88)	-.13→+.59 (.57)	-.59→+.61 (-.59)	-.13→+.97 (.94)	-.22→+.38 (-.10)	

Notes: Figures given are the range of sector-specific (Pearson) correlation coefficients. Figures in parentheses are the correlation coefficients for sector 1 (total manufacturing).

In summary, the following variables enter into equation (4.22): Real R&D expenditure (RRD), openness to trade (OP) defined as exports plus imports over production, real production (RPR), real gross fixed capital formation (RI) and profitability (PROF) defined as value added minus labour costs over capital stock. All variables are in logs. In addition, I test for the inclusion of growth (GR). Group dummies are included in the specification of (4.22) to pick up any country-specific effects (as suggested by the wide disparity between countries in R&D expenditure). I also allow for a time trend in the specification to pick up any omitted trending variables.

Two interesting regularities in the data have been pointed out. These are:

1. More open industries do not appear to be more highly R&D intensive, or spend more on R&D than less open industries.

2. Larger industries tend to be more highly R&D intensive than smaller ones.

The first fact questions the existence of a consistent positive link between R&D expenditure and openness to trade and reflects the ambiguity in the trade/R&D relationship discussed in Section 4.2. The second fact suggests that there is a positive link between scale and R&D intensity (scale effects). Of course, we cannot read too much into these simple pairwise correlations. A multiplicity of factors are at work here which need to be controlled for before having any real confidence about the existence and direction of relationships between the variables. In order to do this, the econometric techniques discussed in the following section must be applied.

4.7: Econometric Considerations

4.7.1: Testing for Unit Roots

The approach adopted to estimating equations such as (4.22) depends on whether we believe the series to be trend-stationary (I{0}) or difference-stationary (I{1}). Well-known procedures exist for testing for unit roots. Unfortunately, these tests are often very inconclusive.

Testing for unit roots with short series is beset by problems of low power, so that rejection of the null hypothesis of a unit root is very difficult. However, recent work (see, for example, Quah, 1994; Im et al, 1996) finds that the performance of standard Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) tests can be improved upon by exploiting the cross-section dimension of the data.[105] I test for unit roots following the Im et al (1996) t-bar testing procedure for panel data series. Of course, with limited observation time-series such as those used in this Chapter (a maximum of 20 annual observations) even these tests are beset by low power problems. They have also been criticised for other reasons, as discussed below.

The Im et al (1996) t-test (hereafter the IPS test) is intuitively a simple extension of the standard DF and ADF tests. The IPS equivalent of the DF test (where there is no lag

---

[105] DF tests amount to an examination of the t-statistic of the lagged y-variable in the equation below, known as the Dickey-Fuller equation:

$\Delta y_t = \pi y_{t-1} + \beta' DR_t + u_t$.

DR above signifies the deterministic regressors in the process (eg: a time trend and constant). The standard test, for which the above t-statistic is used, is as follows:

H_0: $\pi=0$, ie: unit root,

H_1: $\pi<0$, ie: stationary.

The ADF test is an extension of the standard DF test to allow for additional serial correlation in the data. it is carried out in precisely the same way as the DF test, except lagged differences are included on the right hand side of the DF equation, as below:

$$\Delta y_t = \pi y_{t-1} + \sum_{i=1}^{p} \Delta y_{t-i} + \beta' DR_t + u_t.$$

The order of the last lagged difference included in the regression is the order of the ADF test (order p). See Campbell and Perron (1991) for a full discussion of these tests in practice.

structure in the DF equation) is simply the mean of the t-statistics for the individual series in the panel. In this case, the basic IPS test statistic for the R&D series for total manufacturing is the mean of the t statistics for the total manufacturing R&D series for each individual country. The null and alternative hypotheses are defined in the same way as for the standard DF/ADF tests outlined above, ie: a one-sided test with the null of a unit root. The IPS test equivalent of the ADF test where there is a lag structure of order p in the ADF equation is based on the same principle. The mean t-statistic value from the separate ADF equations is normalised and distributed according to (4.24):

$$(4.24)\quad \Gamma = \sqrt{N}\{\bar{t}_{NT} - E(t_T)\}/\sqrt{Var(t_T)} \Rightarrow N(0,1),$$

where N is the width of the panel and T the length. The mean and variance of the t-bar distribution under different conditions are tabulated in Im et al (1996). Once again it is a one-sided test with the null of a unit root.

The IPS tests for the sectors 1-13 have a maximum power of under 40% and a minimum of under 10% (see Im et al, 1996), so they are still open to the likelihood of falsely accepting the null hypothesis. In addition, Matyas (1996) points out that the nature of these tests is to accept the null of a unit root for all series of cross-sectional units even if only one unit's series is truly non-stationary. In other words, if the series for the US, for example, displays unit-root like behaviour, but none of the other countries do, the IPS tests may conclude that all the series have unit roots.

It is important not to miss-specify the testing equations for the individual series in terms of the presence or otherwise of a time trend or the order of the ADF tests. Omission of a trend from the regression if a trend is present in the data generating process (DGP) causes the power of the tests to tend to zero as the sample size increases (Campbell and Perron, 1991). Power is also reduced by the inclusion of a trend if no trend is present in the DGP. The reason for including tests both with and without trends in Table 4.2 is that it is not clear in all cases whether such trends should be included in the DF/ADF regressions or not. Some idea of the presence or otherwise of a trend can be obtained by running the DF/ADF regressions on the individual series. This evidence suggests the inclusion of a time trend for some country-sectors but not for others. The tests including and not including time trends are therefore both likely to over-accept the null hypothesis.

Specification of the correct lag length for the ADF regressions can also be problematical. Miss-specification can adversely affect the size of the test when too few lags are included or the power of the test when too many lags are included (Campbell and Perron, 1991). Im et al (1996) suggests over-fitting when in doubt. Generally, preliminary exploration of the series suggests lag structures tend to be of orders not more than two. Therefore, I carry out IPS tests of orders zero, one, two and three for all series.

The results of the IPS tests of various orders, with and without time trends are reported in Table 4.2 below, for the total manufacturing sector. The 5% critical values for the DF-equivalent tests with and without trends are –2.55 and –1.91 respectively. The 10% critical values are similarly –2.46 and –1.82. For the ADF-equivalent tests, the normalised test-statistics are distributed standard normal, with means and variances given in Im et al (1996, Tables 3 and 4).

Table 4.2: IPS Unit Root Tests, Total Manufacturing

	DF	DFt	ADF1	ADF1t	ADF2	ADF2t	ADF3	ADF3t
RRD	.02	-1.97	2.54	0.34	1.58	-.15	.38	-2.55**
RI	-1.16	-2.06	-.79	-2.69**	0	-.51	-.28	-1.16
RPR	-.97	-2.29	2.77	-1.76*	.22	-.22	2.82	-1.60
OP	-1.27	-2.57**	2.54	-.11	.29	-2.16**	-.41	-1.53
PROF	-1.84*	-1.49	-2.17**	1.12	-3.42**	.29	-1.83*	1.90
GR	-3.50**	-3.55**	-6.10**	-3.63**	-2.59**	-.62	-3.93**	-.1.73*

Notes: Variable names are defined in Section 4.6. DF denotes the DF-equivalent IPS t-test without trend and DFt with trend. Similarly, ADFp is the IPS t-test of order p without trend and ADFpt is the IPS test of order p with trend. Rejection of the null of a unit root at 5% is marked by ** and at 10% by *. Variables are in logs.

Leaving aside the problems with power and the Matyas critique, the results are still far from conclusive in some cases. The growth and profitability series can be assumed to be stationary with a fair degree of confidence. The evidence is also suggestive of a stationary openness series, although not conclusively so. The evidence for the remaining series is mixed. A priori, we expect the production, investment and R&D series to be the most likely candidates for unit roots. Macroeconomic output (GNP or GDP) is commonly taken as persistent (see, for example, Campbell and Mankiw, 1987), and it is a small step from this to total manufacturing production.[106] Jones (1995) argues that investment at the macro level may also be non-stationary. On the other hand, individual sectors, particularly those dominated by a small number of large firms such as aircraft, are less likely to display such persistence. Firm-level studies usually assume stationarity for these series (see, for example, Lach and Schankerman, 1989).

Given this ambiguity, and given the problems with the tests discussed above, it seems sensible to proceed along two alternative directions. Firstly, assume all series are stationary.[107] Secondly, assume the series for R&D, production and investment are difference stationary. The remainder of this Chapter proceeds along the first route, assuming the series are I(0). The alternative route assuming I(1) series is left for further research.[108] As an exception to this, (4.22) is estimated in first differences by least squares dummy variables (LSDV) to test the robustness of the results to the possible presence of unit roots or near unit roots.

4.7.2: Exogeneity

It is also necessary to establish the exogeneity or otherwise of the regressors in the R&D equation before consistent estimation can be carried out. We have to decide whether to treat the remaining regressors as exogenous (uncorrelated with the error term) or

[106] This is not always the case, however. Arestis et al (1998) suggest GDP for Finland is I(0), for example.
[107] This is assumed by a number of recent empirical R&D studies (see, for example, Nickell and Nicolitsas, 1997).
[108] Performing the IPS tests on the first differences of the series conclusively rejects any possibility of I(2) variables.

endogenous (correlated with the error term). The simplest way to do this is by *a priori* reasoning. Of all the regressors, we would firstly expect the investment term to be endogenous, as it is included in the equation partly to capture unobserved common effects on investment and R&D expenditure. It may also be that a reverse causality relationship exists between R&D and investment in fixed capital, as discussed in the previous sections. It can be argued that the other regressors should be exogenous, primarily due to either the fact that R&D is sectoral only (in the case of growth) or that only past R&D could effect current levels of openness, profitability or production.

Although we can be guided by this kind of reasoning, it is also useful to try to establish exogeneity/endogeneity of variables statistically with a simple testing procedure. A further reason for carrying out these tests is that because R&D is persistent, R&D expenditure today may statistically affect the explanatory variables today because of its close correlation with R&D expenditure yesterday. So, despite the presence of a time lag between R&D and the likely effects of R&D on the regressors, in a purely statistical sense, these regressors may not be truly exogenous. For Sector 1, I apply the Wu-Hausman test to the non-LDV regressors, country by country.[109] In practice, this test amounts to running reverse regressions and adding the resulting residual to the original equation and carrying out an F-test of its significance as an addition. If the added residual is not significant, the relevant variable can be taken as exogenous. For example, to test the investment variable, the following process is carried out:

(i) Estimate, by least squares dummy variables (LSDV) on the pooled series after within transformation, the equation $RI_t = const + b_1RI_{t-1} + b_2 RRD_t$ + other regressors + residual, and save the residual.

(ii) Estimate equation (4.22) by LSDV and carry out an F-test for the addition of the residual to the equation.

---

[109] See Wu (1973).

Table 4.3: Wu-Hausman Tests, Total Manufacturing

	F-statistic
Investment (RI)	27.21***
Openness (OP)	.11
Production (RPR)	4.21**
Profitability (PROF)	.49
Growth (GR)	.81

Notes: Rejection of the null of exogeneity at 1% is denoted by *** and at 5% by **.

The null hypothesis is that the residuals are not a significant addition to the model, which implies the investment variable is exogenous. Table 4.3 reports the F-statistics for total manufacturing. The 5% critical value for this test is 3.94 (F(1,200+)).

The results confirm the expectation that the primary variable of interest in terms of endogeneity is the investment variable, for reasons elsewhere discussed. Investment is therefore treated as endogenous in the estimation. The other variables mostly accept the null, with the possible exception of the production variable. These variables are therefore treated as exogenous, taking the test results together with the *a priori* reasoning.

4.7.3: Estimation of Dynamic Panel Data Models

For each sector, I have a cross section of up to 14 countries, for a time period of up to 20 years, with annual observations. For some sectors, balancing considerably cuts down the dimensions of the sample (see Appendix B for details). In all likelihood, there will be country-specific effects that are not captured by the explanatory variables (eg: availability of skilled workers, policy conditions etc.), so the natural model to estimate would be the least squares dummy variable (LSDV) model, with country-specific fixed effects.[110] It is also likely that the cross-country nature of the panel will imply differences in the scale of the error terms across groups (groupwise heteroskedasticity), and correlation across groups (cross-sectional correlation). The time series nature of the data makes it likely that there may also be some serial correlation of the errors (autocorrelation). The choice of best estimators will depend on the heteroskedasticity and correlation properties of the data, so I carry out simple tests for these properties (details in Appendix B) and use an appropriate Generalised Least Squares (GLS) estimator.[111]

The bias of the LSDV estimator of the LDV is well known and has been shown to be of the order O(1/T), or proportional to the inverse of the time dimension of the panel, and to act in a downward direction (Nickell, 1981).[112] The bias on the other coefficients can act in either direction, but are usually smaller in magnitude than the LDV bias (see Judson and Owen, 1997). So, for anything other than a panel with a large time dimension, LSDV will give biased estimates of γ and β in (4.22).

Recent empirical work in this and related fields has typically encountered this difficulty, and a lot of discussion exists on the problems associated with estimation of dynamic panel data models (see, for example, Van Reenen, 1996; Nickell and Nicolitsas, 1997, for applications; Kiviet, 1997; Judson and Owen, 1997). Broadly speaking, two solutions to the problem of the downward biased LDV are proposed (and applied) in this literature.

The first suggestion, which is stretching the language to describe as a *solution*, is to carry on with LSDV anyway. This is not as careless as it at first appears, especially when the T dimension of the panel is sufficiently large, and the N dimension not so large, and past work (eg: Nickell and Nicolitsas, 1997; Cameron et al, 1998) has used this approach. Judson and Owen (1997) try to address the question of how big T has to be before the

[110] The fixed effects model is the natural specification for such data, with a small cross section dimension of similar units. It also has the advantage over the random effects model if the group effects are not orthogonal to the error term (for a discussion of the relative merits of fixed and random effects estimators, see, for example, Greene, 1993; Judge et al, 1985).
[111] Data is first transformed by the Cochrane-Orcutt transformation and then estimated by GLS. See Greene (1991, p281-283 for details).
[112] The bias arises because the LDV is correlated with the time-invariant group dummies. After the Within transformation, the demeaned LDV is still correlated with the demeaned error term (ε_{it} - ε_i) in (4.22), even if the errors themselves are not serially correlated (Baltagi, 1997).

bias can be ignored. They find that even with a time dimension as big as 30, the bias on the coefficient of the LDV can be as much as 20% of the coefficient value. When T=20, as in the current case, they find a bias of between 2% and 28% of the true parameter value. They conclude that, although a bias of this magnitude would still result in an estimate with the correct sign, and that the biases on the other coefficients are small in comparison, LSDV does not generally provide us with the best possible estimator of the parameters of the model.

A number of alternative estimation techniques have been applied to such models, also evaluated by Judson and Owen (1997), which involve the use of instrumental variables. These are discussed below. A closer alternative involves a two-stage procedure to correct for the bias of the LSDV estimator. The first stage involves a consistent IV-type estimator which is then used to estimate the LSDV biases (see Kiviet, 1997). This technique has been found to perform well in panels with a small time dimension, reducing the bias, although slightly increasing the standard errors of LSDV estimates. However, Judson and Owen (1997) find this corrected LSDV estimator to be outperformed by the simpler Anderson-Hsiao estimator (discussed below) in all cases.

The second solution to the bias-problem outlined above is to find an alternative, consistent estimation technique. Two related alternatives are commonly considered. The first is a relatively simple instrumental variables procedure, first suggested by Anderson and Hsiao (1981). The first step with this estimator is to first difference the data to remove the fixed effects. Equation (4.25) clearly shows the correlation of the first-differenced LDV with the differenced error term.

(4.25) $y_{it}-y_{it-1} = \gamma(y_{it-1}-y_{it-2}) + \beta(x_{it}-x_{it-1}) + (u_{it}-u_{it-1})$.[113]

However, it is possible to find a variable uncorrelated with the error term, although correlated with the differenced LDV, which can be used to instrument for the LDV in estimation of (4.25). Anderson and Hsiao (1981) suggest the difference of the second lag,

[113] It should be clear that estimating in levels or first differences (or with demeaned data) makes no difference to the parameter estimates.

(y_{it-2}-y_{it-3}), is the appropriate estimator, being unrelated to the error in (4.25). Arrelano and Bond (1991) find this instrument leads to high standard errors and suggest using the second lagged level, highly correlated with the first-differenced LDV but uncorrelated with the error, as a more efficient alternative. The other variables can be instrumented for by themselves or by lagged values, depending on assumptions of exogeneity. In this case, the set of instruments would consist of the second lagged level of R&D, the first lagged difference of investment, with all other variables (assumed predetermined) instrumenting for themselves. Calling this matrix of instruments **Z**, the Anderson-Hsiao (AH) estimator is given by:

(4.26) $\delta_{AH} = (\mathbf{Z'X})^{-1}\mathbf{Z'Y}$,

where **X** is the matrix of (differenced) regressors and **Y** the stacked differenced dependent variable.

This AH estimator has been found to outperform all alternative estimators, at least in terms of reducing the bias in small panels (see Judson and Owen, 1997; Kiviet, 1997). However, the reduction in bias comes with a significant reduction in efficiency. Judson and Owen (1997) find standard errors generally at least twice as big with AH as with LSDV. For applications, see Van Reenen (1996) and Judson and Owen (1997).

The second alternative procedure for dealing with panels including LDVs is to apply a Generalized Method of Moments estimator (GMM) of some form, as suggested by Arrelano and Bond (1991), again on first-differenced data (to remove the fixed effects), again instrumenting for troublesome variables with lags. In this sense, the GMM estimator is closely related to the AH estimator (in fact the AH estimator is a special case of the more general GMM estimator). GMM estimation utilises additional moment conditions in order to estimate the model, which can improve efficiency relative to the AH estimator.[114] In practice, this involves using additional lags as additional instruments

[114] For a detailed analysis of the use of GMM estimators in panel data models, see Ahn and Schmidt (1995) or Baltagi (1997). See, for example, Van Reenen (1996), Judson and Owen (1997) for applications.

over and above the AH instruments, giving the estimator below in (4.26).[115] These moment conditions, or *orthogonality* conditions, exist between the error terms and all lags of all variables (and contemporaneous values of exogenous variables). For example, from (4.25) we know we can use y_{it-2} as an instrument for y_{it}-y_{it-1}, since it is correlated with y_{it}-y_{it-1} but uncorrelated (orthogonal) with the error term u_{it}-u_{it-1}. Now consider the next period (t+1). Here, y_{it-1} is a valid instrument in the same way as the second lag at period t. However, an additional orthogonality condition exists between the error and y_{it-2} (they are also uncorrelated), which means we can also use this as an instrument. This process can be repeated until we see that all available lags may be used as instruments.

$$(4.27)\ \delta_{GMM} = \{\mathbf{X'Z^*}(\frac{1}{N}\sum_{i}^{N} \mathbf{Z_i^{*\prime}HZ_i^*})^{-1}\mathbf{Z^{*\prime}X}\}^{-1}\mathbf{X'Z^*}(\frac{1}{N}\sum_{i}^{N} \mathbf{Z_i^{*\prime}HZ_i^*})^{-1}\mathbf{Z^{*\prime}Y},$$

where **X** and **Y** are defined as in the AH case above, and **H** is a T-2 square matrix with twos and minus ones on the diagonals and sub-diagonals (zeroes everywhere else). $\mathbf{Z_i}^*$ is a block diagonal matrix, whose *s*th block is given by $(y_{i1}...y_{is}x_{i1}...x_{i(s+1)})$, for s=1,...,T-2. $\mathbf{Z^*}=(\mathbf{Z_1}^*,...\mathbf{Z_N}^*)$.[116] N is the number of cross-sectional units.

The GMM estimator above can use all the available lags of variables as instruments, or a subset of the available lags in a restricted GMM estimator (like the AH estimator). The choice of instruments is an act of balancing bias and efficiency, since there is a trade-off between adding efficiency but possibly worsening the bias by adding more instruments (Judson and Owen, 1997). In short, although the efficiency of the AH estimator in small panels can be improved upon by using a less restricted GMM estimator, this may come at a cost in terms of bias reduction. Indeed, Judson and Owen (1997) show that GMM estimators with 2 and 5 lags used as instruments give a bias only slightly smaller, and in some cases larger, than standard, uncorrected LSDV, for data sets with similar dimensions to the one here.

[115] The estimator presented is the 1-step GMM estimator (GMM1 in Judson and Owen, 1997). The related 2-step estimator is not presented here, and is found by Judson and Owen (1997) to be outperformed by the 1-step estimator.

[116] For more details of these definitions, see Judson and Owen (1997).

The instruments suggested in the above discussion of the AH and GMM estimators are only valid if the errors themselves are not serially correlated (Baltagi, 1997). The validity of these instruments rests on their orthogonality to the error term. If u_{it} is correlated with u_{it-1} and so on, then y_{it-2} will be correlated with (u_{it}-u_{it-1}) and therefore not a valid instrument.[117] In order to be satisfied that such instruments can be used in the estimation, it is therefore necessary to test for serial correlation in the errors. In Appendix B, evidence of first order serial correlation is presented, suggesting standard AH and GMM estimators may lead to inconsistent estimates.

Another problem with GMM/AH estimators, which is a distinct possibility in this case, is that they do not perform well where variables are persistent. This is because lagged levels will only be weakly correlated with subsequent first differences, leading to large finite-sample biases (Blundell and Bond, 1999). Given that R&D is widely accepted to be highly persistent, and that the IPS tests discussed in Section 4.7.1 cannot rule out unit roots (suggestive of a high degree of persistence), these estimators appear less attractive still, in this particular application.

In summary, estimation of dynamic panel data models is problematical because the LDV is correlated with the error term by construction. This results in a downward bias on the estimated coefficient of the LDV and a smaller bias, which can be of either direction, on the coefficients of the other variables, with LSDV estimation. This bias decreases as T (the time dimension) increases, although it may still be significant with T=20. However, despite this bias, coefficients will be of the correct sign and not too far from their true value, when estimated with LSDV. Since the bias is less severe on the coefficients of the other variables, we would expect the openness variable to be estimated with the correct sign and a reasonable degree of accuracy.

Alternative estimators based on the application of instrumental variables techniques exist which can improve the consistency (reduce the bias) of the estimates. However, these estimators will suffer from inconsistency and finite-sample bias when errors are serially

[117] If the errors are correlated, estimation using the suggested instruments will be inconsistent (Baltagi, 1997).

correlated and the dependent variable is persistent. I therefore estimate (4.22) by LSDV, allowing for heteroskedasticity, cross-sectional correlation and serial correlation in the errors. Although this estimator is not unbiased, it should provide efficient estimates close to the true parameter values. I report the LSDV estimates of (4.22) as it stands (ignoring the likely endogeneity of investment)[118] and also replacing the contemporaneous investment variable with the first lag.

I test for the inclusion of a time trend in the LSDV estimation, to check any apparent relationships are not merely picking up the action of time (or omitted trended variables). This is positive and significant for all sectors, so is included in the specification of (4.22) for the estimations the results of which are given in the tables below. I do not report the coefficient on the trend, since it is of little interest in itself. I also test for the inclusion of growth. It is difficult to carry out further tests of the robustness of the results given that I am already using all the available data. However, the fact that (4.22) is estimated separately over 13 sectors is a good test of robustness in itself. I also estimate (4.22) in first differences by LSDV, given the difficulty in ruling out the presence of unit roots.[119]

4.8: Results and Discussion

I carry out LSDV estimation of equation (4.22) by first demeaning the data by the Within transformation to get rid of the fixed effects dummies and then estimating the demeaned equation by 2-step FGLS as outlined in Section 4.7.[120] In this way, I control for groupwise heteroskedasticity, cross-sectional correlation and first-order correlation, all of which are found to be present in all 13 sectors. For further details of these tests, see Appendix B. The results of this estimation for the 13 sectors are given below as LSDV1 in Table 4.4. The LSDV estimator is also applied to an alternative specification of equation (4.22), where investment (RI) is replaced by the first lag of investment (RI[-1]),

[118] This, if investment is truly endogenous, as suggested by the testing procedure and by *a priori* reasoning, may also bias the results in an uncertain direction.
[119] The time trend disappears with the differencing transformation.
[120] I also carry out AH and restricted GMM2 estimation of (4.22) as discussed in the previous section. Unsurprisingly, given the serial correlation and persistence of R&D, the results are very poor, so are not reported.

for all sectors, given below as LSDV2 in Table 4.5. Both estimators are carried out in the same way, apart from this one change, with the LIMDEP package.[121]

All the estimators display poor results for Sector 9 (petrol refineries and products). This is perhaps to be expected given the small size of the panel for this sector; it is the smallest with NT=6*18. In the analysis that follows, I largely discount this sector.

Table 4.4: LSDV1 Estimation of (4.22)

Sec	**RRD[-1]**	**OP**	**RPR**	**RI**	**PROF**	**GR**
1	.89 (.02)**	.08 (.03)**	.09 (.05)*	-.04 (.02)**	.04 (.01)**	.33 (.10)**
2	.86 (.01)**	-.02 (.01)**	.40 (.04)**	-.07 (.02)**	-.03 (.01)**	.00002*
3	.76 (.02)**	.22 (.04)**	.42 (.08)**	.04 (.03)	-.08 (.02)**	-.93 (.24)**
4	.76 (.03)**	.08 (.03)**	.28 (.05)**	.09 (.01)**	-.09 (.02)**	.36 (.20)*
5	.70 (.03)**	.24 (.05)**	.53 (.04)**	-.07 (.01)**	-.06 (.01)**	.36 (.19)*
6	.89 (.01)**	.08 (.02)**	.16 (.02)**	-.02 (.01)*	.05 (.01)**	.55 (.09)**
7	.72 (.03)**	.17 (.04)**	.24 (.04)**	.03 (.02)	.06 (.02)**	.74 (.21)**
8	.71 (.03)**	.10 (.03)**	.41 (.05)**	.02 (.02)	.01 (.02)	.19 (.18)
9	.27 (.04)**	.16 (.11)	.68 (.08)**	.11 (.06)*	-.04 (.02)	2.13 (1.19)*
10	.66 (.04)**	.13 (.03)**	.21 (.05)**	.07 (.02)**	-.02 (.01)	.27 (.24)
11	.69 (.03)**	.21 (.03)**	.28 (.04)**	-.02 (.01)	.01 (.01)*	.85 (.26)**
12	.62 (.02)**	.22 (.03)**	.35 (.03)**	.01 (.01)	.004 (.005)	.32 (.12)**
13	.69 (.05)**	.07 (.04)*	-.02 (.07)	.10 (.04)**		1.03 (.57)*

Notes: Standard errors are given in parentheses. A double star denotes significance at 5% and a single star denotes significance at 10%. Figures can be read as short run elasticities (all variables are in logs) but not long run elasticities. The profitability variable is undefined for Sector 13 (aircraft) because of non-positive values in levels. Estimation is robust to groupwise heteroskedasticity, cross-sectional correlation and first order autocorrelation.

The LSDV1 estimator appears to perform reasonably well. The coefficients on the LDV are always around a plausible magnitude and always positive and significant. The range for the coefficient on the LDV is .62-.89, always less than 1, but suggesting a quite high degree of persistence nonetheless. Openness to trade is significant in all cases, and with

[121] I use LIMDEP 6.0, using the TSCS (time-series cross-section command). This package tests automatically for groupwise heteroskedasticity, cross-sectional correlation, and informally for autocorrelation. The Within transformation can be carried out straightforwardly from inside the program using the Xbr command. For further details, see Greene (1991).

the exception of food, beverages and tobacco, is always positive. Production is positive and significant in all but one case. The other variables display more mixed patterns, with some significance and some contrasts in terms of sign. The effect of growth is generally positive however.

Table 4.5: LSDV2 Estimation of (4.22)

Sec	RRD[-1]	OP	RPR	RI[-1]	PROF	GR
1	.89 (.02)**	.07 (.03)**	.09 (.04)**	-.03 (.02)**	.03 (.01)**	.21 (.11)*
2	.86 (.01)**	-.02 (.01)**	.37 (.05)**	-.05 (.02)**	-.03 (.01)**	.00002**
3	.77 (.04)**	.35 (.10)**	.55 (.20)**	.0001 (.07)	-.07 (.07)	-.90 (.88)
4	.76 (.03)**	.09 (.03)**	.32 (.05)**	.05 (.01)**	-.07 (.02)**	.12 (.21)
5	.71 (.25)**	.29 (.05)**	.49 (.04)**	-.05 (.01)**	-.05 (.01)**	.49 (.18)**
6	.87 (.01)**	.09 (.02)**	.17 (.02)**	.01 (.01)	.04 (.01)**	.54 (.10)**
7	.73 (.03)**	.16 (.04)**	.24 (.04)**	.02 (.02)	.06 (.02)**	.76 (.20)**
8	.71 (.03)**	.10 (.03)**	.41 (.05)**	.02 (.02)	.01 (.02)	.19 (.18)
9	.40 (.05)**	.05 (.11)	.63 (.07)**	-.18 (.04)**	-.04 (.03)	3.14 (1.18)**
10	.65 (.04)**	.17 (.03)**	.32 (.05)**	-.02 (.02)	-.03 (.01)**	.28 (.26)
11	.70 (.03)**	.20 (.03)**	.27 (.04)**	-.003 (.01)	.01 (.01)**	.85 (.27)**
12	.62 (.02)**	.21 (.03)**	.34 (.02)**	.02 (.01)**	.01 (.005)	.33 (.08)**
13	.73 (.05)**	.08 (.04)*	.10 (.07)	-.03 (.04)		.68 (.57)

See notes to Table 4.4.

The LSDV2 estimator shows a pattern of results very similar to that for LSDV1. There are some slight changes in magnitudes and significance, but no significant sign changes. Unsurprisingly, the estimates that change the most are for the investment variable, although these too display no significant sign changes. This does not mean that the investment variable is not particularly endogenous; it could be that the lagged investment variable is similarly endogenous to the current investment variable (particularly given the possible persistence of the series). However, since the primary interest of this Chapter is the relationships of R&D with other explanatory variables, the general robustness to this specification change is encouraging.

Despite the plausibility of the results presented in Tables 4.4 and 4.5, we have to sound a note of caution on the possibility of spurious regression. I have not been able to rule out the possibility of a unit root in the series for R&D and production, and also perhaps for investment. The high t-ratios on the LDV and in some cases on the production variable look suspiciously high. Even though we should expect significant LDV effects because of the widely accepted persistence of R&D, and significant size effects just from proportionality, we have to treat t-ratios like this with a certain amount of scepticism. As a first step to testing the robustness of these results to the possible presence of unit roots, I estimate (4.22) in first differences by LSDV.[122] The results of this estimation are presented in Table 4.6.

Table 4.6 below shows the results from LSDV estimation of (4.22) in first differences. The group dummies and the trend drop out with differencing, although there is still significant groupwise heteroskedasticity and cross-section correlation.[123]

Table 4.6: LSDV Estimation on First Differences

Sec	DRRD[-1]	DOP	DRPR	DRI	DPROF	DGR
1	.42 (.05)**	.02 (.08)	.22 (.09)**	.02 (.03)	-.02 (.02)	.27 (.14)*
2	.27 (.05)**	-.02 (.01)	.18 (.05)**	-.02 (.02)	-.02 (.01)**	.10 (.10)
3	.31 (.05)**	-.19 (.08)**	.72 (.12)**	.02 (.03)	-.15 (.03)**	-.97 (.23)**
4	.26 (.06)**	-.03 (.05)	.31 (.06)**	.06 (.02)**	-.08 (.02)**	-.48 (.18)**
5	.28 (.05)**	.17 (.08)**	.64 (.06)**	-.03 (.02)*	-.05 (.02)**	-.24 (.18)
6	.39 (.05)**	-.01 (.04)	.15 (.03)**	-.01 (.01)	.004 (.02)	.47 (.09)**
7	.35 (.05)**	.02 (.07)	.28 (.06)**	.02 (.02)	.03 (.03)	.45 (.15)**
8	.28 (.07)**	.04 (.06)	.46 (.11)**	.001 (.02)	-.06 (.03)**	-.18 (.23)
10	.07 (.06)	.15 (.06)**	.35 (.06)**	.06 (.02)**	-.02 (.01)	.12 (.18)
11	.47 (.03)**	.35 (.04)**	.54 (.04)**	-.01 (.01)	.001 (.005)	.54 (21)**
12	.25 (.04)**	.21 (.04)**	.44 (.03)**	.04 (.01)**	-.02 (.01)**	.58 (.13)**
13	.46 (.09)**	.03 (.05)	.01 (.10)	.09 (.05)*		1.29 (.61)**

[122] The group dummies and trend disappear, but I still allow for the possibility of groupwise heteroskedasticity and cross-sectional correlation.

[123] I do not report the test statistics for this case. The tests are the same as for the LSDV1 specification and all fail to reject groupwise heteroskedasticity and cross-sectional correlation except for Sectors 8 and 13, where there is no evidence of cross-sectional correlation.

Notes: See notes for Table 4.4. Breusch-Godfrey tests rule out serial correlation in the first differences (see Appendix B). The investment variable is contemporaneous. The regression contains an unreported constant term.

Comparing Table 4.6 with Table 4.4 shows a number of broad similarities. Firstly, the coefficient on the LDV is always positive and significant, with the exception of Sector 10 only. Secondly, although the significance of the openness effect is down overall, there is still evidence of a significant positive relationship in 4 sectors (one sector is again negative). Thirdly, the production variable displays consistently positive and significant effects. Finally, the other variables again display a mixed pattern of significance and signs. There are also a number of differences between the two sets of results. As mentioned above, openness is significant in a minority of sectors in the first-differences specification compared to nearly all of the sectors in the standard specification. Of course, differencing loses a lot of information contained in the levels of the variables, which might explain this. It is possible that the explanation is less benign, however, and that some spurious regression is being uncovered. Secondly, the coefficients on the LDV are considerably smaller in size. The t-ratios are still large, however. In general, the estimation of (4.22) appears *reasonably* robust to specification in levels or differences. Further research is needed to explore these issues in more depth.

4.8.1: Discussion

Consider the LDV and its effects on contemporaneous R&D expenditure. In all sectors, the LDV is significant and positive, with all estimators. The coefficients tend to be quite high, between .62 and .89, generally speaking, suggesting substantial persistence in R&D expenditure across time.[124] The least R&D-persistent sector is Sector 12 (motor vehicles). The most persistent sectors are Sector 1 and Sector 6 (total manufacturing and chemicals). The results from the first-difference specification of (4.22) suggest the persistence of R&D falls below this range, with coefficients ranging from .07-.47. Nonetheless, the positive significance of the LDV is consistent across sectors and robust.

[124] The exception to this is the LSDV estimates for Sector 9, which appears poorly specified. This sector suffers from a very small sample (108 observations in total, 96 observations used in estimation). Nonetheless, coefficients are positive for this sector, for both LSDV1 and LSDV2, and significant.

As discussed in Section 4.5, there are clear reasons why we should expect R&D expenditure to be persistent in this way. Around 90% of R&D expenditure goes on skilled labour and lab-equipment, which is likely to be difficult to adjust in the short term (Nickell and Nicolitsas, 1997). The above estimates suggest most of this skilled labour and lab-equipment is unadjustable from one year to the next. The literature has generally found similar results for the persistence of R&D expenditure. For example, Nickell and Nicolitsas (1997) estimate a significant coefficient of 0.7 on log R&D lagged one year for the UK. Lach and Schankerman (1989) suggest a coefficient of 1.19 on lagged R&D in a VAR model of R&D and investment. Other related evidence also exists in the literature for the persistence of R&D. For example, in a firm level study, Geroski et al (1997) find that innovation is generally driven by a few firms who innovate persistently.

The presence, size and significance of the LDV has implications for the interpretation of the other variables (see Section 4.5 for a discussion). Reading off the coefficients gives us a short run elasticity, or in other words, the effects of explanatory variables on R&D expenditure in a given year. Whilst these coefficients are of interest in themselves, the long run effects of these variables can also be calculated with a simple transformation. That these variables have long run effects can be seen by repeated substitution of the LDV in (4.22). A change in one of the other explanatory variables will have a long run effect that builds up over time. This long run elasticity is given by the coefficient/(1-coefficient on the LDV), or $\beta/(1-\gamma)$.

Turn now to the question of primary interest in this paper. Is there a relationship between R&D expenditure and openness to trade? The LSDV estimates show a clear pattern of mostly positive coefficients, ranging between -0.02 (Sector 2, the one negative coefficient) to 0.24 (or .35 including LSDV2). Of these coefficients, all but Sector 9 are significant at the 10% level and most at the 5% level. In the first-difference specification, 5 of the sectors display significant coefficients, with one negative, and the coefficients are generally of a similar magnitude. That a positive and significant relationship exists between openness to trade and R&D expenditure in at least *some* industries appears to be a robust conclusion. That the relationship may also be negative in some cases cannot be ruled out.

Reading off the elasticities from the tables gives us the short run effects of openness on trade. For example, for total manufacturing (using LSDV1 estimates), a coefficient of 0.08 implies that, other things being equal, if openness increases by 1% in a year, R&D expenditure will increase by just under 0.1%. Of course, this increase in openness will then also affect next year's R&D expenditure through this year's R&D expenditure, and the next year and so on. The *long run* elasticity for total manufacturing is closer to 0.7. So, the 1% increase in openness in that year, will, in the long run, drive a 0.7% increase in R&D expenditure. Across the sectors, these long run elasticities of R&D to openness range from –0.14 (food) to 0.92 (textiles). Clearly openness to trade has a big effect on R&D expenditure both in the short run and the long run in a number of sectors. For example, in the total manufacturing sector, the various LSDV estimates explain between 0 and 30% of the world increase in real R&D expenditure over the sample period.[125] The long run elasticities from the first-difference specification of (4.22) fall within this range.

A great deal of literature exists suggesting a positive relationship between trade and growth, much of which draws on the argument that trade increases R&D incentives.[126] Evidence has been presented in this literature supporting the existence of such a positive relationship, but *direct* evidence remains scarce. The above results generally support the existence of such a trade/R&D incentives relationship, directly. By doing so, they support those models of growth and openness to trade, driven by R&D (for example, Rivera-Batiz and Romer, 1991a; Chapter 3 of this thesis). The results also make a contribution to the wider growth/trade literature, showing the existence of a mostly positive relationship between openness and R&D, which hints at a positive relationship between trade and growth.

There are many reasons why such a positive relationship might exist, which are outlined in Section 4.2 and discussed in Chapter 2. Unfortunately, due to data limitations I am not

[125] World (sample) R&D expenditure has increased by 110% over the sample period, with trade intensity up by an average of 46%.

able to disaggregate the estimated relationship into the individual driving forces identified. Rather, the results identify the existence of an *overall* relationship, generally positive, which could be driven by any or all of the forces previously discussed. There is, however, a scale variable in the equation, which might have the effect of capturing some of the potential scale effects due to increased trade. If all scale effects were captured by this size variable, we would be left with evidence of strong positive competition effects and strong positive duplication-reduction effects. Unfortunately, there is no way of knowing how the two variables interact in this respect, and it is unlikely that any scale effects are captured wholly by one or the other variable. However, if any trade-driven scale effect is captured by the size variable, then the true trade openness/R&D relationship may in fact be stronger than that suggested by the results here.[127]

Turning to the investment variable, the results are considerably less clear. LSDV1 is estimated on equation (4.22) with contemporaneous investment. LSDV2 is estimated on (4.22) with lagged investment replacing contemporaneous investment, because of the possibility of the investment variable being endogenous. There are some differences between the two estimators in terms of the investment coefficients, although in all other respects they are very close. LSDV1 suggests a significant positive relationship in 4 sectors and a negative relationship in 3. LSDV2 suggests a similar positive relationship in 2 sectors and a negative relationship in 4. Generally, the coefficients themselves are small, with a maximum of 0.1 (discounting sector 9). This corresponds to a maximum long run elasticity of 0.3 for the aircraft sector and a minimum of –0.2 for the paper and paper products. A similar pattern is displayed for the first-differenced specification.

The prior for this variable was for a positive relationship (see Section 4.5), and this has generally been found in previous studies. However, the results are so mixed as to give little support to the argument that R&D and investment spending come out of the same pot and therefore might be negatively related, or to the common shocks argument. In order to say more about the R&D/investment relationship, I would need a far richer econometric structure than the single dynamic equation. Previous studies have tended to

[126] See earlier sections of this Chapter.

[127] Equally, if there are negative scale effects, our R&D/openness relationship may be overestimated if the size variable captures some of this.

adopt multiple equation or VAR approaches (see, for example, Nickell and Nicolitsas, 1997; Lach and Schankerman, 1989). This is another possibility for further research.

As expected, the size variable enters with a positive and significant coefficient in all sectors under LSDV estimation, with the exception of the aircraft sector. This is true for all three alternative LSDV specifications, which suggests the results cannot be put down to spurious regression. Coefficients range from just 0 to 0.64. The existence of a positive relationship between scale and R&D expenditure is not surprising. A bigger industry will almost inevitably spend more on R&D than a smaller one, other things being equal. That is not to say that a bigger industry will be more R&D *intensive* than a smaller one. That is a much more open question, and one that has not been established either way in the literature. It is question that can be addressed by looking at the long run elasticities.

The range of long run elasticities of the size variable on R&D expenditure is from 0 to 2.9 (food).[128] For LSDV1, there is a roughly even split between sectors displaying a long run size elasticity below 1 (6 sectors) and those displaying an elasticity above 1. This mixed pattern is largely reflected in the LSDV2 and first-differenced estimates.[129] For a purely proportional relationship (bigger industry spends more on R&D, but is no more or less R&D intensive) we would expect a long run elasticity of around unity. Many sectors are close to this, such as total manufacturing, wood products, chemicals excluding drugs, office machinery and computers and motor vehicles. But some sectors show significant positive scale effects. In particular, food, textiles, paper and drugs and medicines all have long run size elasticities of greater than 1.5. So, for these industries, the results suggest that the larger an industrial sector in a given country, the more R&D intensive it is in the long run.

This mixed pattern of results does not allow us to question or support the validity of those models of R&D and growth built on scale effects. Of course, my measure of scale is not perfect for the purpose of proving or disproving scale effects. For example, the demand-led scale effect (bigger market therefore bigger potential reward to successful

[128] It is possible that this large scale effect for the food sector is partly responsible for the negative coefficient on openness to trade, by capturing positive openness-related scale effects.

innovation) is driven by scale in a *potential* sense and may not therefore be captured by production data. However, the differences across sectors suggest at least that we should look for scale effects at a disaggregated level as well as at aggregate level.[130]

The profitability variable displays a mixed pattern of estimates, with some positive significant effects (not robust to first-differencing), some negative significant estimates and many insignificant estimates. Perhaps the most interesting pattern in these results is an apparent low tech/high tech contrast. It is the *low-tech* sectors that display significant effects here. Do financial constraints matter in some way in low tech industries all the more than high tech industries, or is there something else being captured by this variable? Theory suggests, and other studies have found, the presence of a negative relationship between financial constraints at the firm level and R&D expenditure. This would suggest a positive relationship between profitability and R&D expenditure. However, at industry level, this effect may be much less well defined. The proxy for profitability is also not satisfactory. The capital stock estimates are not entirely convincing, for example. Unfortunately, there is no clear way to improve this variable with the current industry level data. Indeed, previous studies have tended to use firm level data, which are much more suitable to this sort of exercise (see, for example, Himmelberg and Petersen, 1994). Given these difficulties, I am unwilling to draw any firm conclusions here.

Finally, the growth of GDP variable enters with a generally positive and sometimes significant coefficient. The exception is sector 3 (textiles) which appears to show a counter-cyclical relationship between R&D expenditure and GDP. For the other sectors, short run elasticities range from between 0 and 1.29 (aircraft). The largely pro-cyclical nature of R&D is unsurprising, and is probably reflecting macroeconomic demand and expectations effects.

Separate estimation of sectors allows us to see whether there are any systematic differences in these relationships across sectors. For example, high technology and low technology industries may behave differently in a number of ways, or low-bulk and high-

---

[129] Although the majority display a long run size effect greater than one in the latter specification.

bulk product industries may behave differently to openness to trade. Of the sectors above, 2-5 and possibly 9-10 can be classified as low technology industries, with 1, 6-8, 11, 12 and 13 high or medium-tech industries (see Section 4.6 for a discussion). Classifying sector by the bulk of their products in transportation would give us a similar breakdown. However, there appears to be no clear contrast in terms of the openness/R&D relationship across sectors. In terms of scale effects, the contrast does not seem to be between high and low technology sectors, either. Of course, trying to conclude anything else about sectoral differences would perhaps be stretching the data set a little far.

4.9: Summary and Conclusions

A number of recent papers present models in which R&D incentives or outcomes are affected by the degree of international trade or more generally by integration between economies (see for example, Rivera-Batiz and Romer, 1991a, 1991b; Grossman and Helpman, 1990, 1991a, 1991b; Chapter 3 of this thesis). Although not exclusively so, this relationship is usually characterised as a positive one, with increased integration or increased trade leading to increased R&D incentives or faster technological progress. A number of alternative mechanisms through which this positive effect takes place have been suggested, such as spillovers of technical knowledge, market enlargement and changes in the intensity of competition. In some sense, it is almost conventional wisdom to believe that integration or increased trade will drive increased R&D. This belief lies behind many explanations of possible integration/growth effects, as found by Henrekson et al (1996), for example. However, until now, little direct evidence to support (or otherwise) these theoretical ideas suggesting a positive integration/R&D relationship has been presented in the literature.

In this Chapter I have looked for evidence of a relationship between R&D expenditure and openness to trade, using sectoral level data across a number of OECD countries between 1973 and 1992. A simple dynamic R&D expenditure equation, with openness to

[130] Backus et al (1992), which is perhaps the best known example of a search for scale effects, looks for

trade included as an explanatory variable, is estimated for each of thirteen sectors as a time series cross-section of countries. Estimation is complicated both by the small size of the 'panels' and by the presence of a lagged dependent variable. LSDV results are presented, which are efficient and for which the bias is known to be relatively small.

Despite the problems in estimating the model, I find evidence of a significant, positive relationship between R&D expenditure and openness to trade in a number of sectors. In one sector (food), there is evidence of a weak negative relationship. The strength of this positive trade/R&D relationship, in terms of long run elasticities, can be quite high, at up to 90%, on average. For example, in total manufacturing, up to a third of the increase in R&D expenditure over the sample period can be put down to the effects of increasing world openness to trade.

There is strong evidence for persistence in R&D expenditure from one year to the next, a fact found to be the case by a number of previous studies. Other explanatory variables largely paint a mixed picture. There is some evidence of positive scale effects on R&D expenditure for some industries. Equally, some industries' R&D expenditure is roughly proportional to size (no scale effects) and a few industries have R&D *intensity* decreasing with size. R&D is generally pro-cyclical. I can draw no strong conclusions about the relationship between R&D and investment and R&D and profitability from the results. This is perhaps unsurprising given the potential endogeneity problems with investment in a single equation model and given the nature of the profitability proxy. Ideally, a study of R&D and financial conditions needs to be carried out with firm level data.

Overall, these results go some way towards providing direct evidence in support of models of trade integration and R&D in the literature. The evidence of a positive relationship is quite strong, being generally robust across a large number of sectors and to some alternative specification of the model. Despite complications with estimation, my interpretation of the results is one of support for the hypothesis of a positive integration/R&D relationship in at least some sectors. At the very least, the results

evidence at country level.

suggest the need for further research to be carried out into these issues at a disaggregate level.

However, I cannot claim too much from this exercise. Rather, I look on this Chapter as a tentative first step towards finding evidence on the existence of this R&D/integration relationship. Further work will be needed to provide a more convincing case of its existence, as there are numerous areas of weakness in the current approach, which result in a degree of ambiguity in the results and conclusions that can be drawn. Perhaps the most important limiting factor in this study and many other studies of its kind is the nature of the available data. Although improved data is becoming available all the time, there is not yet a sufficiently detailed, sufficiently long (in the time dimension) and sufficiently wide data set to enable consistent and efficient estimation of a well-specified empirical model of R&D expenditure across countries. Large and detailed panel data sets do exist at firm level, however, and this may be the obvious place to look for better data in the near future.

Even given the available data, there are avenues for further research that should be explored. The difficulty in ruling out the presence of unit roots in the R&D, production and possibly investment series is problematical. I argue that the best approach is to firstly assume stationarity and then non-stationarity and compare results. This Chapter makes the former assumption, but the possibility of some spurious regression with the LSDV estimator gives extra urgency to the need for a non-stationary approach. It may also be that a single equation approach is not the best way to model this problem. Further work could adopt a VAR approach with fixed capital investment, trade flows and R&D expenditure jointly determined, for example. Taken together, these arguments suggest a panel-cointegration approach is the next step for this research.

Finally, a last word on how I see this Chapter in relation to the literature. By attempting to establish the empirical existence and nature of the integration/R&D relationship, and meeting with at least some measure of success, support is given to the models that hypothesise such a relationship. There is a clear way forward to further robustify these conclusions. Such research is key to providing direct empirical support for what are potentially crucial building blocks in our understanding of economic growth.

Chapter 5

Knowledge Spillovers, Trade and Foreign Direct Investment in UK Manufacturing

5.1: Introduction

The existence of knowledge spillovers has significant implications for economic growth by raising the social rate of return to research and development (R&D) above the private rate of return. Recently, numerous studies have found evidence of such spillovers at various levels of aggregation. These studies typically assume that technology is transmitted though trade in goods between countries, industries or firms. Alternative transmission mechanisms, such as foreign direct investment (FDI), have been widely suggested, but little research has so far been carried out into their effects.

In this Chapter, I examine knowledge spillovers in a panel of UK manufacturing industries, allowing for technological transmission through both direct investment and trade in goods. The results are somewhat counter-intuitive in that they show foreign technological progress to mostly have a *negative* effect on productivity in UK manufacturing, through both trade and FDI. This result is robust to a variety of specification changes. One possible explanation is that the UK's position as a leading technology source country makes it unlikely to gain from R&D in technologically inferior countries and makes it vulnerable to *outward* spillovers, which might act as a disincentive to invest in technological advancement. Alternatively, increasing competition with technologically superior economies, such as the US and West Germany, might have a negative effect on incentives to invest in R&D in the UK which is not linked to knowledge spillovers.

The remainder of this Chapter is set out as follows. The next section discusses the background to the study. Section 5.3 reviews the existing empirical literature on knowledge spillovers. Section 5.4 presents a model of how FDI and trade in goods can affect productivity with endogenous R&D and derives an estimable empirical model. Section 5.5 discusses the data and provides empirical background for UK manufacturing. Section 5.6 discusses the econometric issues arising from the estimation of the model. Sections 5.7 present and 5.8 discuss the results. Section 5.9 concludes. A data appendix is provided in Appendix C with detailed notes on the construction of the series used. A further appendix (Appendix D) presents statistical details and tests of robustness not reported in full in the main text.

5.2: Background

Recent models of endogenous growth have stressed the role of profit-driven innovation by entrepreneurs as the driving factor behind an economy's growth performance (see, for example, Romer, 1990). Innovation is enabled by the accumulation of technical and scientific knowledge through R&D, and feeds back into this stock of knowledge, thereby driving long-run productivity growth. Cumulative R&D is generally taken as the best approximation to this stock of knowledge. Strong evidence exists of a positive relationship between measures of productivity and cumulative R&D expenditures at firm level (see, for example, Bernstein, 1988; Hall & Mairesse, 1995), industry level (see, for example, Bernstein & Nadiri, 1988; Keller, 1997) and country level (see, for example, Coe & Helpman, 1995; Nadiri & Kim, 1996).

The key characteristics of this kind of technical knowledge are that it is, to an extent, non-rivalrous and non-excludable. In other words, it can be used in many applications and locations at the same time and you cannot prevent others from using your ideas without full monetary compensation (Grossman & Helpman, 1991a). So, not only does R&D have a potential effect on own productivity, but through its contribution to the wider stock of knowledge it also has a potential effect on productivity elsewhere. The diffusion of such knowledge from one body to another has become known as knowledge spillovers.

The existence of knowledge spillovers suggests that productivity gains can be made not only through costly own R&D, but also through exploiting the benefits of others' R&D. This possibility has enormous implications for policy. For example, if FDI into the UK delivers significant benefits in terms of knowledge spillovers to domestically owned firms, then we should encourage such investment, perhaps through subsidies. If this is not the case, then resources spent on encouraging such inward FDI may be better spent elsewhere. Clearly, we must understand the extent of, and sources of knowledge spillovers if we are to derive maximum long run benefit from them.

Taking a country-sector as the basic unit, knowledge spillovers may exist from other sectors within the country, from the same sector outside the country and from other sectors outside the country. In principle, technological progress from any of these sources may spill over and lead to improved productivity in our given sector. In practice, it may be that knowledge spillovers are more likely between bodies that are close in some sense, either geographically (see, for example, Krugman & Venables, 1995) or in technology-space (see, for example, Branstetter, 1996). The extent and direction of any spillovers will depend on the strength of those mechanisms through which knowledge may be transmitted and the relative technology levels across country-sectors.

The list of possible transmission mechanisms for technical knowledge is long. For example, knowledge may be transmitted from one body to another through migration of labour or through direct investment. However, researchers have tended to concentrate on trade in physical goods, both because of the existence of supporting theoretical models and because of the availability of quantifiable goods trade data. In short, if trade from one sector to another, or from one country to another increases, this can act to increase the transmission of knowledge from sector to sector or from country to country and therefore impact on productivity in the given country-sector. Grossman and Helpman (1991a) identify four channels through which such trade might affect productivity:

(1)A direct effect where a country-sector can increase the productivity of its own resources by employing intermediates from other country-sectors.

(2)Through the opening of communications channels between country-sectors which can lead to cross-border learning.

(3)Through copying or modifying external technologies discovered through external contacts.

(4)By raising the productivity of a country-sector in developing new technologies or imitating the technologies of others.

However, as suggested above, knowledge may spill over independently of trade in goods. This is recognised by most authors, although the difficulty in pinning down and quantifying other possible transmission mechanisms has not yet been overcome. It is, however, an unavoidable fact that, in order to satisfactorily examine the nature of spillovers we must correctly specify (at least) the most important mechanisms through which knowledge may be transmitted.

The direction of knowledge spillovers is likely to depend on the relative technology levels in different country-sectors. A sector with a substantial stock of technological knowledge is unlikely to learn much from a sector with little technical knowledge. It is more likely that knowledge spillovers will flow in the other direction (from high knowledge to low knowledge sectors). Some researchers have argued that this relationship is not straightforward in that the extent and direction of knowledge spillovers may not be linear in terms of the knowledge-gap (see, for example, Engelbrecht, 1997).

Differences in technology levels and the strength of transmission mechanisms suggest that the extent and direction of knowledge spillovers are likely to vary considerably across countries and across sectors. This supports the need for studies of knowledge spillovers to be carried out at as disaggregated a level as possible. Equally, the likely heterogeneity across countries and sectors implies that general conclusions cannot be drawn from studies limited to one country or one sector. In this Chapter, I examine how total factor productivity in a number of UK manufacturing industries is related to stocks of technical knowledge both domestically and outside the UK. In addition to the standard goods-trade transmission mechanism, I allow knowledge spillovers to be transmitted through FDI, both into and from the UK.

5.3: The Existing Empirical Literature on Knowledge Spillovers

5.3.1: Macro-Level Studies with Goods Trade

Probably the most widely cited paper on knowledge spillovers in recent years is Coe and Helpman (1995). Theirs was the first paper to attempt to measure knowledge spillovers, generally accepted to exist at the micro level, at the macroeconomic level. Building on Grossman and Helpman's model of technology embodied in tradable intermediates, they estimated an empirical model where total factor productivity (TFP) could depend on both the domestic stock of R&D and foreign R&D stocks, through bilateral imports for a panel of 21 OECD countries and Israel. They provided evidence not only of a high return to R&D domestically, but also of significant spillovers where domestic R&D stimulates the productivity of trading partners. Because of the seminal nature of this work, I briefly present their model and main findings below.

For a pooled panel of 21 OECD countries and Israel, Coe & Helpman estimate the following equation:

(5.1): $\log F_{it} = \alpha_i + \alpha_i^d \log S_{it}^d + \alpha_i^f \log S_{it}^f + \varepsilon_{it}$,

where i is the country index, F_{it} is TFP, S_{it}^d is the domestic R&D stock and S_{it}^f is the import-share weighted average of foreign R&D stocks. The elasticity of TFP to the domestic R&D stock is estimated at around 10%, with the elasticity to foreign R&D stocks around 9%. A G7 dummy multiplying the domestic term in equation (5.1) gives a G7 elasticity to domestic R&D of around 13% and slightly lowers the estimates for the remaining countries. They then argue that equation (5.1) above will not capture the *level* of imports since the weights add to 1. Therefore they estimate a further variation of (5.1) with the log weighted foreign R&D stock multiplied by the ratio of imports to GDP. This considerably increases their estimate of the elasticity of TFP to foreign R&D stocks to around 29%.

A number of studies followed on directly from Coe and Helpman's 1995 paper. Coe et al (1997) and Bayoumi et al (1996) extend the analysis to examine the extent of spillovers from the 22 industrial countries of Coe and Helpman's original sample to a number of developing countries. They find evidence of significant spillover effects to these developing countries through trade.

Nadiri and Kim (1996) place Coe and Helpman's study in a structural model allowing spillovers to affect a country's production structure through factor-bias effects. They also allow the parameters to vary across countries and find significant heterogeneity in the relative importance of domestic and foreign R&D across countries. For example, the elasticity of TFP with respect to domestic R&D is around six times that of TFP with respect to foreign R&D in the US. In Italy, however, the effects of international spillovers are around three times as big as the effects of domestic R&D in terms of driving productivity growth.

Engelbrecht (1997) also finds evidence of a wide diversity of productivity effects of foreign R&D, with technology source countries (eg: US, Canada, West Germany) displaying a negative relationship between domestic productivity and foreign spillovers through trade.[131] Countries with very low R&D stocks also display a far weaker (and sometimes negative) productivity effect from foreign R&D, which Engelbrecht speculates could reflect the need to have reached a certain threshold level in domestic capability to be able to benefit from technological advances abroad.[132]

5.3.2: Goods Trade at Firm and Industry Level

A long tradition of spillover studies exists at a more disaggregated level. Bernstein and Nadiri (1988), for example, find evidence of significant intra-industry spillovers, from firm to firm, within Canada. The extent of these spillovers varies from industry to industry, with social rates of return to R&D being higher than private rates of return by a factor of 1-3. Bernstein (1988) also finds evidence of significant inter-industry

[131] This may be reflecting the effects of spillovers on R&D incentives. Chapter 3 of this thesis shows how the presence of spillovers across countries can reduce the incentives to engage in R&D and how this disincentive effect gets stronger the more integrated the countries become.

[132] This explanation is consistent with the idea of appropriate technology discussed in Section 2.3.2.

spillovers in Canada, which also vary considerably across industries in terms of strength and direction of flow. Input/output flows of goods provide the transmission mechanism at industry level.

Branstetter (1996) analyses spillovers at firm level in the US and Japan. He finds spillovers to be primarily intra-national in scope for these two countries. There is no evidence of positive spillovers from Japanese firms to US firms (indeed the effect is sometimes negative) and little evidence of spillovers flowing in the opposite direction. In addition, he argues that the considerable technological heterogeneity at country level and even at 2-digit industry level calls for research into spillovers to be carried out at as disaggregate a level as possible. He also argues for the use of technology-proximity measures to capture the propensity of knowledge to spill over, in preference to the more standard goods trade measures. In support he refers us to Keller's random trade matrix which gives even higher TFP/R&D elasticities than the actual trade matrix (Keller's paper is discussed below). Finally, Branstetter provides us with a warning on the issue of mis-measurement spillovers; firms spanning several sectors may give us a false impression of inter-industry spillovers through trade.[133]

Keller (1997) examines industry-level spillovers across a panel of 8 OECD countries. The paper provides an important bridge between the macro studies of Coe and Helpman and others and the smaller scale studies such as those discussed above. Spillovers can exist between industries within each country, across the same industry in different countries and between different industries in different countries, and Keller modifies the Romer model to formally set out these alternative sources of spillovers. His non-linear empirical equation is derived directly from this model. In this Chapter, I will closely follow Keller's approach. It is therefore instructive to summarise his paper in a little more detail (see Box 5.2 below).

Foreign intra-industry spillovers occur through goods trade (imports) as in Coe and Helpman (1995). Domestic inter-industry spillovers occur through the input/output matrix (the US input/output matrix is used for all countries) or alternatively through a

[133] Equally, this may be the case across countries, which provides a strong argument for the inclusion of FDI in any studies of international spillovers. This is discussed in the review of Hejazi and Safarian (1998) in the text.

technology proximity matrix. Foreign inter-industry spillovers occur through imports and the import use input/output matrix. Keller estimates the following equation and a number of restricted versions of the equation (dropping foreign inter-industry spillovers etc):

$$(5.2):\ \log F_{ijt} = \mu_{ij} + \beta_1 \log(b_{ijt} + \beta_2 b_{ijt}^{io} + \beta_3 b_{ijt}^{f} + \beta_4 b_{ijt}^{f,io}) + \varepsilon_{ijt},$$

where i and j denote country and industry respectively and b_{ijt}, b_{ijt}^{io}, b_{ijt}^{f} and $b_{ijt}^{f,io}$ denote domestic own sector, domestic I/O-weighted outside sector, and foreign import-weighted within sector and outside sector R&D stock respectively. An alternative specification replaces the I/O matrix with a technology flow matrix estimated by Scherer (1984) and Evenson et al (1991).

Keller's results again show a positive R&D/productivity relationship within country-sector. Domestic inter-industry spillovers are around one fifth to one half as strong and foreign intra-industry spillovers are between half and equally as strong. He is unsuccessful in attempting to quantify technology flows from other sectors abroad. These estimates are from the technology flow matrix specification, which is found to perform considerably better than the pure goods flow specification.

A further important aspect of Keller's paper is the way he calls the goods trade approach to modelling spillovers into question, by applying a randomly created goods flow matrix to his model. This random matrix performs better, giving larger and more significant spillover estimates, than the actual goods flow data. This, he suggests, shows the need to incorporate alternative transmission mechanisms in future research.

5.3.3: Studies Introducing FDI

The need to include alternative transmission mechanisms over and above physical goods trade has recently led to a small number of papers which allow spillovers through FDI. Lichtenberg and Pottelsberghe de la Potterie (1996) extend the macro-level study of Coe and Helpman (1995) in this way, allowing technology to be transmitted through both outward FDI (technology sourcing) and inward FDI.

Although such FDI data is difficult to come by, they manage to construct a 4-year moving average series for each country. In addition to significant trade effects, they find a significant spillover effect through outward FDI, but surprisingly, no significant effect from inward FDI. They argue this shows how UK firms, for example, can invest in the US and gain access to US technology, but that technology does not spill over to UK firms from US affiliates in the UK, perhaps because of a tendency to concentrate R&D in the home country. In other words, the knowledge spillover effects of FDI flow to the people who *send* it and not to those who *receive* it. An additional innovation of this paper is that they move away from using index numbers, as in Coe and Helpman (1995). They argue that this is a mis-specification of the equation including the import/GDP ratio. This change in specification gives an insignificant foreign spillover estimate.

Hejazi and Safarian (1998) build on Lichtenberg and Pottelsberghe de le Potterie's paper, by obtaining annual bilateral FDI data, albeit for a smaller sample of 6 G7 countries (minus France). Broadly following the Coe and Helpman methodology, they find significant effects of both inward and outward FDI. They argue that this result is more appealing intuitively, and more in line with the micro literature on FDI, than Lichtenberg and Pottelsberghe de le Potterie's result where inward FDI is not a significant source of spillovers. In addition, they find that when FDI is included in this way, the importance of goods trade is considerably diminished as a spillover mechanism. Given this, and the fact that total spillovers seem to increase when FDI is allowed, they argue that FDI is both complementary to trade as a spillover mechanism and substitutable for trade as a spillover mechanism. Studies that do not include FDI are therefore likely to both underestimate the extent of spillovers overall and to overestimate the importance of spillovers through trade in physical goods.

5.3.4: An Alternative Explanation

Bertschek (1995) provides an alternative explanation for the relationships implied by the above studies. He argues that imports and FDI may have positive effects on domestic productivity levels through their effects on domestic innovation incentives because of the increase in the intensity of competition. Both import share and FDI are found to have a positive effect on product and process innovation in Germany, which will not necessarily be picked up by German R&D stock. This argument suggests that regressions including import share and FDI share terms may over-estimate the importance of spillovers, a point also noted by Lichtenberg and Pottelsberghe de le Potterie (1996).

5.3.5: UK Studies

A number of recent papers provide arguments for, and evidence of spillovers from trade and FDI for the UK. Barrell and Pain (1997) provide a great deal of useful background analysis of FDI in the UK. They stress how the characteristics of the host country affect the type of investment it receives, with the UK tending to attract less research-intensive and capital-intensive affiliates than Germany, for example. They also point out that, nonetheless, labour productivity in US affiliates in the UK is around a third higher than domestically owned firms.[134] They go on to estimate a simple regression where growth in TFP is dependent on an exogenous trend and lagged inward FDI for the UK. They find significant effects of FDI on productivity growth in the UK, mainly in manufacturing, with an elasticity of around 26%. From this they deduce that around 30% of UK manufacturing productivity growth since 1985 is down to inward investment.

Cameron et al (1998) study UK sectoral level productivity and its relation to sectoral openness. Openness is measured in a number of ways, including openness to inward investment and trade, and tends to be positively correlated with labour productivity and TFP growth. It is not clear whether this reflects spillovers or intensity of competition effects. However, they do find that the rate of productivity growth is related to the US/UK technology gap, which may suggest the former is an important factor behind this apparent relationship.

Cameron (1995) examines the evidence of spillovers in a panel of 19 UK manufacturing industries from 1970 to 1992. These spillovers tend to be intra-national rather than international in scope, and different for different industries. Although knowledge spillovers are found to play a part in explaining TFP movements,

[134] The US is the largest single inward investor into the UK. See Section 5.5 for a discussion of UK data.

especially for higher technology industries, much of the variation in TFP over the sample period can be explained by other factors, such as de-unionization or the 'Thatcher effect'.

In summary, a great deal of recent literature has looked for evidence of spillovers across firms, industries and countries. These spillovers are generally found to be significant, although there appears to be a great deal of cross-sectional variation in their strength. Early estimates may have overstated the importance of trade in transmitting technology across industries or countries by omitting other possible transmission mechanisms. Recent studies introducing FDI have found it to be a significant transmitter of technological spillovers, partly over and above trade in physical goods, but also partly instead of trade in physical goods. Studies for the UK have suggested the presence of spillovers from inward FDI and the existence of a more general openness/productivity relationship that may reflect spillovers. To the best of my knowledge, there is no UK study of spillovers that explicitly analyses the role of trade and FDI together.

5.4: A Model of Trade, FDI and Endogenous Technological Progress

I set up an extension of Keller's model, originally based on Grossman and Helpman (1991) and Rivera-Batiz and Romer (1991), allowing for exogenous technical knowledge through FDI-driven spillovers to enter the production function in the 'A' term. Whilst this does not represent any significant step forward in terms of modelling technological progress (FDI is left as purely exogenous) it does allow the specification of a simple empirical model with potential spillovers from both trade in goods and from FDI. The presentation here is brief and readers are referred to Keller (1997) and the references above for a more detailed description.

Good z_j, which can be either an intermediate or consumption good, is produced in sector j with the following production function:

(5.3): $z_j = A_j \Psi^{\eta}_j L_j^{\alpha} d_j^{1-\alpha}$, where A_j is a constant, Ψ_j is a measure of technology embodied in FDI, L_j is the labour used in production of output, α and η are [0,1] parameters and d_j is a composite input of horizontally differentiated goods x of variety s:

$$(5.4): d_j = \{ \int_0^{nj(de)} x_j(s)^{1-\alpha} ds\}^{1/(1-\alpha)}$$

where n_j^{de} is the range of intermediates employed in sector j.

The range of intermediates produced in the sector is denoted by n_j^P, which is some function of cumulative R&D in sector j, given by $\Delta n_j^P = \phi_j^P$. The range (ie: the total number of varieties) of intermediates produced in sector j at time T is therefore given by:

$$(5.5): n_j^P(T) = \int_0^T \phi_j^P(t)\, dt = b_j(T),$$

that is, equal to the cumulative R&D resources at time T.[135]

Assume that producing one unit of intermediates requires the same amount of inputs as producing one unit of the consumption good. Therefore, the sector's capital stock is simply the foregone output of the consumption good (or the total amount of intermediates produced in the sector):

$$(5.6): k_j = \int_0^{nj(p)} x_j(s)\, ds,$$

In symmetric equilibrium, all intermediates x are produced at the same level. We can therefore write down:

(5.7): $k_j = n_j^P x_j$.

Rearranging for x and substituting into equation (5.4) gives output:

(5.8): $z_j = A_j \Psi_j^{\eta} (n_j^{de})^{\alpha} L_j^{\alpha} k_j^{1-\alpha}$.

Define TFP as $f_j = z_j / L_j^{\alpha} k_j^{1-\alpha}$, and taking logs and substituting equation (5.8) gives an expression for the log of TFP in sector j:

(5.9): $\log f_j = \log A_j + \eta \log \Psi_j + \alpha \log n_j^{de}$.

So TFP in sector j is a positive function of a sector-specific constant, (exogenous) sectoral FDI and the number of intermediates employed in the sector.

Following Keller (1997), we assume the range of intermediates *employed* in sector j in country i is given by a weighted sum of all the ranges of intermediates produced by all country-sectors. In other words, a weighted sum of other country-sector R&D stocks. The number of intermediates employed is therefore made up of the number of intermediates produced in the given country-sector and a weighted sum of the number of intermediates produced in the sector abroad, the number of intermediates produced in other domestic sectors and intermediates produced in other sectors abroad. The weights are given by import shares, input shares from input/output tables and import-use shares from the input/output tables. Therefore n_j^{de} can be written as:

(5.10): $n_j^{de} = b_j^d + \beta_1 b_j^f + \beta_2 b_{jio}^d + \beta_3 b_{jiu}^f$,

where b_j^d denotes domestic within-sector R&D stock, b_j^f denotes the import-weighted sum of foreign within-sector R&D stocks, b_{jio}^d denotes the input-weighted sum of domestic outside-sector R&D stocks and b_{jiu}^f denotes the import-use weighted sum of foreign outside-sector R&D stocks.[136] The betas are parameters capturing the strength

[135] The range of intermediates produced is the total number of varieties of intermediates produced by sector j at time T, which, given intermediates never become obsolete, is equal to the total cumulative number of varieties produced up to and including time T.
[136] These variables and their derivation are described in greater detail in the following section and Appendix C.

of spillovers from these three possible sources, where $\beta_i = 1$ denotes a technology effect of equal strength to that from domestic within-sector R&D stock.

From (5.9) and (5.10) we can write down an estimable TFP equation, which is the basis of the empirical model:

$$(5.11): \quad \log f_j = \mu_j + \eta \log \Psi_j + \alpha \log(b_j^d + \beta_1 b_j^f + \beta_2 b_{jio}^d + \beta_3 b_{jiu}^f) + \varepsilon_{jt}.$$

Equation (5.11) is estimated by non-linear least squares. The empirical model is unsurprisingly similar to that of Keller (1997) with the addition of the FDI term and restricted to the UK only.

5.5: The Data

The data are annual, for 8 broad UK manufacturing sectors between 1973 and 1992. I have data on transactions between these UK sectors themselves (input/output) and between the UK sectors and similar sectors abroad in 13 OECD countries, by country. R&D stocks are calculated for the 13 OECD countries along with the UK. Since there are a large number of series included in the analysis, many of which required non-trivial construction, a detailed Data Appendix (Appendix C) is included with full notes on sources, construction and descriptions of all the main series. Technical details on these series are therefore not presented in the main text.

5.5.1: TFP

The dependent variable is the log of TFP (expressed as an index, 1973=100) in the eight UK sectors. Figure 5.1 below shows log TFP for these sectors over the sample period. Productivity has generally risen over the sample period, although not smoothly. The 1980s were the most important period in terms of these rises, with most sectors displaying steady year on year increases in TFP. The 1970s and early 1990s, in contrast, display either static or falling TFP in most sectors. The sectors displaying the fastest growth in TFP over the period are mechanical engineering, chemicals and electrical engineering. The sectors displaying the slowest TFP growth are food and metal manufacturing. These patterns conform to standard classifications of sectors into high technology and low technology, with high tech industries experiencing more rapid TFP growth.

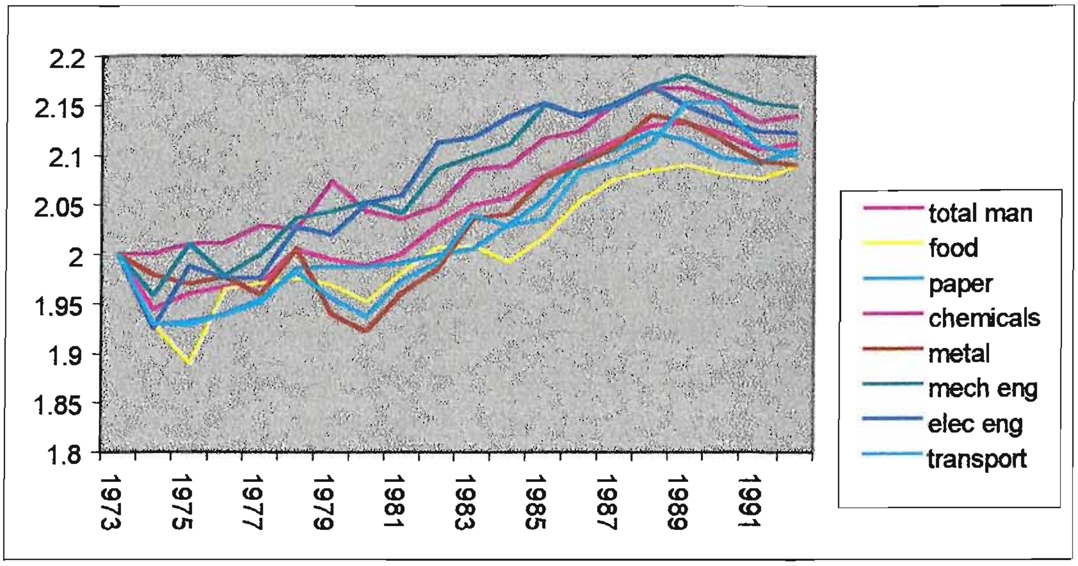
The pattern of TFP growth in UK manufacturing shown above reflects the widespread slowdown in productivity growth experienced in the 1970s in most advanced economies. Many commentators have put this down largely to the joint effects of oil price rises and the slowdown in the growth of R&D in the 1960s and early 1970s (see, for example, Griliches, 1988). The early 1980s saw a dramatic decline in the size of the UK manufacturing sector, with corresponding TFP growth as unproductive firms were either forced out or forced to improve their performance to survive in tighter markets.[137]

5.5.2: R&D Stocks

What of R&D in UK manufacturing? Figure 5.2 below shows log cumulative R&D expenditure for the eight UK manufacturing sectors. In general, R&D stocks have risen steadily over the sample period, although not rapidly. The most rapid growth has been in chemicals, electrical and mechanical engineering. No sectors show a fall in R&D stocks between 1973 and 1992. The correlation here with TFP growth is immediately apparent. Indeed, the pairwise correlation coefficient between TFP and R&D stock is positive, although quite small (see Table 5.2 below for panel correlations). At first glance therefore, there seems to be a positive relationship in the raw data between R&D stocks and TFP in UK manufacturing, albeit a weak one.

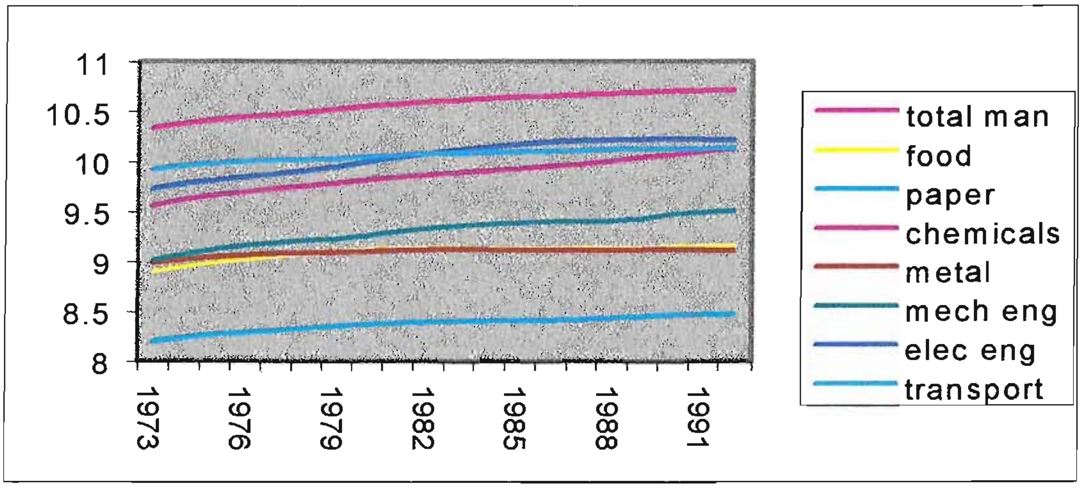
[137] An alternative to the standard empirical model (Equation (5.11)) includes a business cycle variable (log (GDPt/GDPt-1)) to control for such effects. See Section 5.7.

Figure 5.1: Log TFP in UK Manufacturing Sectors, 1973-1992.



Notes: Log TFP for 8 UK manufacturing sectors, 1973-1992. The TFP series shown are constructed as described in the Data Appendix, with time-invariant factor shares.

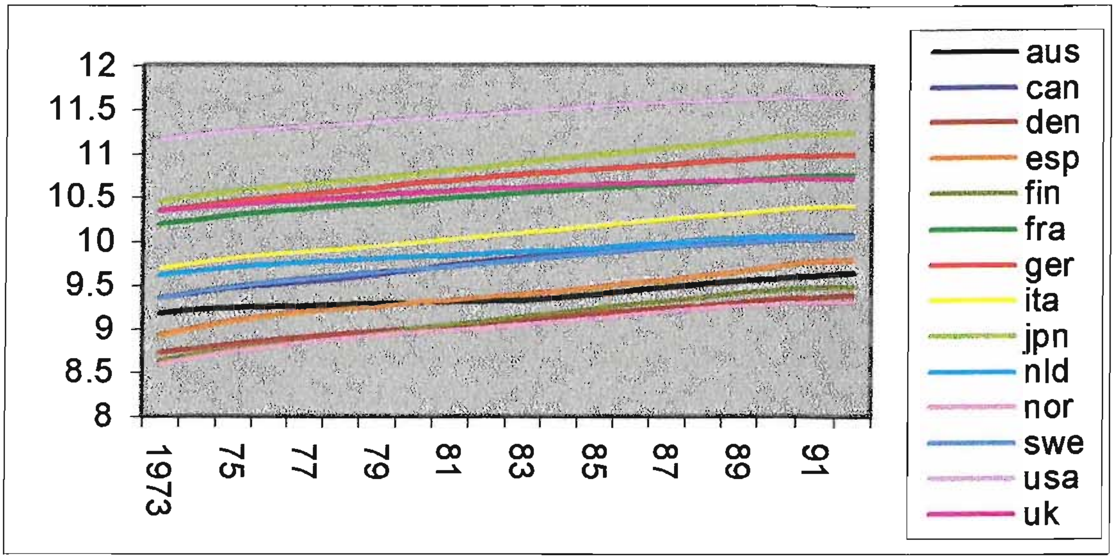
Figure 5.2: R&D Stocks in UK Manufacturing, 1973-1992.



Notes: The series shown are log R&D stocks for the 8 UK manufacturing sectors. Details on construction of the series are given in the Appendix C. The series shown above assume a depreciation rate of 10%.

How do UK R&D stocks compare to R&D stocks elsewhere? Figure 5.3 below shows log R&D stocks in total manufacturing for the 13 OECD countries in the sample plus the UK over the sample period.

Figure 5.3: R&D Stocks, Total Manufacturing, 14 OECD Countries, 1973-1992.



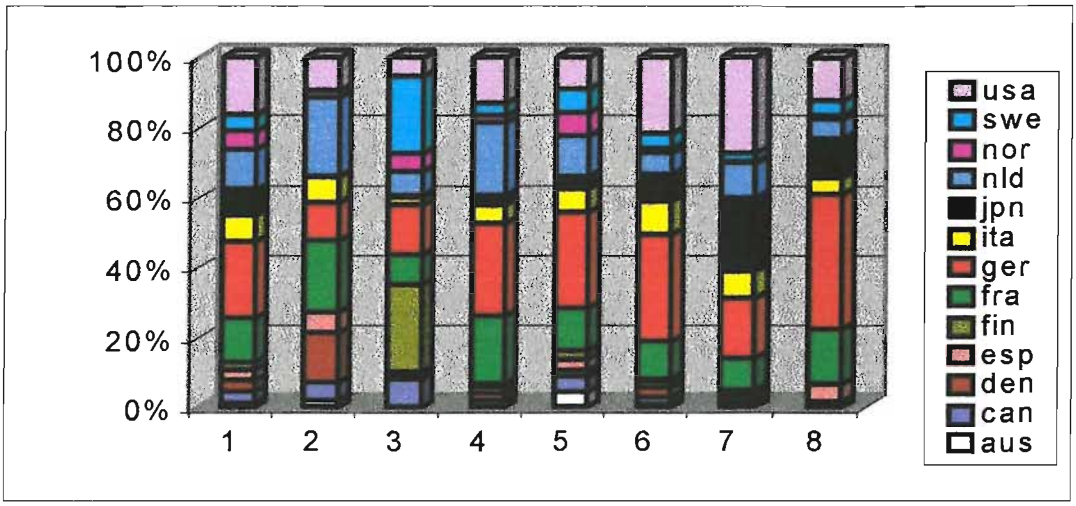
Notes: Series shown are log of total manufacturing R&D stocks for 14 OECD countries, 1973-1992. The series construction is described in Appendix C. The depreciation rate is taken as 10%.

It is clear from Figure 5.3 that R&D stocks in total manufacturing have been growing consistently over the sample period in all 14 OECD countries. The UK has the fourth largest R&D stock in 1973 and the fifth largest in 1992, having been overtaken by France. However, the UK is one of a group of three countries (UK, Netherlands, Australia) that displays the slowest growth rate of R&D stock in manufacturing. In the first three years of the 1990s, R&D stock in UK manufacturing has been static. Although the UK remains the world's fifth most important technology source country in manufacturing, its decreasing share in the world stock of technical knowledge may suggest a growing potential role for inward international spillovers over recent years.

5.5.3: Imports

Import shares vary considerably across sectors and across time. Figure 5.4 below shows the average import shares over the sample period (time-invariant import shares) of the 13 sampled OECD countries for each of the eight UK sectors. In general, the major sources of imports for the UK are Germany and the USA. Given the position of these two countries in the world's top three technological leaders, this suggests the possibility of the UK benefiting from knowledge spillovers through imports. Overall, however, the UK imports a considerable proportion of its manufactures from countries with a lower level of technical knowledge. Total manufacturing imports, for example, are split about 50/50 between higher and lower technology countries.

Figure 5.4: Average UK Import-Shares, Sectors 1-8, 1973-1992.

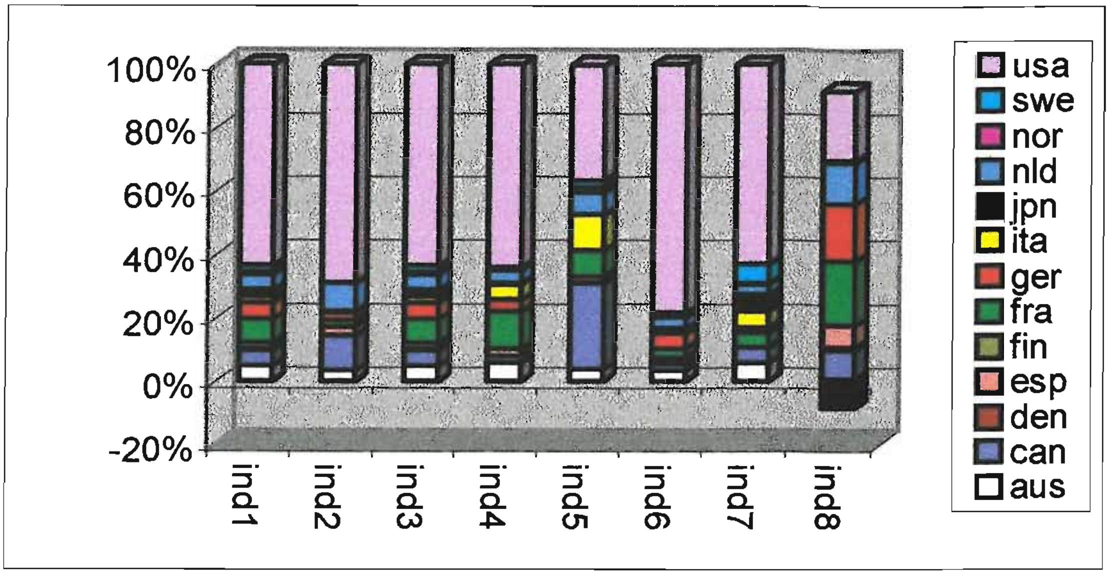


Notes: Average import-shares of the 13 OECD countries, 1973-1992, for the eight UK manufacturing sectors. Construction notes are presented in Appendix C.

5.5.4: FDI

Shares of (total) FDI are shown in Figure 5.5 below. Clearly the biggest source of, and destination for FDI is the US. In all industries except metals and transport, investment from and to the US makes up a majority of total UK FDI. The next biggest contributors to UK FDI are France and Germany. The clear majority of FDI therefore comes from, or is directed at, countries with a higher technology level than the UK, which suggests the possibility of positive spillover effects. Interestingly, the figures show that Japan has had a net dis-investment in transport over the sample period (see Appendix C for an explanation of how FDI might be negative in a given year).

Figure 5.5: Average FDI Shares, UK Manufacturing, 1973-1992.



Notes: Figures shown are for time-invariant total FDI shares in UK manufacturing sectors (inward + outward). Details of construction and sources are contained in Appendix C.

5.5.5: Variable Definitions and Correlations

The main variables used in the econometric analysis are listed in Table 5.1 below. Table 5.2 shows the correlation matrix of these variables (for the panel of all the industries).

Table 5.1: Variables in the Estimation, Including Alternative Specifications

Variable	Name	Notes
TFP	**LTFP**	Time-invariant and time-varying
Domestic R&D stock within sector	$\mathbf{LRD_j^d}$	$\delta = 0.1, 0.2$.
Domestic I/O weighted outside sector R&D stock	$\mathbf{LRD_i^d}$	$\delta = 0.1, 0.2$. Time-invariant input shares.
Foreign import-weighted R&D stock, within sector	$\mathbf{LRD_j^f}$	$\delta = 0.1, 0.2$. Time invariant and time-varying import shares.
FDI-weighted foreign within sector R&D stock	$\mathbf{LRD_j^{f,fdi}}$	$\delta = 0.1, 0.2$. Time-varying and time invariant FDI shares. Variables for total, inward and outward FDI.
Foreign outside-sector import-use weighted R&D stock	$\mathbf{LRD_i^f}$	$\delta = 0.1, 0.2$. Time invariant import-use shares. Time invariant import shares.

Notes: Construction and source details of all above variables are contained in Appendix C. All variables exist for all sectors, with the exception of outside-sector domestic and foreign R&D stock variables, which are not specified for total manufacturing.

Table 5.2: Correlation Matrix of Key Variables.

	LTFP	LRD_j^d	LRD_i^d	LRD_j^f	$LRD_j^{f,fdi}$	LRD_i^f
LTFP	1	.27	.29	.33	.31	.17
LRD_j^d	.27	1	-.27	.96	.95	.95
LRD_i^d	.29	-.27	1	-.12	-.17	.91
LRD_j^f	.33	.96	-.12	1	.98	-.14
$LRD_j^{f,fdi}$	.31	.95	-.17	.98	1	-.20
LRD_i^f	.17	.95	.91	-.14	-.20	1

Notes: Correlation coefficients shown are simple pairwise correlations for the key variables as defined in Table 5.1. Depreciation rates are set at 0.1 in all cases. Time-invariant figures are used wherever there is a choice between time-invariant and time-varying figures. The correlations are calculated across industries 2-8 for variables LRD_i^d and LRD_i^f, ie: excluding total manufacturing, for which no outside sector figures are specified. All other correlations are for all sectors, 1-8.

Log TFP is positively correlated with all the explanatory variables, as expected. However, the correlation coefficients are not large, suggesting the relationships are not particularly strong. The weakest correlation is between TFP and foreign outside sector R&D stocks, also as expected. The other main conclusion to draw from Table 5.2 is that many of the R&D stock variables are highly correlated with each other. For example, domestic within-sector R&D stock is highly correlated with foreign import-weighted and FDI-weighted foreign R&D stocks.[138] Surprisingly, domestic within-sector R&D stock is weakly negatively correlated with input-weighted domestic outside-sector R&D stock.

Before concluding this section, it is worth noting the limitations placed on such an exercise as this by the availability and accuracy of the data. For example, I am assuming that inward FDI in a particular sector originates from the same sector abroad and vice versa. In other words, the possibility of FDI *across* sectors is not accounted for in the data. Also, the aggregation to the 8 sectors is often difficult, with much of the data specified for different sectors or for commodities. The prevalence of missing or outlying observations, in the FDI data in particular, also limits us in carrying out empirical analysis of these issues.

[138] In Section 5.7.3, FDI-weighted R&D stock is dropped from the equation to examine the effects on the performance of the model in the light of this multicollinearity.

5.5.6: Stylized Facts

To conclude this section, I draw together the different strands of the analysis into a list of regularities in the raw data, which I call stylized facts. I identify 6 such facts, listed A-F below. The UK is used as a general term for the 8 manufacturing sectors covered by the data.

A) UK TFP displays moderate growth over the sample period, particularly in the 1980s.
B) UK R&D stock grows over the period, but at a slower rate than most other OECD countries.
C) UK imports are split between countries with higher technology levels than the UK and countries with lower technology levels.
D) UK FDI is largely from (to) higher technology countries.
E) UK TFP is positively correlated with both domestic and foreign R&D stocks.
F) R&D stocks across industries and countries are in many cases highly correlated.

5.6: Econometric Considerations

5.6.1: Stationarity

The approach adopted to estimating equation (5.11) depends on whether the series are believed to be trend-stationary (I{0}) or difference stationary (I{1}). Well-known procedures exist for establishing the stationarity or otherwise of such series (see Dickey & Fuller, 1979, 1981). However, these procedures (unit root tests) are often very inconclusive for short series because of low power (see, for example, Campbell & Perron, 1991). In other words, it is very difficult to reject the null hypothesis of a unit root even if it is not true.

However, recent work (see, for example, Quah, 1994; Im et al, 1996) finds that the performance of standard Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) tests can be improved upon if one exploits the cross-section dimension of the data. Of course, with limited observation time-series such as in the current study, with only 20 annual observations, these (improved) tests are still beset by problems of low power. Nonetheless, I carry out these tests following the t-bar test procedure set out by Im et al (1996). These tests, and issues relating to the correct specification for the DF and ADF regressions, are discussed in Chapter 4 of this thesis (Section 4.6). My treatment here will therefore be relatively brief.

Preliminary analyses of the pooled series, by regressing current values of each variable on an intercept, time trend and lagged values of the variable, suggest a mixed picture as regards the presence of significant time trends in the series. These analyses are also used to estimate the autocorrelation structures of the individual series in order to help specify the correct order for ADF regressions. Although it is difficult to specify a precise order from these regressions, the evidence points to lag structures of between 1 and 4 years for the series. If in doubt, applied econometricians usually suggest over-fitting rather than under-fitting (Im et al, 1996). I therefore carry out ADF tests of orders two, three and four, both including and excluding trends.[139]

The results of the Im et al (1996) t-bar tests are displayed below. Means and variances for the t-bar statistics are given in Im et al (1996). The standardised t-bar statistic is distributed standard normal under the null hypothesis of a unit root. Significant test statistics (rejecting the null of a unit root at a 10% significance level) are marked with an asterisk.

Table 5.3: Unit Root Tests

	ADF2	**ADF2t**	**ADF3**	**ADF3t**	**ADF4**	**ADF4t**
LTFP	3.29	3.88	3.09	3.33	2.69	2.51
LRD_j^d	-4.93*	2.10	-2.76*	-1.75*	2.60	3.38
LRD_j^f	-6.07*	7.59	-2.96*	.22	2.85	4.18
$LRD_j^{f,fdi}$	-5.7*	2.10	-1.70*	.16	2.81	3.92
LRD_i^d	-6.19*	8.48	-4.67*	-.57	2.86	4.16
LRD_i^f	-4.90*	5.99	-2.05*	1.64	2.86	4.04

Notes: The 10% critical value of the standard normal distribution is –1.64. ADFp is the ADF test of order p without time trend. ADFpt is the ADF test of order p with time trend.

[139] The ADF regressions also include an intercept.

There are clear differences between the results of the alternative specifications of the tests. To a certain extent this is to be expected. Where the series are trended, the tests without trends are incorrectly specified. This results in reduced power. Where the series are not trended then the trended tests will also suffer from reduced power. Under-fitting the ADF regressions reduces the size of the test. In other words, if the true dynamic structure of the series is greater than order two, the ADF(2) tests will tend to over-reject the null. Over-fitting the ADF regressions results in reduced power.

All the 4th order tests accept the null of a unit root. With the exception of log TFP, for which all tests accept the null, the picture from the 2nd and 3rd order tests is mixed. 2nd and 3rd order tests without trends reject the null of a unit root for all series, whereas 2nd and 3rd order tests with trends mostly accept the null. The pattern of results is consistent with under-fitting reducing size and over-fitting reducing power. It is more difficult to decide whether the 2nd and 3rd order tests are over-rejecting the null or whether the 4th order tests are under-rejecting the null, however. Given the mixed picture from the preliminary analyses, it is likely that the answer is different for different variables. The contrast between the trended and non-trended test results suggests the trended tests may be under-rejecting the null. To conclude, the evidence on stationarity is mixed, although in many cases tests point towards possible rejection of the null hypothesis of a unit root for all series except log TFP. Of course, with a maximum power of around 10% for these tests, given the sample size, even the TFP series tests may be falsely accepting the null.

Given the ambiguity of the unit root tests above, the most sensible approach to estimating the model is to allow for both stationarity and non-stationarity. In this Chapter, however, I focus almost entirely on the stationary approach, leaving estimation of the model assuming non-stationary series for further research. This further research has the potential to provide an interesting applied comparison of stationary panel estimation and non-stationary panel estimation, the econometrics of which have recently begun to attract more attention in the literature (see, for example, Kao, 1999).[140]

[140] The previous lack of econometric literature on panel cointegration was one of the problems faced by Coe and Helpman (1995).

In support of the assumption of stationary series, assuming series to be I(0) as opposed to I(1) is supported by the fact that I(1) series must have a root *exactly* equal to 1. Even a root of .99 gives us a stationary series. When there is doubt as to the presence of a unit root in a series, leaning towards the assumption of stationarity is therefore more attractive. Secondly, and most importantly, this study builds closely upon that of Keller (1997), which assumes stationarity in similar series. Dropping the assumption of stationarity moves us away from Keller's model and the extension of it as described in Section 5.4. I also estimate the model (equation (5.11)) in first differences as a test of robustness, just to be on the safe side.

5.6.2: Endogeneity

It is necessary to establish the exogeneity or otherwise of the regressors in equation (5.11) before we can carry out consistent estimation. In general, non-linear least squares estimation will be biased and inconsistent if the RHS variables are endogenous (jointly determined within the model). If this is the case, these variables may need to be instrumented to obtain consistent estimates.

Although relatively straightforward tests exist for establishing exogeneity or endogeneity, working assumptions can be made based on *a priori* reasoning. All of the explanatory variables, with the exception of domestic within-sector R&D stock, can be treated as exogenous if we make the following assumptions:

1) that the UK is a small open economy (ie: the UK does not have a significant effect on the rest of the OECD) and,
2) that individual manufacturing sectors are small compared to manufacturing as a whole.

For example, assumption (1) would say that TFP in the UK food industry is unlikely to have a significant effect on R&D stocks in the US food industry. It is more reasonable to assume that US food industry R&D stock can be taken as being determined outside the model. Equally, assumption (2) would say that TFP in the UK food industry is unlikely to effect the R&D stock in the UK electrical engineering sector. Again, it is more reasonable to assume UK electrical engineering R&D stock is determined outside the model. Unfortunately, domestic within-sector R&D stocks cannot be assumed exogenous in this way.

However, *a priori* reasoning can be used to suggest the exogeneity (or at least predetermination) of within-sector R&D stocks when we consider the time dimension of the variables. R&D stock at time t consists of that part of the R&D stock at time t-1 that has not depreciated and R&D investment at time t-1. Clearly, therefore, productivity at time t would not be expected to have a causal effect on R&D stock that is determined prior to this. Of course, *expected* productivity might affect R&D stocks prior to time t, especially where R&D investment responds to perceived technological opportunity. However, of a significant number of recent empirical R&D papers in the literature, none include TFP as an explanatory variable (see Section 4.3 for a review of this literature). There are sufficient *a priori* grounds, therefore, to assume that domestic within-sector R&D stock can be treated as exogenous in the model. In addition to these arguments, I carry out a quick Wu-Hausman test of the exogeneity of this variable, the details of which are presented in Appendix D1. The results of this test support the assumption of exogeneity.

Despite *a priori* reasoning, on which I proceed, there is always the chance of some simultaneity in such a model. For example, it might be that country-sectors with relatively low initial productivity attract more imports or less FDI than other sectors. This possibility of some reverse causality is more likely to be a problem when sectors are estimated separately. In the panel, fixed effects control for initial productivity levels, which should reduce any such simultaneity.[141]

5.6.3: Non-Linear Least Squares Estimation of Equation (5.11)

Equation (5.11), as derived from the model outlined in Section 5.4, has a clear non-linear structure. The natural technique to estimate such an equation is non-linear least

[141] See Section 2.3.2 for a discussion of a similar point in the context of cross-country growth regressions.

squares (NLSQ), which I carry out on the pooled data.[142] In this, I follow the estimation approach of Keller (1997).

NLSQ is an iterative technique for the estimation of non-linear equations such as (5.11). It works by minimising the sum of squared residuals (SSR) through iterations, which is equivalent to maximising the likelihood function if the error term in the equation is additive and normally distributed (see Hall and Cummins, 1997, for a brief but clear discussion of the iterative technique behind NLSQ). A more formal exposition is given in Greene (1997), of which the following is a brief summary, as applied to the empirical model.

For equation (5.11), the values of the parameters that minimise the sum of squared residuals,

$$S(\beta) = .5\Sigma_i\Sigma_t\varepsilon_{it} = .5\Sigma_i\Sigma_t[y_{it}-h(X_{it},\beta)]^2,$$

where h denotes the RHS of equation (5.11), are the NLSQ estimators. The first order conditions (FOCs) for minimisation of S(β) are:

$$\partial S(\beta)/\partial\beta = -\ \Sigma_i\Sigma_t[y_{it}-h(X_{it},\beta)](\ \partial h(X_{it},\beta)/\partial\beta) = 0.$$

In this case, this is a set of non-linear equations that do not have an explicit solution and therefore must be solved by iteration.

A potentially important consideration with such estimation is the choice of starting values for the iterative procedure. Iteration may stop at a *local* minimum as well as a *global* minimum. It is therefore necessary to select starting values for the parameters carefully and to test the robustness of estimates to alternative sets of starting values. Starting values are initially based loosely on the estimated parameter values found by Keller (1997) in the goods-trade specification of his empirical model. The FDI parameter (not present in Keller's specification) is assumed to be of similar value as

[142] As a test of the robustness of the results to differences in specification, I also estimate a linearised version of equation (5.11) as a Fixed Effects (FE) model. The econometric details of the FE model need little or no explanation here, and results are presented in Appendix E.

the trade parameter. A number of alternative sets of starting values are also used to test the robustness of the estimates. These make no difference to the results, but are listed in Appendix D2 for completeness.

The NLSQ estimates will be consistent and asymptotically normal. They will also be efficient if the errors are normally distributed (eg: in the absence of heteroskedasticity or autocorrelation, for example). Given this, NLSQ estimation on the pooled data can proceed with similar properties to OLS estimation.[143]

5.6.4: Multicollinearity

Section 5.5 shows how the R&D stock variables generally follow a smooth rising trend over the sample period. Given this, it should come as no surprise that some of these R&D stock series are closely correlated, as suggested by Stylized Fact F in Section 5.5. Table 5.2 presents simple pairwise correlation coefficients for the series in the empirical model. One of these coefficients, in particular, is close to 1, and this may cause problems with estimation.

Foreign within-sector FDI-weighted R&D stock is correlated with foreign within-sector trade-weighted R&D stock with a coefficient of .98. This is to be expected, to a certain extent, as the series are based on the same R&D stocks in the same countries. However, it does suggest that the weighting matrices derived from import flows and from FDI flows, whilst different, are not that different in their *effect* on the final series. Given that the R&D stocks of the different countries all follow a similar pattern, different linear combinations (which are nonetheless time-invariant) of these series will also inevitably follow a similar pattern. Of course a correlation coefficient of .98 is not a disaster. It is not so close to 1 as to render estimation impossible[144]. However, it does suggest that there may problems with separating the explanatory effects of the two variables.

---

[143] It is worth noting that the fit of the regression, R^2, is no longer *guaranteed* to fall between 0 and 1, although Greene (1997) suggests it is still a useful descriptive measure (see Greene, 1997, for more details).

[144] NLSQ estimation requires the columns of the regressor matrix to be linearly independent (Greene, 1997). This is discussed later.

Lichtenberg and Pottlesberghe de la Potterie's (1996) results are not inconsistent with the notion that the correlation between FDI-weighted and import-weighted R&D stocks can cause problems with estimation (see Section 5.3). Hejazi and Safarian (1998) implicitly consider the downward-biasing effect on the trade-weighted variable of introducing FDI to the Coe and Helpman (1995) equation, which is found to be substantial. However, neither paper provides much analysis of the robustness of their final results to small changes in the specification of the equation including FDI. If there is a high correlation between these variables,[145] as suggested by Hejazi and Safarian's findings, then it is difficult to be certain that parameter estimates are reflecting the true nature of the relations in the DGP and not some largely arbitrary allocation of inseparable effects to the two parameters.

This multicollinearity problem between FDI and import-weighted R&D stocks implies a need for caution in interpreting results of regressions including both variables separately. There are potential solutions to the multicollinearity problem, but all are unsatisfactory in some way. For example, principal components (estimating linear combinations of the explanatory variables) leads to parameter estimates that may be largely meaningless (Greene, 1997, p427). It is even doubtful that improved or different data would reduce the problem much. Selective omission of variables and extensive robustness testing to slight changes in specification would seem to be the best way forward, though far from satisfactory. In the end, the correlation between FDI and import-weighted foreign R&D stocks may make it difficult to pin down the relative spillover effects of FDI and trade with any certainty for some time to come.

Although the correlation is less close, within-sector trade-weighted foreign R&D stock is also highly correlated with domestic within-sector R&D stock (ρ=0.96), as is within-sector FDI-weighted foreign R&D stock (ρ=0.95). Therefore the analysis presented above also holds, albeit to a lesser extent, for the domestic within-sector R&D stock variable. This, too, will be an important consideration in the analysis of the results.

[145] Without access to the data used in these studies I cannot make any formal comment on the correlation between these variables. Given that these are both OECD samples, however, such correlation is likely.

There is one final, but very important point to make on multicollinearity in equation (5.11). Greene (1997) points out that the columns of the regressor matrix must be linearly independent if the sample asymptotic covariance matrix is to converge to a positive definite matrix. If this is not the case (ie: there is a serious multicollinearity problem) then the iterative estimation procedure may be unsuccessful. During estimation therefore, given the possibility of a collinearity problem, the iterations are checked thoroughly to make sure nothing is amiss.

5.6.5: Heteroskedasticity and Autocorrelation

NLSQ estimation will provide us with unbiased, consistent and efficient parameter estimates so long as the errors are distributed normally. Heteroskedasticity (error variances correlated with regressors) or autocorrelation (error terms correlated over time) in the model may interfere with the distribution of the errors and therefore reduce the efficiency of the coefficient estimates, although they will remain consistent. Unfortunately, there is a potential for both problems in the data.

In such a study of a panel of different industries, there is a clear potential for two types of heteroskedasticity in particular. Firstly, it is possible that the variance of the error term may be different for different sectors, because, for example, some sectors are bigger than others. This is called groupwise heteroskedasticity and is discussed in more detail in Chapter 4. Similarly, some sectors may be more alike than others, which can give rise to cross-sectional correlation of the error terms. Heteroskedasticity is tested for in each model estimated (using a simple LM test). The evidence is mixed, pointing to some cases where there is significant heteroskedasticity and others where there is not. The robustness of the standard model is tested by estimating (5.11) with a heteroskedasticity-robust GLS procedure (see Section 5.7.2). None of the main conclusions are altered, although standard errors are slightly reduced in many cases.

In time series data of this nature we typically expect to find some correlation of the errors across time. The simplest test for this is the Durbin Watson (DW) test, which is reported for the sector by sector estimations in Section 5.7.6.[146] The evidence from

[146] The DW test is not applicable to the pooled regressions.

these DW tests suggests that autocorrelation is not a significant problem in the model. However, only four of the seven sectors are estimable individually, so this evidence is not entirely conclusive.[147] As a test of the robustness of the results of the standard model, I attempt (unsuccessfully) to use an autocorrelation-robust estimation procedure, details of which are given in Section 5.7.2. This attempt is confounded by the high degree of collinearity between some of the explanatory variables, which makes it difficult to estimate ρ. In the end, the possibility of some inefficiency in the estimates and possible inaccuracy with inference has to be accepted. Estimates continue to be unbiased, however.

5.7: Results

In this section I report the results of the estimation of the standard non-linear model in log-levels on the pooled data. Alternative specifications and tests of robustness are then detailed, including:

1. Heteroskedastic-robust estimation of the standard model (see Section 5.7.2),
2. Dropping FDI, alternative depreciation rates (see Section 5.7.3),
3. Estimation in lags and in 1st differences (see Section 5.7.4),
4. Estimation of the standard model with alternative starting values (see Appendix D2),
5. Fixed Effects (FE) estimation of an augmented Coe and Helpman linear model (see Section 5.7.5),
6. Estimation of the standard model for individual sectors (see Section 5.7.6),
7. Estimation of the standard model with a business cycle variable added to capture cyclical effects on productivity (see Section 5.7.7).

[147] Evidence from the Fixed Effects estimation suggests ρ (assumed common across sectors) is quite low for the sample, again not suggestive of a serious degree of serial correlation (see Section 5.7.5).

5.7.1: Standard Model

The standard model is equation (5.11), estimated by NLSQ on the pooled data from sectors 2-8.[148] A time trend is included on the RHS along with sector dummies. The results of this estimation are presented below in Table 5.4. Unless otherwise stated, all models are estimated with TSP's LSQ procedure.

To recap, I estimate the following equation by NLSQ:

STANDARD MODEL: $LTFP_{jt} = \mu_j + \beta_0\ TREND + \beta_1 LRD_j^{f,fdi} + \beta_2 \log(RD_j^d + \beta_3 RD_j^f + \beta_4 RD_i^d + \beta_5 RD_i^f) + \varepsilon_{jt}$.

Table 5.4. Standard Model: NLSQ Estimation of Equation (5.11)

Parameter (vble)	Coefficient	Standard Error	T-Ratio
β_0 (trend)	0.029	0.005	6.06**
β_1 ($lrd_j^{f,fdi}$)	-0.121	0.081	-1.50
β_2 (rd_j^d)	0.032	0.015	2.08**
β_3 (rd_j^f)	-0.454	0.012	-38.07**
β_4 (rd_i^d)	4.06	1.28	3.18**
β_5 (rd_i^f)	-1.63	.943	-1.72*
$R^2 = 0.85$	SSR = 0.57	LM(het) = 3.28*	

Notes: Significant coefficients/test statistics at 5% are marked with ** and at 10% with *. The coefficients cannot be read as elasticities (with the exception of β_0, β_1 and β_2) because of the non-linear nature of the estimated equation.

The model appears generally well specified, with a satisfactory R^2 (although this should be regarded as an upper bound) and plausible parameter estimates. There is evidence of heteroskedasticity detected by the LM test. DW tests on individual sectors suggest little evidence of autocorrelation of the errors.[149] However, some sectors cannot be estimated separately (see Section 5.7.6), so this evidence is not entirely conclusive. The same model is estimated with a heteroskedastic-robust estimator and

[148] Sector 1 is omitted from the sample as the outside-sector variables are undefined for total manufacturing.

[149] The DW test statistics for Sectors 3, 6, 7 and 8 range from 1.67 to 2.65.

details are presented in Section 5.7.2. Very little changes. Section 5.7.2 also discusses attempts to estimate the above model with an autocorrelation-robust estimator.

The estimation process itself is reasonably straightforward. Convergence is achieved after 19 iterations. However, there is a small problem due to non-positive evaluations of the term in brackets, at some points in the data (5 observations) during the iterative process. When the bracketed term (see STANDARD MODEL above) is non-positive, the logarithm of that term is not defined. TSP treats these cases as missing observations for the particular iteration concerned. This arises because the negative coefficients for foreign import-weighted within-sector R&D and outside-sector R&D sometimes outweigh the positive value of domestic within and outside sector R&D. This is difficult to avoid without changing the specification of the equation by omitting variables or altering its non-linear structure.[150] However, given that an undefined log term only occurs very rarely in the iterations, this problem is unlikely to have any significant influence on the results. This is reflected by the robustness of the results to small changes in the model specification.[151]

The results presented above are unchanged for the alternative sets of starting values listed in Appendix D2, which suggests the iterations reach a global rather than a local minimum. They are robust (at least in sign) to all the specification changes listed above, although small changes in coefficients and standard errors do occur. The main point of sensitivity in the model specification is the sample coverage and sample period. Reducing the sample period (to estimate in lags or in order to specify an autocorrelation structure) increases the linear dependence between some of the explanatory variables, to the point where the estimation procedure breaks down (the iterations do not converge). There is little I can do, within the confines of the model and the current data set, to alleviate this problem. Estimating separately on individual industries, or small groups of industries, where possible, highlights some interesting heterogeneity behind the panel results.

[150] Section 5.7.5 outlines estimation of a linear FE version of the model, based on Coe and Helpman's empirical model, in which this problem does not arise. Variables could also be entered in the equation in nonnegative form (eg: squared). Expressing the RHS R&D stock variables as index numbers does nothing to alleviate this problem.

[151] Hall and Cummins (1997) support this assumption. They state that... 'During iterations, numeric errors ...(in the SSR) are not a problem, since...(the stepsize) is squeezed repeatedly until the numeric errors stop.' See Hall and Cummins (1997), pp86, paragraph 1.

5.7.2: GLS Estimation of the Standard Model

As discussed in Section 5.6, if the errors are not normally distributed then NLSQ estimation may be inefficient and we may have problems with inference. Two reasons why this might be the case are that the error variances might be correlated with the regressors (heteroskedasticity) or that the errors might be correlated across time (serial correlation). It is straightforward to obtain heteroskedastic-consistent (White-robust) estimates of the variance-covariance matrix in TSP by adding the HETERO command to the LSQ command.

Table 5.5 below shows the standard model estimated with heteroskedasticity-robust standard errors. Coefficient estimates are unchanged, but standard errors are slightly smaller/larger than the standard model in some cases (compare with Table 5.4). Overall, the presence of heteroskedasticity makes no significant difference to the analysis.

Table 5.5: Heteroskedasticity Robust Errors

Parameter (vble)	Coefficient	Standard Error	T-Ratio
β_0 (trend)	0.029	0.006	4.94**
β_1 ($lrd_j^{f,fdi}$)	-0.121	0.101	-1.20
β_2 (rd_j^d)	0.032	0.014	2.28**
β_3 (rd_j^f)	-0.454	0.007	-67.45**
β_4 (rd_i^d)	4.06	0.841	4.83**
β_5 (rd_i^f)	-1.63	.632	-2.57**
$R^2 = 0.85$	SSR = 0.57		LM(het) = 3.28*

Obtaining efficient estimates robust to serial correlation in the non-linear model is unfortunately problematical. In theory we can specify a lag structure and estimate equation (5.11) by Generalised Method of Moments (GMM), which will give us consistent standard errors. In practice, however, this is unsuccessful. The difficulty arises because, using pooled data, the first observation has to be dropped from each sector in order to estimate ρ. The iterations do not converge with this smaller sample (see Section 5.7.4, where a similar problem arises when trying to estimate the model with first lags). The most probable explanation for this is that reducing the sample in this way accentuates the multicollinearity problem to the point where the columns of the regressor matrix are insufficiently linearly independent for the asymptotic covariance matrix to converge (see Section 5.6.4). I therefore have to live with the possibility that the standard errors may not be 100% accurate, although there is no particular evidence of serial correlation in the sample. Coefficients, however, will continue to be unbiased.

5.7.3: Alternative Specifications and Variables

A number of alternative specifications of the standard model are estimated to check robustness. For example, in Table 5.6 below, I drop the FDI-weighted R&D stock. Such alternative specifications are important to check the robustness of the estimates because of the high degree of correlation between the explanatory variables. The estimates given below are heteroskedasticity-robust.

Table 5.6. Omitted Variables: Drop FDI

Parameter (vble)	Coefficient	Standard Error	T-Ratio
β_0 (trend)	0.022	0.002	14.42**
β_2 (rd_j^d)	0.030	0.014	2.19**
β_3 (rd_j^f)	-0.453	0.004	-113.07**
β_4 (rd_i^d)	4.294	0.203	21.19**
β_5 (rd_i^f)	-1.794	0.159	-11.26**
$R^2 = 0.85$	SSR = 0.58	LM(het) = 6.42*	

The coefficients after dropping FDI-weighted foreign R&D from the standard model are essentially unchanged (some are larger in significance and slightly larger in absolute value). Therefore the model is generally robust to this change. It is interesting that the magnitude of the foreign trade-weighted variables do not change. This suggests the FDI and the trade-weighted variables may be complementary rather than substitutes (see discussion in Section 5.8.5, along with the discussion of Hejazi and Safarian, 1998, in Section 5.3).[152]

I have so far assumed a knowledge depreciation rate of 10%. Alternative variables are constructed using a 20% depreciation rate and the estimation results of the standard model using these variables is outlined below in Table 5.7. In the sense that the directions of the relationships between the variables are the same with both depreciation rates, the model is robust to this change. However, the size and strength of these relationships are in many cases sensitive to the chosen depreciation rate of knowledge. For example, outside sector R&D is now insignificant (both foreign and domestic), although coefficients are broadly similar. Own R&D appears to have a greater influence on TFP with a depreciation rate of 20%.

Table 5.7. Alternative Knowledge Depreciation Rate

Parameter (vble)	**Coefficient**	**Standard Error**	**T-Ratio**
β_0 **(trend)**	0.025	0.004	5.56**
β_1 **($lrd_j^{f,fdi}$)**	-0.078	0.088	-0.89
β_2 **(rd_j^{d})**	0.052	0.031	1.65*
β_3 **(rd_j^{f})**	-0.613	0.303	-2.02**
β_4 **(rd_i^{d})**	6.964	9.199	0.76
β_5 **(rd_i^{f})**	-2.145	5.345	-0.401
$R^2 = 0.85$	*SSR = 0.58*	*LM(het)=6.44**	

Notes: Heteroskedasticity-robust standard errors.

5.7.4: Lags and First Differences

The standard model is specified in contemporaneous values. Implicitly, this assumes that spillovers occur simultaneously (or at least rapidly) after R&D has taken place. An alternative (and possibly more realistic) assumption is that technology only spills over with some time delay (see, for example, Mansfield, 1985; Keller, 1997). To capture this possibility, I regress TFP on the first and second lags of the explanatory

[152] Omission of the foreign import-weighted outside-sector R&D stock variable form the standard model makes little difference to the results.

variables.[153] Unfortunately, neither the first lag nor the second lag model is well specified in the sense that neither is estimable by NLSQ. The iterations do not converge in either case.[154] The most likely explanation for this is that reducing the sample size (T=18 or 19 as opposed to 20) accentuates the multicollinearity problem, which prevents the variance-covariance matrix from converging correctly (see Section 5.6.4 for a brief discussion of this point). This explanation is supported by the correlation coefficient between FDI-weighted and trade-weighted R&D stocks for the restricted sample (.985 as opposed to .980, for T=19). As discussed in Section 5.7.2, this also has problematical implications for autocorrelation-consistent estimation (iterations break down while trying to estimate ρ).

Although I have assumed that the variables in the standard model are stationary (I{0}), the nature of unit root testing means there is always some element of doubt with small time-dimensions (T=20 in this case). Consequently, I estimate the standard model in first differences to check the robustness of the results against the possibility of unit roots.

First differencing removes all the sector dummies and the time trend, leaving us with a simple constant. We also lose levels information in the other variables, which may have considerable explanatory power. Therefore, although we would expect the parameters of the variables in the first differenced equation to be similar (if the stationarity assumptions are correct) we might expect some reduction in the explanatory power of the model. The first differenced equation is given below, as the non-linear structure means the transformation is not straightforward.

First Differenced Equation: $(LTFP_{jt} - LTFP_{jt}[-1]) = \mu + \beta_1(LRD_j^{f,fdi} - LRD_j^{f,fdi}[-1]) + \beta_2 \{\log(RD_j^d + \beta_3 RD_j^f + \beta_4 RD_i^d + \beta_5 RD_i^f) - \log(RD_j^d[-1] + \beta_3 RD_j^f[-1] + \beta_4 RD_i^d[-1] + \beta_5 RD_i^f[-1])\} + (\varepsilon_{jt} - \varepsilon_{jt-1})$.

The pattern of the results is generally similar to that for the standard model in levels, but the explanatory power of the model is considerably lower without the levels information of the standard model. There is no strong evidence against the assumption

[153] I do not lag further back in time because of the shortness of our sample period.

of stationarity, given the general robustness of the results. The small changes in the coefficients and standard errors are to be expected as the sample is restricted (we lose the first observation from each sector) for the first differenced equation. The diagnostics for the first differenced model are encouraging (no evidence of heteroskedasticity). In summary, the results from the first-differenced model are quite encouraging. Although I may be overestimating the significance of variables in the standard model, the pattern of results is generally robust. Nonetheless, further research should examine the model assuming non-stationarity in far greater depth.

Table 5.8: First Differences

Parameter (vble)	Coefficient	Standard Error	T-Ratio
Constant (μ)	0.033	0.013	2.67**
β_1	-0.364	0.195	-1.87*
β_2	0.024	0.040	0.61
β_3	-0.487	0.184	-2.65**
β_4	3.925	23.680	0.166
β_5	-1.262	17.286	-0.07
$R^2 = 0.13$	$SSR = 0.44$	$LM(het) = 0.75$	

5.7.5: Coe and Helpman's Linear Model Extended

Although the non-linear specification of the empirical model is similar to that of Keller (1997), the majority of other preceding papers in this field recently have adopted a more *ad hoc* linear empirical model, following on from that of Coe and Helpman (1995). Of course, these studies (Coe and Helpman, 1995; Hejazi and Safarian, 1998, for example) are at macro-level and not industry level, and focus on a number of OECD countries rather than just the UK, so direct comparison with the current study is not possible. However, for rough comparative purposes and also to escape the non-linear problems associated with the bracketed logarithmic expression which can be undefined, I carry out a Fixed Effects (FE) estimation of an extension of

[154] Nor do they explode.

Coe and Helpman's equation.[155] A further bonus of the FE model is that it can account for autocorrelation without problem.

The *ad hoc* extension of Coe and Helpman's model is given below:

(CH) $LTFP_{jt} = \mu_j + \beta_0\, TREND + \beta_1 LRD_j^{f,fdi} + \beta_2 logRD_j^d + \beta_3 logRD_j^f + \beta_4 logRD_i^d + \beta_5 logRD_i^f + \varepsilon_{jt}$.

The equation is estimated as a FE model using LIMDEP7 for all sectors (except total manufacturing). The results are given below in Table 5.9.

Table 5.9 Linear (FE) Model

Parameter (vble)	**Coefficient**	**Standard Error**	**T-Ratio**
β_0 (trend)	0.032	0.011	2.94**
β_1 ($lrd_j^{f,fdi}$)	-0.058	0.125	-0.47
β_2 (rd_j^d)	0.003	0.052	0.06
β_3 (rd_j^f)	-0.300	0.208	-1.44
β_4 (rd_i^d)	1.690	0.898	1.88*
β_5 (rd_i^f)	-1.489	0.917	-1.62
$R^2 = 0.84$	$SSR = 0.62$	$\rho = 0.45$	$LogL = 181.28$

Notes: Autocorrelation-consistent estimation.

Although some parameters are close to being significant, none (with the exception of the parameter on the TREND and domestic outside-sector variables) are significant at 5%. The hypothesis that $\beta_1 = \beta_2 = 0$ is not rejected[156] giving us the parsimonious model presented below in Table 5.10.

[155] I consider only their most simple equation not including import/GDP ratios.
[156] Using an F[2,127] test.

Table 5.10: Parsimonious Linear Model

Parameter (vble)	Coefficient	Standard Error	T-Ratio
β_0 (trend)	0.034	0.010	3.31**
β_3 (rd_j^f)	-0.348	0.178	-1.96**
β_4 (rd_i^d)	1.716	0.886	1.94*
β_5 (rd_i^f)	-1.538	0.905	-1.70*
$R^2 = 0.84$	$SSR = 0.62$	$\rho = 0.46$	$LogL = 181.16$

Notes: Autocorrelation-consistent estimation.

The table above looks rather different from that presented in Section 5.7.1 for the standard non-linear model. For example, the elasticity of domestic outside-sector R&D stock is now around 20 times greater than previously. However, there are some striking similarities, despite the substantial differences in the structure of the two models. Firstly, the FDI variable does not seem to have any significant effect on TFP. Secondly, although own R&D was weakly significant in the standard model and is now insignificant, the actual coefficients on both this and the FDI variable are remarkably similar in size. Thirdly, and perhaps most importantly, the pattern of signs of the effects is identical in the non-linear and linear models. Thus, although the *strength* of the various effects on TFP is different, the *direction* of these effects is the same. The overall explanatory power of the linear model is very close to that of the non-linear model. I conclude that, at the very least, the direction of the effects identified in the non-linear model is robust to this substantial specification change.

In common with the non-linear model, the linear model suggests foreign R&D spillovers play a significant *negative* role in UK TFP. Interestingly, in contrast with the majority of previous research in the field (for example, Coe and Helpman, 1995), I find no significant effect of own-sector domestic R&D on TFP in the UK, in the linear model. As suggested in Section 5.8, the UK appears to behave in a peculiar way in terms of the sources of TFP growth over the sample period. A full discussion of this question is presented in Section 5.8.

5.7.6: Individual Sectors

The panel covers 7 sectors of UK manufacturing industry. This is far from an exhaustive coverage, but lack of FDI data prevents a more comprehensive study at this stage. So far, I have imposed common parameter estimates on all sectors. However, we might expect some heterogeneity in spillover effects across sectors. The most obvious stratification of the sample is into high technology and low technology industries. R&D intensity in the food and paper industries is considerably lower than that in the chemicals and electrical engineering industries, for example. We might expect the TFP effects of own R&D to have a greater effect in the high tech industries than in the low-tech industries. Equally, it is unlikely that high tech industries will benefit much from spillovers from low tech industries (we would expect the opposite). I estimate the standard model on each sector individually to examine whether there is such heterogeneity across sectors.

Given the multicollinearity-induced difficulties of cutting the sample encountered elsewhere (see Section 5.7.4, for example) it should come as no surprise that it is not straightforward to estimate the model on all industries separately. Indeed, I cannot obtain estimates for Sectors 2, 4 and 5 at all (the iterations do not converge). However, it is possible to obtain estimates for Sectors 3 (paper), 6 (mechanical engineering), 7 (electrical engineering) and 8 (transport). These results support the argument that we might expect different effects in different sectors. These estimates for Sector 3 and for Sectors 6-8 pooled (engineering sectors) are presented below.

The model for Sector 3 above appears generally well specified, although most of the explanatory variables are not significant. This is interesting, as it suggests that technology-related factors are not of much importance for the productivity of a low technology sector such as paper. Rather, it is the trend variable, perhaps capturing other effects such as de-unionization, for example, that is the major explanatory factor. However, the coefficient on the FDI variable is significant. It is negative, as found in the complete panel, although considerably larger (an elasticity of –0.56 as opposed to –0.12). Therefore, at least for the paper industry, FDI-weighted foreign technological advances have a negative effect on UK productivity. Possible explanations of this negative effect are discussed in Section 5.8.2.

Table 5.11: Sector 3

Parameter (vble)	Coefficient	Standard Error	T-Ratio
β_0 (trend)	0.061	0.022	2.73**
β_1 ($lrd_j^{f,fdi}$)	-0.560	0.283	-1.98**
β_2 (rd_j^d)	0.060	0.110	0.54
β_3 (rd_j^f)	-3.379	4.639	-0.728
β_4 (rd_i^d)	-0.404	0.397	-1.02
β_5 (rd_i^f)	0.466	0.511	0.91
$R^2 = 0.94$	SSR = .03	DW = 1.70	LM(het)=.54

Table 5.12: Sectors 6-8 (Engineering)

Parameter (vble)	Coefficient	Standard Error	T-Ratio
β_0 (trend)	0.014	0.001	9.43**
β_1 ($lrd_j^{f,fdi}$)	0.141	0.029	4.87**
β_2 (rd_j^d)	0.057	0.032	1.75*
β_3 (rd_j^f)	-0.422	0.139	-3.03**
β_4 (rd_i^d)	0.073	12.303	0.01
β_5 (rd_i^f)	0.924	8.901	0.10
$R^2 = 0.89$	*SSR = 0.27*	*LM(het)=4.00*	

Notes: Heteroskedasticity-robust estimator.

In contrast to the results for the paper industry presented in Table 5.11, more of the explanatory variables are significant in the engineering sectors. Only the two outside-industry variables are insignificant. This, again, is very interesting. It is consistent with the hypothesis that higher technology industries do not have much to learn from lower technology industries (in terms of knowledge spillovers). The significant effects found in the complete sample are therefore likely to be capturing technology flows *from* the engineering sectors to the other sectors.

Leaving aside the FDI variable for the present, the coefficients of the other explanatory variables are similar to those obtained for the full sample. The elasticity to own R&D in the engineering sectors is 4.6%. This is around three times that for the whole sample, which I argue above might be expected for higher technology sectors. The coefficient on trade-weighted foreign R&D stocks is almost identical to that for the whole sample: The same negative foreign spillover effect is coming through. However, the elasticity is greater for the engineering sectors (around twice that for the whole sample), which again is consistent with the greater intra-industry role of technology for high tech industries.

Turning now to the FDI variable, I find a clear contrast with the complete sample: The FDI variable has a positive effect on TFP in UK engineering sectors as opposed to a (weak) negative effect for the wider sample of manufacturing sectors. Put together with the results for Sector 3 above, this is consistent with a pattern of positive FDI-transmitted spillovers for high technology industries and negative FDI-transmitted spillovers for low technology industries. These effects are balanced out (although slightly more to the negative side) in the complete sample of sectors, hence the insignificant FDI coefficient in Table 5.4. A fuller discussion of this point is presented in Section 5.8.2.

5.7.7: Introducing A Business Cycles Variable

The series for TFP display a significant degree of cyclicality (see Figure 5.1). This is not the case for the R&D stocks. There are two possibilities to try to control for this cyclicality. First, the TFP series could be averaged over a number of years. The two peak to peak periods in the sample period are 1973 to 1979 and 1979 to 1989. Therefore, TFP would need to be averaged over 7 to 10 years. This would considerably reduce the length of the sample period, so is rejected. The second possibility, which has less dramatic implications for the sample period, is to include an additional explanatory variable to capture cyclical effects. In this, I follow Engelbrecht (1997), by including a variable of GDP growth on the RHS of Equation (5.11).[157] The

[157] Outside the brackets.

variable is labelled CYCLE, and is defined as $\log GDP_t - \log GDP_{t-1}$.[158] Table 5.13 below presents the results from this extension to the standard model.

Table 5.13: Standard Model Including CYCLE

Parameter (vble)	Coefficient	Standard Error	T-Ratio
β_0 (trend)	0.030	0.004	7.21**
β_1 ($lrd_j^{f,fdi}$)	-0.143	0.070	-2.04**
β_2 (rd_j^d)	0.049	0.025	1.97**
β_3 (rd_j^f)	-0.451	0.095	-4.75**
β_4 (rd_i^d)	-2.24	5.76	-0.39
β_5 (rd_i^f)	2.75	4.48	0.61
β_6 (cycle)	1.71	0.184	9.29**
$R^2 = 0.92$	$SSR = 0.33$	$LM(het)=6.63$**	

The results above are broadly similar to those for the standard model. Spillovers through FDI and foreign intra-industry trade are very similar in size. Own R&D also displays a very similar effect. The major change is that domestic inter-industry and foreign inter-industry spillovers are no longer significant, when the cycle variable is included. Once again, this does not contradict the signs of the estimated effects from the standard model, but it does question their significance. The cycle variable itself is significant and positive. In other words, productivity is pro-cyclical.

5.8: Discussion

5.8.1: Own R&D

The idea that a great deal of economic growth can be explained by technological progress is widely accepted (see, for example, Solow, 1957; Denison, 1985; Jorgenson, 1990). This is true for productivity growth at industry level as well as for output growth at macro level. For example, Griliches (1980) finds an industry

158 Engelbrecht's variable is defined in the same way, but labelled BC for business cycle.

elasticity of output to R&D stock of between 0 and 7% for the US. Englander et al (1988) finds a broader range for this elasticity for the G5 of 0-50%. Cameron (1995) finds a range of 0-27% for the UK, over a similar sample period to that considered here.

The standard model presented above in Table 5.4 shows an elasticity of productivity to own R&D stock of just over 1.5% for the sample of 7 UK manufacturing sectors. This is well within the bounds of the elasticities found by Cameron (1995), as well as broadly consistent with studies of productivity elsewhere. Taking chemicals as an example, TFP has increased over the sample period by 38% (the second highest of the 7 sectors). UK R&D stock in the industry has increased by 276% over the sample period (the highest of the 7 sectors). Of the 38% increase in TFP, the standard model assigns just over 4% to the increase in own R&D stock. In other words, between an eighth and a tenth of the productivity increase in the UK chemicals sector between 1973 and 1992 has been driven by technological knowledge accumulation internal to the sector. The figures for other sectors show some contrast. Productivity in the food sector, for example, has grown by only 23% over the sample period. A lower R&D stock growth means only 1.5% of this 23% growth (or around one sixteenth) is down to own R&D, however. Table 5.14 presents similar figures for the remaining sectors, based on the standard model (see Section 5.8.6). The calculation of the elasticities in the standard model is explained below.

The non-linear structure of the standard model means that all coefficients cannot be read as elasticities (the exceptions are β_0 and β_1 for the trend and FDI-weighted foreign R&D). Take the simple example where $\beta_0=\beta_1=\beta_4=\beta_5=0$, so that:

(5.12): $\text{LogTFP}_{jt} = \mu_j + \beta_2\log(RD_j^{\,d} + \beta_3 RD_j^{\,f}) + \varepsilon_{jt}$.

Relabel with y=TFP, x= $RD_j^{\,d}$ and z=$RD_j^{\,f}$ so that:

(5.13): $\partial\log y/\partial x = \beta_2/(x+\beta_3 z)$.

$\Rightarrow\partial\log y = [\beta_2/(x+\beta_3 z)]\partial x \Rightarrow.\partial\log y = [\beta_2 x/(x+\beta_3 z)]\partial\log x$.

The term in square brackets is the elasticity of y with respect to x. The elasticity of y with respect to z is similarly given by $[\beta_2\beta_3 z/(x+\beta_3 z)]$. These elasticities are not invariant to the values of x and z. Therefore, the elasticities of TFP with respect to own R&D, domestic inter-industry spillovers, foreign within-industry spillovers and foreign inter-industry spillovers are all calculated at sample means.

The estimate for the productivity effect of own R&D is robust to a number of specification changes. Dropping the FDI variable makes no difference. Using a knowledge depreciation rate of 20% as opposed to 10% increases the estimated elasticity from 1.5% to 2.5%. In the model including a business cycle variable (see Section 5.7.7), the elasticity of TFP with respect to own R&D rises to 2.2%. Own R&D drops out of the linear model and the first-differenced model. Perhaps the most interesting alternative estimate is that obtained for the sample restricted to engineering sectors (6-8), which we can classify as relatively high tech. In this case, the elasticity rises to 4.6%. Using this figure, nearly one third of the productivity increase in the UK electrical engineering sector (33%) can be explained by the increase in own R&D stock over the sample period (209%), for example. The stronger elasticity for higher technology industries (and the more rapid growth in R&D stock for these industries) is consistent with the idea that higher technology industries display greater technological opportunities than lower tech industries (see Chapter 4 for a discussion of this point).

To sum up, estimates of the TFP elasticity of own R&D for the sample of UK manufacturing industries range from 0% to 5%. This range is consistent with a variety of estimates from previous research. It is below the range found by Keller (1997) of 7% to 17%, for a similar sample period. This suggests the UK falls towards the bottom of the OECD countries sampled by Keller in terms of its recent R&D performance, an argument consistent with the evidence of falling relative R&D stocks presented in Figure 5.3.

5.8.2: FDI-Transmitted Spillovers

The potential for spillovers through FDI has been widely commented on. However, with a few notable exceptions, little research has been carried out to date that quantifies these spillover effects. Section 5.3.3 (and 5.3.5) discusses these exceptions. Lichtenberg and Pottelsberghe de la Potterie (1996) find evidence of significant positive spillovers through *outward* FDI only, for a sample of OECD countries at national level. Hejazi and Safarian (1998) find evidence of total (inward and outward) FDI-transmitted spillovers for the G6, also at national level. Indeed, Hejazi and Safarian (1998) find a country level TFP elasticity to FDI-weighted foreign R&D stocks of around 10%.

Hejazi and Safarian (1998) argue that we might expect FDI-transmitted spillovers from inward FDI because such FDI usually takes place in order to exploit a firm-specific capability (a particular piece of scientific knowledge, for example) which may be new to the host economy. This particular capability may be partly appropriable by other (indigenous) firms. Equally, outward FDI might give rise to spillovers to the source country through decentralization of the R&D functions of the firm, or through the desire of firms to access knowledge in the host country.

The FDI-weighted foreign R&D stock of this study is for total inward and outward FDI (the variable cannot be defined separately for inward and outward FDI because of non-positive log terms) at industry level or the UK. Only intra-industry FDI is considered, due to data limitations.[159] My prior, based on the arguments above and the evidence presented by Hejazi and Safarian (1998), amongst others, is that this variable might have a positive effect on TFP in UK industries.

However, as can be seen from Table 5.4, this is not the case. The estimated effect on TFP of the FDI-weighted foreign R&D stock variable is *negative*, although not significant at standard levels. Leaving aside the standard error for a moment, the coefficient suggests a TFP elasticity of FDI of around –12% for the sample of UK

[159] This may lead to measurement errors where firms span several sectors. Branstetter (1996) refers to this as the possibility of mis-measurement spillovers.

manufacturing industries. In other words, an increase of 1% in foreign within-industry R&D stocks, weighted by FDI shares, leads to a *fall* in TFP in the UK sector of .12%.

Of course, the high standard error suggests the coefficient is not significantly different from zero, giving us no effect of FDI-weighted foreign R&D stocks on UK TFP. This insignificant effect (albeit with a negative coefficient) is consistent across the majority of the alternative specifications estimated as test of robustness. There are three exceptions to this. Firstly, the model including the business cycle variable displays a significant elasticity of –14%. Secondly, the model estimated in first differences displays a negative coefficient which is significant at the 10% level. The elasticity implied by this model is –36%. In other words, a 1% increase in the rate of growth of FDI-weighted foreign R&D stocks causes a .36% fall in the rate of growth of TFP. The other exception to the pattern of insignificant FDI results is when industries are estimated separately.

Restricting the model to the paper sector only gives us an elasticity of TFP with respect to FDI-weighted foreign R&D of –56%. This is considerably larger in magnitude to the estimate for the sample as a whole and is also significant at the 5% level. For reasons discussed above (see Section 5.8.1), we might expect a smaller positive effect for low tech industries than for high tech industries, but this does not imply that we should see a stronger negative effect for low tech industries than for high tech ones.

How can we explain such a negative effect? Foreign spillovers through trade have been found to be negative in a few cases in previous research. For example, Branstetter (1996) finds no evidence of positive Japan-US spillovers through trade and some evidence of negative spillovers, at firm level. There are three explanations offered for these apparent negative spillover effects. Firstly, they could be caused by inaccuracies in the data. Branstetter largely discounts this possibility by throwing away outliers from the data set and still getting the negative result. Secondly, Branstetter's data are characterised by a high degree of multicollinearity. As is the case in the present paper, some of the spillover variables are more correlated with each other than they are with the TFP variable. Finally, it is suggested that the negative elasticities could be reflecting the dominance of a negative competition externality over the positive technological spillover externality. In other words, an increase in foreign rivals' R&D might affect the domestic firm's R&D incentives and thus its technological progress. Bertschek (1995) argues that many apparent spillover effects might be capturing such competition effects (see Section 5.3.4).[160]

Engelbrecht (1997) finds evidence of negative spillovers through trade for some countries, at national level. These countries are the US, Canada and West Germany, which are described as technology-source countries. He argues that this may reflect the fact that these technology-source countries may have little to learn from foreign R&D. This would lead us to expect insignificant international spillover effects, but not *negative* effects.

All the above arguments are possible explanations for the insignificant or negative FDI-spillover effects found in the present paper. For most of the sample period, the UK was the world's 4th most R&D-intensive country. Therefore, it is possible that the UK has less to learn from foreign R&D than many other countries because of its position as a technology-source country. This might explain the lack of a positive effect.[161]

The possible explanations of a negative effect suggested by Branstetter are also all applicable here. The data are somewhat imprecise in places and the spillover variables, particularly the FDI-weighted and trade-weighted foreign R&D stock variables, display a high degree of correlation. Against this explanation is the fact that the insignificant or negative result seems robust to a number of specification changes, which we might not expect if multicollinearity were the prime cause. The most plausible (and intuitively appealing) explanation is that competition effects are outweighing, or at least cancelling out, spillover effects.

Chapter 3 presents a model where increased competition (because of increasing openness to trade) from a technologically superior foreign firm has a positive incentive

[160] A fourth possible explanation is that FDI might be attracted to sectors with poor productivity and the coefficient on FDI-weighted R&D is picking up such a reversed causal effect. This simultaneity problem is more likely to be present in the individual sector estimation, as the panel allows for initial productivity levels in the fixed effects. See Section 5.6.2 for a brief discussion of this point.

effect on the domestic firm's R&D investment. However, a variant of the model of Chapter 3 with increasing returns to R&D predicts the opposite (see McVicar, 1998). In the absence of any random factor or perfect spillovers, the low-technology firm essentially gives up trying to compete on level terms with the technology leader. The vast majority of FDI from and to the UK goes to/comes from the US. Figure 5.3 shows the US's technological superiority to the UK in manufacturing. Therefore increasing R&D in the US relative to the UK, coupled with the large US share of UK FDI, might be expected to have a negative incentive effect on UK R&D investment.[162]

Backtracking a little, the negative spillover effect from FDI is much stronger for the paper industry than for the whole panel of industries. The paper industry is relatively low tech in our sample. In contrast, sectors 6-8 (engineering industries) are relatively high tech. I find a very interesting FDI-spillover result when the sample is restricted to these engineering industries. Table 5.12 shows the estimation of the standard model for the three engineering industries. In this case, the FDI variable has a significant *positive* coefficient. In fact, the engineering sectors display a TFP elasticity of FDI-weighted foreign R&D of +14%. Consider the electrical engineering sector, for example. FDI-weighted foreign R&D in this sector increases over the sample period by 178%. TFP in the sector in the UK increases by 33% over the sample period. Of this 33%, 25% can be explained by spillovers transmitted through FDI. So, at least for these industries, the arguments of Hejazi and Safarian (1998) seem true: Knowledge spillovers from abroad can be transmitted through FDI and have significant effects on TFP.

[161] Purely econometric reasons also might explain the insignificance of the coefficient (eg: autocorrelation), but not the negative sign.

[162] A further prediction of the Chapter 3 model, along with numerous other models, is that spillovers themselves might reduce the incentives to invest in R&D, particularly when knowledge spills over to those in direct competition with the researching firm. Increasing trade with technologically close economies might therefore be expected to reduce incentives to invest in R&D if spillovers increase with trade. Infact, a simple spillovers/trade relationship in this model gives rise to an inverted U-shape relationship between integration and R&D, where beyond an optimal tariff point, increased trade results in reduced R&D because of the increased spillovers to closer competitors. Where countries are close in technology space (eg: UK, France and Germany, for example), increased trade (or FDI) between them might be expected to reduce R&D incentives, and thus productivity growth, because of knowledge spillovers. However, given my use of time-invariant FDI-shares, this offers little further explanation of the current results.

Although there are only 7 sectors in the current sample, the results are consistent with a systematic difference in the importance of international knowledge spillovers between high tech and low tech industries. Following the discussion of Branstetter (1996), we appear to see the conflicting effects of positive knowledge spillovers and negative competition effects driving the FDI variable results. In the sample overall, the two effects seem to counteract each other. For the paper industry (the only low-tech industry I can obtain estimates for), the negative competition effects dominate. For the engineering industries, competition effects are out-weighed by spillover effects.

Summing up, the results suggest a mixed pattern of FDI-transmitted spillovers. Elasticities vary between a low of –56% and a high of +14%, although for the complete sample, the effect is not significantly different from 0. The overall sense of the results is that, at least for low and medium technology sectors, foreign spillovers through FDI are not an important determinant of TFP. In contrast, for high technology sectors, FDI is an important transmission mechanism for international spillovers.

5.8.3: Spillovers Through Intra-Industry Trade

Since Coe and Helpman (1995), a large number of studies have found evidence of significant positive international spillovers through trade, particularly at country level (see Section 5.3). At industry level, the evidence is more mixed. Some previous research suggests that knowledge spillovers are primarily intra-national rather than international in scope (see, for example, Branstetter, 1996; Cameron, 1995). However, the general picture is that there is some evidence of knowledge spillovers, transmitted through trade, between industries in different countries. Keller (1997), for example, finds foreign intra-industry spillovers to have roughly the same TFP effect as domestic own R&D (an elasticity of between 7% and 17%).

The evidence presented in Table 5.4 for the standard model is that import-weighted foreign R&D has a *negative* (and highly significant) effect on TFP in UK manufacturing industries. The coefficient of –0.45 corresponds to an elasticity of -1.5% (evaluated at the sample means). This is clearly not consistent with Keller's results for the OECD sample. In the chemicals sector, for example, this corresponds to a 4% *fall* in TFP over the sample period (see Table 5.14). This estimate is highly robust to specification changes. In particular, it hardly changes when FDI is dropped. This suggests the estimated negative effect is not related to multicollinearity in any substantial way. Nor does the elasticity change much when the business cycle variable is added to the model. An alternative knowledge depreciation rate slightly *increases* the magnitude of the coefficient. In first differences, the coefficient is again very close. Finally, in the linear model, the coefficient is again negative, but the elasticity is much larger (around –35%).

In Section 5.8.2 I speculated as to why the TFP effect of FDI-weighted foreign R&D stocks might be negative in some detail. That discussion applies equally here. The possibility of data errors or multicollinearity cannot be completely discounted here. However, it is the combined effects of the UK's position in the global economy as a technology source country and the possibility of negative competition effects on technological progress that provide the most plausible and intuitively appealing explanations of these consistently negative foreign spillover effects.

Where Hejazi and Safarian (1998) found that including FDI in the macro linear model had a big effect on the TFP elasticity of trade-weighted foreign R&D, I find no such effect in the industry-level non-linear model. The two variables are highly correlated, but the elasticity of the trade-weighted variable is highly robust to dropping the FDI variable. In other words, including FDI or not in the standard model seems to have very little effect on the estimated strength and direction of trade spillovers.[163] Where Hejazi and Safarian (1998) find evidence that FDI and trade are substitute knowledge transmission mechanisms, as well as complementary, I find evidence only of their complementarity. Of course, the UK does seem to be somewhat peculiar, compared to broader samples of OECD countries, at least in terms of spillovers. It would be unwise to draw any wider conclusions about the relationship between FDI and trade-transmitted spillovers from this study alone.

For the FDI variable I found evidence of a significant sign difference between a low tech industry (paper) and high tech (engineering) industries. The difference is less acute here. The paper sector displays no significant trade-related foreign spillovers at

[163] Although the size of the standard errors is generally affected (see Section 5.8.2).

all. The engineering sectors, however, display slightly larger negative spillovers than the sector as a whole (an elasticity of –2.5% as opposed to –1.5%). It is entirely possible that in the paper sector (with just 20 observations), the collinearity between the FDI and the trade variables is driving the high negative coefficient on FDI and the insignificant coefficient on trade. Although I argue that this is not the case for the sample as a whole, it may be that the trade and FDI variables are acting as substitutes in this much smaller sample.[164]

Summing up, estimates of the elasticity of TFP with respect to import-weighted foreign within-sector R&D stocks range from 0 to –35%, although if we discount the linear model, the estimates fall in the much tighter range of 0 to –2.5%. Nowhere in the sample is there any evidence of positive spillovers through international trade for the UK. As with the FDI variable, this negative effect may be reflecting the combination of the UK's position as a technology source country and some competition disincentive effects to invest in technological progress.[165]

5.8.4: Spillovers Through Domestic Inter-Industry Trade

Previous research is undecided whether intra-national spillovers are more important than international spillovers (see, for example, Branstetter, 1996; Cameron, 1995) or whether domestic inter-industry spillovers are weaker than foreign intra-industry spillovers (Keller, 1997). Of course, this contrast may be due to the different samples used by these authors. Keller has a sample of eight OECD countries, whereas Branstetter examines only the US and Japan and Cameron only the UK. It is likely that technology source countries, of which the US, Japan and perhaps the UK are examples, are less likely to benefit from foreign spillovers than from domestic spillovers.

Table 5.4 shows a coefficient of 4.06 on the domestic inter-industry spillover variable for the standard model, which is significant at 5%. This corresponds to an elasticity of

[164] In fact, dropping the FDI variable from the paper sector has little effect on the foreign within-sector trade-weighted R&D effect. However, it does cause the coefficient on own R&D to become negative, although not quite significantly so at 10%.

[165] Again, the possibility of some simultaneity effect cannot be ruled out, particularly for the individual sector estimation.

+7.6%. In other words, a 1% increase in the input-weighted R&D stock of the other 6 industries in the sample leads to a .076% increase in TFP in the industry being considered. This elasticity is reasonably robust to specification changes, ranging from zero to 7.6%.

An elasticity of 7.6% in the standard model is interesting in a number of respects. Firstly, it is positive, in contrast to most of the foreign spillover effects. Thus, the conventional explanation that industries can appropriate some of the benefits from other industries' R&D seems to hold here. The explanations of the negative foreign effects are not applicable to domestic spillovers. For example, I argued that the UK might not have much to learn from foreign R&D because of its position as a technology source country. This does not mean that industries in the UK will not learn from each other. I also argued that there might be a negative competition effect on R&D incentives that in some industries was dominating any positive spillover effect. If we make the reasonable assumption that, in general, competition within industries is more intense than competition across industries, then we would not expect this effect to dominate inter-industry spillovers to the same extent.

Secondly, the estimate of the elasticity of TFP to domestic outside-sector R&D is considerably larger than that for own R&D. This is consistent with Keller's findings from the goods trade specification of his model (Keller, 1997). This might be expected for low technology sectors, where the filtering down of significant technological progress from elsewhere might have considerable TFP effects. For example, the availability of rapidly improving computers (electrical engineering) or improved pesticides (chemicals) might have a big impact on the food industry. Indeed, almost all of the increase in TFP in the UK food industry over the sample period can be explained by technological improvements elsewhere in the UK that have spilled over (see Table 5.14). We might expect own R&D to be more important in higher tech industries, however. Estimating on the paper and engineering sectors separately gives us no insight into these matters, as both give insignificant coefficients.

The third interesting point about the domestic inter-industry spillover elasticity is that, added to the own R&D elasticity, it gives a national own R&D elasticity of around 9 or 10%. This is well within the bounds of estimates of previous macro level research (see, for example, Lichtenberg, 1992; Coe and Helpman, 1995). At least internally, there is evidence that the UK social rate of return to R&D is considerably higher than the private rate of return.

To sum up, the estimates for the elasticity of domestic inter-industry spillovers range from 0 to 7.6% (discounting the rather implausibly high elasticity obtained from the linear model). Perhaps correlation with other explanatory variables plays a part in the relatively high degree of sensitivity of the estimate of this parameter to model specification changes. However, at least for the standard model, the results are not inconsistent with the notion that intra-national spillovers are more important than international spillovers, for the UK at least.

5.8.5: Spillovers Through Foreign Inter-Industry Trade

Just as spillovers are possible from the same sector abroad and from different sectors domestically, so they are possible from different sectors abroad. Keller (1997), however, finds no significant effect of R&D in other sectors abroad on domestic own sector TFP. In the standard model presented in Table 5.4, I do find a (marginally) significant effect, at least at the 10% level. As for FDI-weighted and import-weighted foreign intra-industry R&D spillovers, the effect of foreign inter-industry spillovers appears to be *negative*. The results from the standard model suggest it is also *larger* than the effect of foreign intra-industry spillovers through goods trade. The estimated elasticity of TFP with respect to this variable is around –4.5%.

Given the results and arguments presented in Sections 5.8.2 and 5.8.3, at first glance the negative sign on the foreign inter-industry spillover variable comes as no surprise. Given the apparent importance of domestic inter-industry spillovers it is perhaps also no surprise that the inter-industry foreign trade effect should be greater than the intra-industry foreign trade effect. However, this is not necessarily the case. Although the technology-source argument is equally applicable to within and without-sector foreign R&D, the competition effect, which I hypothesise might partly explain the negative foreign spillovers, is not. Competition from other industries abroad is likely to be less fierce than competition from the same industry abroad, so we would expect a *smaller* negative effect, not a larger one. The argument that low tech sectors are expected to learn more from high tech sectors, discussed in Section 5.8.4, is not relevant to the possible *competition* effect.

The estimated elasticity of the foreign outside-sector R&D variable is not particularly robust to model specification changes, so perhaps the estimate in Table 5.4 should be taken with a pinch of salt. With the exception of the standard model without FDI[166], all other tests of robustness suggest the variable is insignificantly different from zero, which is perhaps more in line with our expectations and with the previous findings of Keller (1997). In summary, I find a range of elasticities between 0 and –5% for foreign inter-industry spillovers. At the very least, this provides further evidence of a lack of positive international spillovers in the UK.

5.8.6: TFP in UK Manufacturing

Cameron et al (1998) examine TFP in a panel of UK manufacturing industries over a similar sample period to that in the present study. They find a great deal of variety in TFP growth performance across industries. Cameron (1995) also finds a great deal of variation in the factors driving TFP movements in UK manufacturing industries. For example, human capital accumulation and de-unionization are found to be important explanations of TFP growth in many sectors. Technological progress, as measured by increasing R&D stocks, although widely seen as probably the most important component of TFP growth, is by no means the only source of TFP growth. In the current model, these other effects on TFP are captured by the time trend, but essentially left un-modelled.

Table 5.14 below outlines the estimated effects of domestic and foreign R&D, within and outside the sector, through trade and through FDI, on TFP growth in the sample of UK industries over the sample period 1973 to 1992. The beginning-to-end period growth of each variable is multiplied by its elasticity, as estimated by the standard model, to obtain a percentage figure for how much growth in TFP is driven by each

[166] The linear model gives an implausibly high elasticity, although barely significant at 10%, so I discount it.

variable for each sector.[167] Clearly, where the elasticity is negative and R&D growth is positive, the effect of R&D stock growth on TFP is negative. FDI does not enter the Table as it's effect is not statistically different from zero, even at 10% in the standard model.

From Table 5.14, it is clear that the R&D in the standard model explains a large part of TFP growth from 1973 to 1992 in the sample (around one quarter on average), although by no means all. The most important component of TFP growth, according to this model, is domestic inter-industry R&D spillovers. This is reinforced by own R&D and partially counter-acted by the negative incentive effects of foreign R&D. The explanatory power of R&D variables differs across sectors. For example, one quarter of the increase in TFP over the sample period in the food sector is explained by the technology variables, compared to just over half that for metal manufacturing. The importance of factors left un-modelled (captured by the trend variable) is not inconsistent with the pattern found by Cameron (1995). Cameron finds de-unionization to be an important factor in TFP growth in many sectors.

Table 5.14: Deconstructing TFP Growth with the Standard Model

Sector	**% ΔTFP 1973-92**	**Own R&D**	**Dom oth R&D**	**For R&D**	**For oth R&D**	**Net %ΔTFP by model**
Food	+23%	+1.5%	+20.5%	-4%	-12%	+6%
Paper	+28%	+1.5%	+20.5%	-4%	-12%	+6%
Chem	+38%	+4.5%	+20.5%	-4%	-12%	+9%
Metal	+23%	+0.5%	+19%	-4%	-12%	+3.5%
Mech	+41%	+3.5%	+19%	-4%	-11%	+7.5%
Elec	+33%	+3.5%	+21%	-3%	-13%	+8.5%
Trans	+26%	+1%	+19%	-4%	-11%	+5%

Notes: Figures given are elasticities multiplied by 1973 to 1992 percentage change in weighted R&D stocks. Elasticities are constrained to be the same across sectors in the standard model.

Of course, such exercises as that presented in Table 5.14 above must not be taken too seriously. There are alternative elasticities suggested by alternative specifications of

[167] In this model, the elasticities are constrained to be the same for all sectors. Variation only arises from the different rates of growth of R&D stocks for the different sectors.

the model that would give completely different figures. Although some of the elasticities fall in a fairly narrow range (own R&D, for example), others fall in a wider range (FDI and domestic inter-industry spillovers, for example). Nonetheless, the results paint a highly consistent picture, at least as far as the *signs* of spillover effects are concerned. That in many cases the magnitudes of these effects seem plausible should perhaps be seen as an added bonus, given the data limitations and econometric difficulties encountered.

5.9: Summary and Concluding Remarks

In this Chapter, I examine knowledge spillovers in a panel of UK manufacturing industries, allowing for technological transmission through direct investment in addition to the usual trade in goods. The incorporation of foreign direct investment (FDI) in studies of international spillovers has been widely suggested, but so far little research has been able to do so, with a few notable exceptions. Perhaps the most important factor holding back such research is the difficulty in obtaining suitable data.

Although far from ideal, I have data on FDI into and out of the UK for 7 manufacturing sectors in addition to total manufacturing. These sectors are not necessarily representative of the UK manufacturing sector as a whole.[168] The volatility of this FDI data and the prevalence of negative observations mean I am unable to specify inward and outward UK FDI separately, but am able to specify a variable for total UK FDI from which average FDI shares of 13 OECD countries can be derived.[169] Marrying this FDI data with data on capital stocks, R&D and goods trade is not straightforward, due primarily to the need for transformations between alternative sectoral and product definitions (ie: SIC, SITC and ISIC). This, coupled with the fact that I am only able to specify TFP series for the UK sectors that assume perfect competition and constant returns to scale (see Appendix C), leaves me not entirely satisfied with the accuracy of the data used in the estimation.

Nonetheless, although undoubtedly important, these potential data inaccuracies do not present an insurmountable obstacle to the research. I have been able to estimate an

[168] There are no textiles, petrol or rubber and plastics sectors in the sample, for example.

empirical model, based explicitly on theory, and obtain results that are not only plausible, but also very interesting. Improvements in the data series on which the estimations are based is something I leave for further research.

On top of the slight unease about the accuracy of the data, there is the thorny problem of the high degree of correlation between many of the explanatory variables used in the analysis. These variables are all measuring R&D stocks in the UK and abroad, weighted by goods trade or FDI, so it is perhaps unsurprising that they are highly correlated with each other given the general smooth upward trend that such stocks follow. If we are to specify a model such as the one specified here, these problems with multicollinearity are likely to be unavoidable. This may be another factor that has held back the pace of research on the issues under consideration in this Chapter.

The multicollinearity problem prevents me from estimating a number of alternative specifications of the model, such as lagged explanatory variables or all sectors individually. The difficulty lies with the lack of convergence in the iterative procedure used to estimate the non-linear model. Nonetheless, although I cannot estimate all the variants of the model that I would like, I am able to estimate a standard model, along with a number of alternative specifications as tests of robustness. In many cases, the estimates are remarkably robust to specification changes considering the high degree of correlation between some of the explanatory variables, which adds a little confidence in their reliability.

In many respects, the results are somewhat counter-intuitive. In particular, they show foreign technological progress to have a *negative* effect on productivity in UK manufacturing, in most cases. This result is highly robust to a variety of specification changes. One possible explanation is that the UK's position as a leading technology source country makes it particularly vulnerable to *outward* spillovers, which might act as a disincentive to invest in technological advancement. Equally, the UK is unlikely to learn much through spillovers from technologically inferior economies. Alternatively, increased competition with technologically superior countries such as the US and West Germany might have a disincentive effect on investing in R&D

[169] Negative observations make the log of FDI unspecified.

which is unrelated to spillovers, but which dominates the positive spillover effect in the data. This negative foreign spillover result is consistent for spillovers through trade, both intra and inter-industry, and spillovers through FDI. UK engineering sectors are the only exceptions, with a positive FDI-transmitted spillover effect on TFP. FDI and trade are found to be complementary transmission mechanisms for technological knowledge rather than substitutes.

In other respects, the results conform to expectations and the findings of previous research. For example, own R&D increases own TFP, with a plausible elasticity. Domestic inter-industry spillovers are found to be very important in driving TFP growth. Intra-national spillovers are generally found to be more important than international spillovers in driving TFP growth.

To conclude, despite data difficulties and problems with multicollinearity, I am able to estimate a model of industry level spillovers, through trade and through FDI, for a panel of UK manufacturing sectors, for which I obtain some very interesting results that appear to be reasonably reliable. There is little evidence of positive international spillovers to these UK sectors and some evidence of *negative* international spillovers, through both trade and FDI. This is likely to reflect some sort of negative competition effect on R&D incentives which dominates any positive spillover effect in the data.

In terms of policy implications, if the negative result for foreign spillovers obtained in the research is correct, and it seems robust, then we clearly need to view increasing economic integration through trade and increased FDI in some sectors as potential threats to long-term industry growth. However, it would appear that this is not the case for all sectors, as can be seen from the positive FDI-transmitted spillover effects found for engineering industries. The sectors most at risk from negative foreign technology externalities seem to be those at the lower end of the technology spectrum, although the current sample is too small to say this with any level of certainty.

Future research on these issues should help us reach a more fundamental understanding, from which we would be in a better position to discuss implications for policy. An obvious avenue for further research is to construct improved data series to check the reliability of the results presented here. Also, the sample is not necessarily representative of UK manufacturing as a whole, therefore data on other sectors would allow us to test whether these results are peculiar to the current sample of sectors or indicative of UK manufacturing in general. Estimating the model assuming the series are I(1) is a good place to start with further research.

Chapter 6

Concluding Remarks

This thesis argues that openness to trade plays a major role in economic growth through technological progress. This is not a new idea, but one that has not yet been exhaustively explored by the literature. The survey Chapter identifies a number of questions that are still unsettled. Three of these questions are studied in greater depth in the remaining Chapters. Given the abundance of research concerned with economic growth and international trade that is built on technological progress foundations, I argue that it is important to learn as much as possible about the inter-relationships of R&D, trade and growth if we are to take our understanding forward with confidence. In a world of increasing globalisation and economic integration this field of research is given all the more urgency.

Chapter 3 sets out to examine the links between openness to trade and R&D incentives through the mechanism of the intensity of competition. In doing so, a model is presented in which economic integration can influence the long run rate of growth purely through such competition effects. The relationships between R&D and the intensity of competition or market structure more generally have been widely studied. However, the Chapter makes a clear contribution to the literature by showing how such relationships can increase growth in the context of economic integration, by providing an alternative to the usual scale effects story, for which empirical evidence is inconclusive. The main weakness of the Chapter is the specialised nature of the model. Further research is needed to generalise the results.

Chapter 4 sets out to search for *direct* evidence of a relationship between R&D incentives and openness to trade. Although this relationship forms the background to a significant body of research in the growth and trade literature, I argue that such evidence has not yet been sufficiently established in the literature. Given the significance of the R&D/openness relationship and the potential ambiguities it displays, this evidence is needed. A methodology for testing the relationship against the empirical evidence is suggested and the first steps are taken. Despite the significant nature of the objective, and despite complications with estimation, some evidence is provided in support of the existence of a positive relationship between R&D incentives and openness to trade relationship at industry level. This relationship is not unambiguously positive, however. Also, the conclusions drawn are not unquestionable, given the problems encountered with the persistence of the R&D expenditure series, amongst other things. There is a clear need for further research to robustify the results.

Chapter 5 sets out to study the nature of international knowledge spillovers in UK manufacturing. Two arguments in particular motivate the study. Firstly, there is some evidence to suggest that spillovers might have different effects on productivity in different countries or different sectors. Secondly, there is a general consensus that alternative transmission mechanisms to goods trade, such as FDI, are significant players in the spillover of technical knowledge from one country to another. The Chapter finds evidence of significant spillovers through FDI, thus contributing support to those papers that suggest such an effect should exist. The Chapter also finds evidence suggesting a possible negative productivity effect from foreign technological progress. This contributes to the literature arguing that spillovers are heterogeneous and that they may not always be positive, or may be obscured by other non-spillover relationships in the data. However, problems with multicollinearity and possible simultaneity leave a great deal of room for further research to explore these issues in more depth.

6.1: Suggestions for Further Research

A number of suggestions for further research are made throughout this thesis. Indeed, one of the most exciting characteristics of the research has been the sense of a clear need for further study along the avenues identified. In this respect, I am glad to have had the opportunity to look at three related, but separate questions, rather than having concentrated in more detail on just one. Of course, this has not come without cost, and the tentative nature of many of my conclusions reflects this.

Chapter 3 concludes with a number of suggestions for further research to generalise the model. Relaxing the assumption of a fixed number of firms must be first among these suggestions. The different relationships between concentration and R&D in under alternative forms of competition that have been found in the IO literature suggest the need for such alternatives to be considered in the context of the model presented in Chapter 3. Allowing the degree of differentiation between products to be endogenous is also an obvious avenue for further research. This would necessitate an alternative method of modelling integration, perhaps following the tariff approach of Appendix A. There is enough potential in the model and its conclusions to warrant at least this further research.

Chapter 4 is very much a first step along a fairly clear research avenue. With the existing data set, the assumption of stationarity of the R&D series and some of the other series needs to be relaxed and the possibility that the series form a cointegrating set explored. With the recent arrival of a number of panel cointegration analyses in the literature, such a study should not only be possible but also a useful addition to the application of this methodology. In the long term, however, it is by exploiting richer data, perhaps at firm level, that this research is likely to be carried forward most successfully. In a sense, Chapter 4 identifies the problem for further research to answer.

Chapter 5 is similar to Chapter 4 in that it provides a first attempt at incorporating FDI into a sectoral-level study of spillovers building directly from the macro spillovers literature. Again, there are immediate issues that require further research, such as the ambiguity of the stationary or non-stationary status of some of the series. Perhaps the over-riding need, however, is for further research into the questions of the role of FDI and the heterogeneous nature of spillovers with an alternative methodology. It is doubtful whether some of the problems identified in Chapter 5, such as the high degree of multicollinearity between R&D stocks, will ever be satisfactorily overcome. An alternative approach to what are very interesting issues is needed.

I argue broadly in this thesis that it is necessary to back up and fill in some of the gaps in the foundations of the trade and growth literature before we can be truly confident that our understanding has developed as much as it could have in the last decade. A great number of unanswered questions remain at this level, and numerous opportunities for further research exist. Perhaps after the rapid development of the growth literature in the last ten years, such an exercise in reflection is needed alongside what could be further rapid development for some years to come.

6.2: Policy Recommendations?

The exploratory nature of much of the research in this thesis and the tentative nature of my conclusions leaves me unable to make any firm policy suggestions. A great deal of further research is needed before such suggestions can be made. However, just as a quick thought experiment, making the substantial assumption that the tentative results in this thesis hold up to further detailed analysis, what would the research suggest?

Firstly, in terms of stimulating economic growth, economic integration is generally a *good thing*, but not in all scenarios. Chapter 3 suggests the possibility that there might be some optimum level of integration, that once passed, could reduce incentives to invest in R&D. Chapter 4 suggests the possibility of a negative effect of openness to trade on R&D expenditure, at least for the food industry. Chapter 5 finds some evidence of a negative effect of foreign technological progress transmitted through imports and FDI on domestic productivity, at least for some manufacturing sectors.

Taken together, the research suggests policy makers should consider the likely costs and benefits of further integration on a *case by case* basis, at least at industry level. This is a necessarily broad suggestion, but particular examples can be drawn from it. For example, government help to attract FDI into the UK may not always be the best way to spend limited resources. The likely benefits in terms of employment should be compared with possibly detrimental productivity effects in those sectors where such negative effects may be present.

Appendix A

R&D in a Two-Country Differentiated Duopoly with Tariff Reduction

The model below is an alternative version of the model outlined in Section 3.3. I do not present any separate motivation or background for this exercise here.

Integration is frequently modelled as an explicit reduction in trade barriers, such as tariffs (see, for example, Venables and Smith, 1986; Driffill and Van Der Ploeg, 1993). There is some question as to the relevance of this approach to the modern developed world, particularly within regional trade blocs such as the EU, with internal tariffs having been set at zero for many years. However, given the non-standard method of modelling integration adopted in Chapter 3, it is a worthwhile exercise to test the general robustness of the model's predictions using a more traditional modelling technique for integration. Much of the detail given in Section 3.3 is skipped over here, but the equations are reworked from the beginning (equation (3.4)). Only the symmetric competitive scenario is considered explicitly.

Consider an industry consisting of two firms, each producing a single differentiated good. Assume that entry into the market is impossible, due to, say, high fixed costs that have been paid in the past by the incumbents. The first firm is located in the first of the two countries (Country A) and the other is located in the second country (Country B). I consider only similar countries, so avoiding any question of comparative advantage. In contrast to the model of Section 3.3, the two countries are explicitly modelled in this case.

A symmetric tariff, t, is imposed on imports in both countries, the revenues from which vanish into the void.[170] Integration is modelled as a decrease in the level of the tariff, which is assumed to be exogenous. A firm's output is separated into domestic and foreign components and each market is treated separately for the purpose of profit maximization, since production costs are constant wherever output is destined. In what follows, price is given as what the firm receives, so the tariff is levied on the consumer.

Identical consumers across both countries have Cobb-Douglas preferences over a composite of the R&D goods and a non-R&D competitive good. They therefore spend a fixed proportion of their income, m, on the two goods, 1 and 2, which they do according to the constant elasticity of substitution (CES) utility function given by equation (A1).

(A1): $U(q_k) = \{q_{1k}^{(\theta-1)/\theta} + q_{2k}^{(\theta-1)/\theta}\}^{\theta/(\theta-1)}$, for all $\theta>1$, for consumer k.

Since the model is symmetric, I concentrate on Country A. Consumers divide their lump of income, m, between the two goods, giving us a representative Country A consumer's demand for each good, as given by equation (A2), where $T= (1+t)$.

(A2): $q_{A1k}(p,m,T) = mp_{A1}^{-\theta}/(p_{A1}^{1-\theta} + (Tp_{A2})^{1-\theta})$,

and $q_{A2k}(p,m,T) = m(Tp_{A2})^{-\theta}/(p_{A1}^{1-\theta} + (Tp_{A2})^{1-\theta})$,

The demand curves for Country A (A3) facing each firm can be found by summing over all consumers, where 'a' denotes the aggregate of the m's. The countries are of equal size, so there is a/2 income available in each market for the two goods.

(A3): $q_{A1} = ap_{A1}^{-\theta}/(p_{A1}^{1-\theta} + (Tp_{A2})^{1-\theta})$,

and $q_{A2} = a(Tp_{A2})^{-\theta}/(p_{A1}^{1-\theta} + (Tp_{A2})^{1-\theta})$.

The corresponding expressions for Country B have the tariff term imposed on good 1. They are otherwise identical.

[170] This is just for simplicity and is OK as the model is a partial equilibrium.

Unit costs of production and the specification of R&D are unchanged from Section 3.3 and are given by (A4), (A5) and (A6), where $f(X_i)$ is the cost reduction resulting from firm i's effective R&D:

(A4): $C_i(q_i) = c_i q_i$.

(A5): $X_i = x_i + \beta x_j$,

(A6): $f(X_i) = c_i X_i^{1/2}$.

At this stage I make the simplifying assumption of setting θ=2. This does not change the model in any way, but makes the maths more straightforward. In this model, θ is just assumed to be exogenous and invariant. Inverting the demand curves (A3) gives the expression for total revenue for each firm from Country A (A7).

(A7): $TR_{A1} = a/\{2[1 + \sqrt{(q_{A2}/q_{A1})}]\}$,

and $TR_{A2} = a/\{2T[1 + \sqrt{(q_{A1}/q_{A2})}]\}$.

Firms play a two-stage game. First they determine R&D expenditure, x, which determines unit costs for stage 2 in which there is Cournot competition. Given R&D, the Country A marginal condition for each firm in Stage 2 is given by (A8) and this defines profit maximising output (A9).

(A8): $c_1(1-X_1^{1/2}) = [a(q_{A1}/q_{A2})^{-3/2}] / [4q_{A2}\{1 + \sqrt{(q_{A2}/q_{A1})}\}^2]$,

and for firm 2, $c_2(1-X_2^{1/2}) = [a(q_{A2}/q_{A1})^{-3/2}] / [4Tq_{A1}\{1 + \sqrt{(q_{A1}/q_{A2})}\}^2]$.

(A9): $q_{A1} = [a\sqrt{(TF(X_1))}] / [4(c_1(1-X_1^{1/2}))\{1+\sqrt{(TF(X_1))}\}^2]$,

and $q_{a2} = TF(X_1)q_{A1}$,

where $F(X_1) = [c_1(1-X_1^{1/2})/c_2(1-X_2^{1/2})]$.

The output expressions for Country B are symmetrical, and for firm 1, profit maximizing output in Country B is given by (A10):

(A10): $q_{B1} = [a\sqrt{(TF(X_1))}] / [4Tc_1(1-X_1^{1/2})\{1+\sqrt{(TF(X_1))}\}^2]$,

Total profits for firm 1 are the revenues from each market less the production and research costs, given by (A11).

(A11): $\Pi_{1total} = [a/2T][(1+T)/(1+\sqrt{(TF(X_1))})]$ -
$[a/4T][(1+T)\sqrt{(TF(X_1))}]/[1+\sqrt{(TF(X_1))}]^2 - x_1$.

The expression for firm 2 follows from the symmetry assumption.

In Stage 1, the firms set the profit maximizing level of R&D expenditure, x_1 and x_2, treating the other firm's R&D as given. Given that the model has been set up with symmetry maintained throughout, there is nothing to make one firm's R&D decision any different to the other's, so $x_1 = x_2$ in equilibrium. The first order conditions give us the following expression for optimal effective R&D:

(A12): $X_i^{1/2} = 1/2 \pm \{\sqrt{[1 - a(1-\beta)\tau]}\}/2$,

where $\tau = (1+T)(3+4\sqrt{T}+T)/4\sqrt{T}(1+\sqrt{T})^4$.

Second order conditions rule out the larger of these two solutions, as in Section 3.3.1. So, the smaller of the two solutions denotes profit-maximizing effective R&D level. The condition stated in Section 3.3.1 to ensure a non-negative expression in the square root applies here also. Since $T \geq 1$, $\tau \leq 1/4$, so for $\theta=2$, all values of $a \leq 4$ are valid.[171]

From (A12), with a little manipulation, it can be seen that:

[171] $\theta \leq a/(a-2)$. So, for $\theta=2$, $a \leq 4$ to ensure that $1 \geq a(1-\beta)\tau$, for all β,τ.

(A13): $\partial(X_i^{1/2})/\partial T < 0$ and that $\partial(X_i^{1/2})/\partial\beta < 0$.

In other words, effective R&D is decreasing in tariffs and decreasing in spillovers. Integration, by lowering tariffs, therefore leads to a higher level of expenditure on R&D, just as in the model in Section 3.3.1. Intuitively, extending the model to the scenarios of Sections 3.3.2 and 3.3.3 will also give similar results, as the outcomes are still driven by the same positive strategic and negative spillover effects as before.[172] In conclusion, the predictions of Section 3.3 are therefore not fragile to using falling tariffs as an alternative measure of integration.

[172] One small difference is that because tariff revenues are not re-distributed in this model, reducing tariffs does have a small market-enlargement effect over and above the strategic effects of integration.

Appendix B

Data and Other Details for Chapter 4

Data Sources and Construction

There were 3 data sources used to assemble the final data set. Firstly, OECD National Accounts from various years. Secondly, the OECD, DSTI (ANBERD database), 1997. Finally, the OECD, DSTI (STAN Industrial Database), 1996. Sources (and definitions) of individual variables are listed below.

i) RRD: Real expenditure on research and development, 1990 $US. R&D expenditure taken from ANBERD. Purchasing Power Parities (PPPs) taken from STAN. Adjusted for US producer prices, taken from Statistical Abstract of the US.
ii) OP: Openness to trade index, defined as imports plus exports over production. All data from STAN.
iii)RPR: Real production, 1990 $US. Production figures from STAN. PPPs from STAN. Adjusted for US producer prices.
iv)RI: Real gross capital formation, 1990 $US. Gross capital formation from STAN. PPPs from STAN. Adjusted for US producer prices.
v) GR: GDP growth, year on year, 1990 prices, per capita. Growth and price level data taken from OECD National Accounts.
vi)PROF: Defined as value added minus labour costs over capital stock. Value added and labour costs from STAN. Capital stock estimated from STAN gross capital formation data, constructed by the perpetual inventories method. Details of this method are presented in Chapter 5 and Appendix C. Depreciation was taken as 3%.

Coverage

The data cover 13 industrial sectors across 14 countries from the OECD ANBERD and STAN Industrial Databases. The countries are Australia, Canada, Denmark, Spain, Finland, France, (West) Germany, Italy, Japan, Netherlands, Norway, Sweden, United Kingdom and the USA. The sectors are listed below with their ISICs.

Table B1: Sectors Covered by Chapter 4 Data Set

Sector	ISIC	Description	Sample Size (T*N)
1	3000	Total manufacturing	20*14=280
2	3100	Food, beverages and tobacco	20*14=280
3	3200	Textiles, apparel and leather	19*14=266
4	3300	Wood products and furniture	20*14=280
5	3400	Paper products and printing	20*14=280
6	3500	Chemical products	20*14=280
7	3512[1]	Chemicals excluding drugs	19*10=190
8	3522	Drugs and medicines	19*10=190
9	3534[2]	Petrol refineries and products	18*6=108
10	3556[3]	Rubber and plastics products	19*13=247
11	3825	Office machinery and computers	17*13=221
12	3843	Motor vehicles	15*12=180
13	3845	Aircraft	16*7=112

Notes: 1=3510+3520-3522, 2=3530+3540, 3=3550+3560.

Some Summary Statistics

The tables below report means across time for total manufacturing by country and for all sectors across countries.

Table B2: Means by Country, Total Manufacturing, 1990 $US

Country	RRD	OP	RI	RPR	PROF	GR
Aus	$6.46*10^8$	.397	$4.50*10^9$	$9.46*10^{10}$	.324	.028
Can	$1.82*10^9$	.656	$1.09*10^{10}$	$2.06*10^{11}$	.243	.031
Den	$3.67*10^8$	.931	$1.83*10^9$	$3.47*10^{10}$	.196	.018
Esp	$8.83*10^8$	.290	$7.13*10^9$	$2.57*10^{11}$	.957	.028
Fin	$4.51*10^8$	.594	$2.58*10^9$	$4.15*10^{10}$	.233	.023
Fra	$9.11*10^9$	.477	$2.67*10^{10}$	$4.94*10^{11}$	.288	.023
Ger	$1.53*10^{10}$	.468	$3.53*10^{10}$	$7.92*10^{11}$	.239	.025
Ita	$3.80*10^9$	.416	$3.11*10^{10}$	$5.14*10^{11}$	.321	.028
Jpn	$2.54*10^{10}$	.169	$9.29*10^{10}$	$1.39*10^{12}$	.152	.038
Nld	$1.88*10^9$	1.12	$6.10*10^9$	$1.20*10^{11}$	.106	.024
Nor	$3.21*10^8$	.798	$1.56*10^9$	$2.84*10^{10}$	.148	.034
Swe	$1.72*10^9$	.712	$4.04*10^9$	$7.70*10^{10}$	.274	.017
Ukm	$8.87*10^9$	.498	$2.01*10^{10}$	$5.20*10^{11}$	.188	.019
Usa	$7.17*10^{10}$	.197	$1.04*10^{11}$	$2.56*10^{12}$	.279	.024

Table B3: Means by Sector, Across Countries

Sector	RRD	OP	RI	RPR	PROF
1	$1.02*10^{10}$	.552	$2.49*10^{10}$	$5.10*10^{11}$	.282
2	$1.90*10^8$	.255	$2.58*10^9$	$7.65*10^{10}$	.424
3	$5.17*10^7$	.937	$1.06*10^9$	$3.29*10^{10}$	.111
4	$2.76*10^7$	.381	$8.61*10^9$	$2.03*10^{10}$	.207
5	$1.04*10^8$	.316	$2.25*10^9$	$4.00*10^{10}$	.272
6	$2.23*10^9$	.635	$5.00*10^9$	$9.71*10^{10}$	.188
7	$1.05*10^9$	.794	$7.41*10^9$	$4.05*10^{10}$	.163
8	$7.26*10^8$	.574	$4.23*10^8$	$7.70*10^9$	.371
9	$9.91*10^7$	.543	$6.37*10^8$	$1.91*10^{10}$	.173
10	$1.77*10^8$	.499	$9.61*10^8$	$1.71*10^{10}$	.161
11	$9.58*10^8$	2.30	$5.57*10^8$	$9.08*10^9$	.515
12	$1.59*10^9$	1.31	$2.92*10^9$	$5.43*10^{10}$	.137
13	$3.60*10^9$	1.02	$6.96*10^8$	$2.11*10^{10}$	.146

Autocorrelation Tests

The LIMDEP package estimates ρ (the autocorrelation coefficient) as part of the LSDV estimation, first assuming common ρ across countries and then allowing ρ to vary across countries. Assuming a common value for ρ, is probably better than group-specific ρs for such short time series (see Greene, 1993, p457-458). Greene (1991) suggests the following informal test for autocorrelation:

(B1): $(T-1)r^2/(1-r^2) = \chi^2_{(1)}$,

where 'r' is the estimate of ρ. Running this test on the demeaned data for all sectors (assuming a common ρ across countries) suggests significant 1st order autocorrelation for all sectors at a 5% significance level and for some at a 1% significance level. Estimated autocorrelation coefficients range from .18 (sector 8) to .48 (sector 11). The LSDV estimation is therefore autocorrelation-robust (see Greene, 1991, for details of LIMDEP's LSDV estimation procedure). There is also sufficient evidence to question the validity of instrumenting with lags in the AH and GMM estimation. Together with persistence-related problems and inefficiency, these estimators are therefore dropped from the study (see Section 4.7.3).

Some Other Diagnostics

A number of diagnostic tests not detailed in the main text are briefly reported here. For the LSDV1 estimator, I report the Wald test for groupwise heteroskedasticity and the likelihood ratio (LR) test for cross-section correlation, carried out automatically under LIMDEP's 'TSCS' command.

The LIMDEP TSCS command (used to estimate the LSDV models) begins under the assumption of groupwise heteroskedasticity and cross-section correlation and then tests the validity of these assumptions. First, the null hypothesis of no cross-sectional correlation is tested for (ie: the off-diagonal elements of the covariance matrix of the error term at time t) with the following LR test statistic:

$\lambda_{LR} = T(\sum_i \ln\sigma_i^2 - \ln|\Sigma_t|)$.

The large sample distribution of the statistic is chi-square with n(n-1)/2 degrees of freedom. Table B4 shows the test results for all 13 sectors in the LSDV1 model. Cross-sectional correlation is significant in all sectors. Second, groupwise heteroskedasticity is tested for with the following Wald test:

$W = \sum_i (s_i^2-\sigma^2)2/Var[s_i^2] \rightarrow \chi^2_n$ under the null of no heteroskedasticity.

Results for this test are given in Table B4. In all cases, the null is rejected. For further details of these tests see Greene (1993).

Table B4: LR and Wald Tests

Sector	LR	Wald
1	179**	1473**
2	170**	212**
3	149**	649**
4	185**	2685**
5	174**	27865**
6	142**	1969**
7	69**	218**
8	61**	112**
9	13.5**	223**
10	137**	357**
11	149**	657**
12	145**	10933**
13	23.1**	25.8**

Notes: Significant rejection at 95% of the null of (i) valid instruments, (ii) no cross-section correlation and (iii) no groupwise heteroskedasticity is denoted by **.

Appendix C

Data for Chapter 5

Data is annual, between 1973 and 1992, for the broad UK manufacturing sectors listed in Table C1 below. The sectors correspond to those for which UK FDI data is available. The full sample of countries (for foreign spillovers) is: Australia, Canada, Denmark, Spain, Finland, France, (West) Germany, Italy, Japan, Netherlands, Norway, Sweden, UK, US.

Table C1: Manufacturing Sectors

Sector	Sector Name	SIC84	SITC70	ISIC	ISIC Names
1	Total manufacturing	2-4	All comms	3000	Total manufacturing
2	Food	41,42	00-12	3100	Food, beverages and tobacco
3	Paper	47	64	3400	Paper
4	Chemicals	25	5	3500	Chemicals
5	Metal Manufacturing	22	67-69	3700,3810	Basic metal industries + Metal products
6	Mechanical engineering	32	71	382X	Non-electrical machinery
7	Electrical engineering	34	72	3825, 383X, 3832	Electrical machinery + Office equipment + Communications equipment
8	Transport equipment	35,36	73	3841, 3843, 3845, 3842A	Ships + motors + aircraft + other transport equipment

Notes: The SIC classifications are based on SIC84. The ISIC classifications are based on the OECD 26 Industry Adjusted ISIC Classification from the STAN database. Special codes are defined in the STAN accompanying notes (OECD, DSTI, STAN Industrial Database, 1996).

C1: Variable Definitions and Construction

A large number of variables are used in the statistical and econometric analysis, although many are variants of the same series. All of the series required some construction and some of the series required a lot of construction. In this section, I define each series and discuss construction and sources.

Total Factor Productivity

Growth in total factor input is given by the Thornqvist index of total inputs, I:

$$\text{Log}(I_{jt}/I_{jt-1}) = .5*[\beta_{jt} + \beta_{jt-1}]\log(N_{jt}/N_{jt-1}) + .5*[(1-\beta_{jt}) + (1-\beta_{jt-1})]\log(K_{jt}/K_{jt-1}).$$

TFP growth is given by output growth (growth in Value Added) minus input growth. TFP levels are expressed as index numbers by setting a value of 100 for each sector in 1973. N is number of workers engaged, K is capital stock and β is the share of labour costs in production.

There are two alternative variables for total factor productivity, one with time invariant labour shares in production and one with time-varying labour shares. The time invariant labour share is the average of the sample period labour shares. This is the preferred specification because of volatility in the time-varying labour shares, probably down to measurement error. The labour share is the revenue-based factor share of labour, defined as:

$$\beta = wN/pQ,$$

where wN is the labour costs of production and pQ is value added (see Hall, 1990). Keller (1997) uses an alternative cost-based factor share measure, which should be more robust to the presence of imperfect competition (see Keller (1997) for discussion). Unfortunately, data limitations constrain me from doing the same.

Sources:

Value Added (VA): STAN (OECD, DSTI, STAN Industrial Database, 1996).

K: See below.

N: STAN.

wN (labour costs): STAN.

Capital Stock

Capital stock is constructed following the standard perpetual inventories method, as in Keller (1997). Following Keller, I use a depreciation rate of 8.19%, as estimated for manufacturing and machinery in the UK in 1980 by Jorgenson and Landau (1993). Capital stock in year t is defined as:

$$K_{jt} = (1-\delta)K_{jt-1} + I_{jt-1},$$

where I is real gross investment in physical capital, deflated by a standard UK Producer Price Index (PPI). Initial capital stock is given by:

$$K_{j0} = I_{j0}/(g_j + \delta),$$

where g is the average annual growth rate of investment, I, over the sample period.

Sources:

I: STAN.

PPI: Economic Trends, various years, HMSO, London.

Domestic Within-Industry R&D Stock

Domestic within-industry R&D stock is constructed following the perpetual inventories method, as for physical capital stock above. All other R&D stocks for other domestic industries and foreign industries are constructed in the same way. Two alternative depreciation rates for knowledge capital are used, $\delta = .1, .2$, to capture the variation in depreciation-rate estimates (see Keller, 1997, for a discussion). R&D stock at time t is defined as:

$n_{jt} = (1-\delta)n_{jt-1} + \phi_{jt-1}$,

where ϕ is real expenditure on R&D (deflated by UK PPI). Foreign R&D expenditures are transformed into £ Sterling by Purchasing Power Parity (PPP). Initial R&D stocks are given by:

$n_{j0} = \phi_{j0}/(\lambda + \delta + 0.1)$,

where λ is the average annual growth rate of R&D expenditure over the sample period. I add 0.1 to the denominator to ensure positive initial stocks (following Keller, 1997).

Sources:
ϕ: ANBERD (OECD, DSTI, ANBERD Database, 1997).
PPP: STAN.

Import-Share Weighted Foreign R&D Stock, Within-Sector

The R&D stocks for sector j in all other countries are constructed as above, converted to £ sterling using PPPs. There are four import-share weighted R&D series constructed in the same way, with the two alternative depreciation rates and with time varying and time-invariant import shares. Time-invariant import shares are taken as the average of the import-shares over the sample period. The series are defined as follows:

$b_{jt}^{f} = \Sigma_{i\neq UK}\ \omega_{it}b_{it}^{d}$,

where b_{it}^{d} is country i's R&D stock in the given sector at time t. The weights are given by ω_{it}, which denotes the share of total (sample) imports from the country i in the given sector at time t.

Data on UK imports over the sample period, from specific countries, is available in 3-figure SITC classification (ie: by broad commodity). I first transform later SITC classifications into 1970 SITC classifications. This is then converted into the 1984 SIC sectors as shown in Table C1 above. Figures are in £ sterling. Import shares are calculated as the ratio of sector imports from country i to the total imports from all sampled countries in each period. The shares thus sum to one.

Sources:
Foreign R&D stocks: Constructed as above from ANBERD R&D expenditure data.
Imports: Overseas Trade Statistics of the UK, 1973-1992, HMSO, London.

Import/Production Ratios (m) and FDI/Production Ratios (f)

In country level studies, such as Coe and Helpman (1995), a variable is typically included to capture the *level* of imports as this is not reflected by the import-share weighted R&D stocks (the shares sum to one). In Coe and Helpman (1995) this is defined as the ratio of total imports to GDP for a given country. In the present sectoral-level study I define this variable as the ratio of total imports to value added in the given sector. Imports totals are derived as above. Value added data sources are as above.

Given that the FDI-weighted foreign R&D stocks are constructed in the same way as the import-share weighted series (see below), the argument for including a variable to capture levels of trade is equally applicable to levels of FDI. I construct such a variable defined as the ratio of sectoral FDI to sectoral value added.

Domestic Input-Weighted R&D Stocks, Outside-Sector

R&D stocks for UK sectors are constructed as described above. Input weights are taken from the 1984 UK Input/Output Tables by SIC Industry. Input/Output Tables are not published every year (for the UK they are published roughly every 4 or 5 years). I use the 1984 Table as it is the closest to the middle of the sample period. Input-shares are therefore time-invariant. There are two series, corresponding to the two alternative knowledge depreciation rates. The series are constructed in a similar way to the import-share weighted R&D stocks discussed above. Thus:

$$b_{jtio}^{d} = \Sigma_{i \neq j}\, \omega_{it} b_{it}^{d},$$

where, in this case, the weights represent input-share as opposed to import-share. The subscript i denotes sector i.

Sources:
Input shares: UK Input/Output Tables, 1984, HMSO, London.

Foreign Outside-Sector Import-Use Weighted R&D Stocks

R&D stocks and import-shares are constructed as described above. To capture the inter-industry nature of the flow of goods I sum the foreign outside-industry import-weighted R&D stocks with weights determined by import-use data from the 1984 UK Input/Output Tables. The weights are given by the share of sectoral import-use in total imports from the outside-sectors, or in other words, sector j's share of purchases of sector i's imports. Import-use data is defined by broad commodity group as for imports and is therefore first transformed into the SIC84 classification. There are two alternative series corresponding to the two alternative depreciation rates, defined as:

$$b_{jtiu}^{f} = \Sigma_{industry\ i \neq j}\ \Sigma_{country\ c \neq UK}\ \omega_{jit}\ \omega_{ict}\ b_{ict}^{d},$$

where ω_{jit} and ω_{ict} represent import-use share of industry j for industry i's imports and import-share of country c for the UK's imports from industry i respectively.
Sources:
Import-use shares: UK Import-Use Tables, 1984 Input/Output Tables, HMSO, London.

Foreign-Direct Investment

FDI is defined by the Office for National Statistics (ONS) in the following passage:

The term 'direct investment' defines a group of transactions between enterprises, usually companies, that are financially and organisationally related and are situated in different countries. Such related enterprises – 'affiliates' – comprise subsidiaries, associates and branches...Direct investment refers to investment that is made to add to, deduct from, or acquire, a lasting interest in an enterprise operating in an economy other than that of the investor and which gives the investor an effective voice in the management of the enterprise. Other investments in which the investor does not have an effective voice on the management of the enterprise (ie: the investor has less than 10% of the voting shares) are regarded as portfolio investments. The estimates of direct investment include the investor's share of the reinvested earnings of the subsidiary or associated company, the net acquisition of equity capital, changes in inter-company accounts and changes in branch/head office indebtedness.

(MA4 Business Monitor, 1997, ONS, London).

From this definition, we can write down an equation for investment flows, both inward and outward:

Investment flows = reinvested profits + acquisitions of share and loan capital – disposals of share and loan capital + net inter company account change + net branch/head office indebtedness changes.

I have data for net flows of inward and outward FDI to and from the UK, by country, between 1973 and 1992 for the 8 sectors described in Table C1. This data is very volatile, with large swings from year to year in many cases, and often large negative values. There are also a significant number of missing values in this data, which have had to be replaced by sectoral sample means. These data problems lead me to favour time-invariant FDI-shares over time-varying shares, although both are calculated.

Sources:

UK FDI flows, by country, by sector: Office for National Statistics (ONS), London.

FDI-Share Weighted Foreign Within-Sector R&D Stocks

Taking the flows of FDI to and from the UK, by sector, by country, as described above, we can calculate FDI shares in the same way as import shares or input shares are calculated. I favour time-invariant shares (the average share over the sample period) because of the volatility of the data and the incidence of missing values. The FDI-share weighted R&D stocks are then calculated in the same way as the import-weighted R&D stocks described above. Data is unavailable for inter-industry FDI, which limits us to intra-industry FDI variables only. In all, 12 FDI-weighted R&D series for each industry are constructed. One group of four is the sum of inward and outward FDI, with the two alternative depreciation rates for both time-varying and time-invariant FDI shares. The second and third groups are similar, but for inward and outward FDI separately.

Appendix D

Some Specification Tests for Estimating (5.11)

D1: Exogeneity of domestic within-sector R&D stock

In Section 5.6.2 I present *a priori* arguments for treating the explanatory variables in the empirical model as exogenous. These arguments are at their least strong when applied to domestic within-sector R&D stock. I cannot appeal to the small open economy or small industrial sector assumptions in the case of this series, although it can be argued that timing considerations and existing R&D literature suggest exogeneity.

As a further examination of the exogeneity status of this series I carry out a quick Wu-Hausman test[173] on the pooled data. This test consists of estimating a reverse regression (for LRD_j^d on LTFP), saving the residual and then testing the significance of this residual in the LTFP regression. For simplicity, I stick to the linear (Fixed Effects) specification of the LTFP equation. The test and its results are outlined below.

(i) FE regression of LRD_j^d on TREND, $LRD_j^d[-1]$, LTFP, Other Exogenous Variables. ($LRD_j^d[-1]$ acts as the identifying variable).

(ii)Save residual from (i) as HAUS.

(iii) FE regression of LTFP on TREND, LRD_j^d, LRD_i^d, HAUS, Other Exogenous Variables. (LRD_i^d acts as the identifying variable, giving an exactly identified pair of simultaneous equations (i) and (iii)).

(iv) H0: β(HAUS) = 0, ie: LRD_j^d is exogenous.

[173] See Wu (1973).

Table D1: Exogeneity of LRD_j^d

Variable	Coefficient	St. Error	T-Ratio
TREND	-0.052	.015	-3.391
LRD_j^d	-0.036	0.045	-0.785
LRD_j^f	-0.174	0.218	-0.796
$LRD_j^{f,fdi}$	-0.259	0.118	-2.191
LRD_i^d	-0.565	0.946	-0.597
LRD_i^f	2.237	1.045	2.141
HAUS	0.512	0.301	1.701

The results of (iii) are presented above. It is clear from the t-test that β(HAUS) = 0 (at 5% significance level). Therefore, there is no evidence to suggest that the *a priori* reasoning as to the exogeneity of domestic within-sector R&D stocks is false. I therefore proceed with the estimation of the empirical model treating all RHS variables as exogenous.

D2: Starting Values

I use starting values loosely based on the parameter estimates in Keller (1997), as discussed briefly in the text of Section 5.3. A number of alternative sets of starting values are tried to test the robustness of the results (to test the minimum of the RSS reached is global and not local). These alternative sets of starting values are listed below.

Table B2: Starting Values

Parameter	Keller	Same Strength	Same Strength 2	Some 0s	All 0s
β1	.1	.1	.3	.1	0
β2	.12	.12	.2	.12	0
β3	.7	1	1	.7	0
β4	.35	1	1	.35	0
β5	.25	1	1	0	0

The estimates of the standard model with all sets of starting values are identical.

Bibliography

Aghion, P., Harris, C. and Vickers, J. (1996). 'Competition and growth with step-by-step innovation: An example.' *European Economic Review*, **41**, 771-782.

Aghion, P. and Howitt, P. (1992). 'A model of growth through creative destruction.' *Econometrica,* **60**, 2, 323-351.

Aghion, P. and Howitt, P. (1998). Endogenous Growth Theory. MIT Press.

Ahn, S.C. and Schmidt, P. (1995). 'Efficient estimation of models for dynamic panel data.' *Journal of Econometrics*, **68**, 5-27.

Anderson, T.W. and Hsiao, C. (1981). 'Estimation of dynamic models with error components.' *Journal of the American Statistical Association*, 1981, 589-606.

Arestis, P., Demetriades, P. and Kelly, C. (1998). 'Time series evidence on the endogenous growth debate' Mimeo, University of East London.

Arrelano, M. and Bond, S. (1991). 'Some tests of specification for panel data: Monte Carlo evidence and an appliction to employment equations.' *Review of Economic Studies*, **58**, 277-297.

Arrow, K. (1962a). 'The economic implications of learning-by-doing.' *Review of Economic Studies*, **29**, 155-173.

Arrow, K. (1962b). 'Economic welfare and the allocation of resources for inventions', in R. Nelson (ed), *The Rate and Direction of Inventive Activity*, Princetown, NJ, PUP.

Atkinson, A. B. and Stiglitz, J. E. (1969). 'A new view of technological change.' *Economic Journal*, **79**, 573-578.

Backus, D., Kehoe, P. and Kehoe, T. (1992). 'In search of scale effects in trade and growth.' *Journal of Economic Theory*, **58**, 377-409.

Bagwell, K. and Staiger, R. (1994). 'The sensitivity of strategic and corrective R&D policy in oligopolistic industries.' *Journal of International Economics*, **36**, 133-150.

Baldwin, W. and Childs, G. (1969). 'The fast second rivalry in research and development.' *Southern Economic Journal*, **36**, 18-24.

Baltagi, B.H. (1997). Econometrics of Panel Data. John Wiley & Sons, New York.

Barrell, R. and Pain, N. (1997). 'Foreign Direct Investment, Technological Change, and Economic Growth Within Europe.' *Economic Journal*, **107**, No. 445, 1770-86.

Barro, R. (1990). 'Government spending in a simple model of endogenous growth.' *Journal of Political Economy*, **98**, 5, S103-S125.

Barro, R. (1991). 'Economic growth in a cross-section of countries.' *Quarterly Journal of Economics*, **106**, 2, 407-443.

Barro, R. and Sala-I-Martin, X. (1995). Economic Growth. McGraw-Hill.

Baumol, W. J. (1986). 'Productivity growth, convergence and welfare: What the long run data show.' *American Economic Review*, **76**, 5, 1072-1085.

Bayoumi, T., Coe, D. and Helpman, E. (1996). 'R&D Spillovers and Global Growth.' *NBER Working Paper,* No. 5628.

Beath, J., Katsoulacos, Y. and Ulph, D. (1994). 'Strategic R&D and Innovation,' in *Current Issues in Industrial Economics*. MacMillan.

Ben-David, D. (1993). 'Equalising exchange: Trade liberalization and income convergence.' *Quarterly Journal of Economics*, **108**, 3, 653-679.

Berman, E., Bound, J. and Griliches, Z. (1994). 'Chnages in the demand for skilled labour within US Manufacturing: Evidence from the Annual Survey of Manufactures.' *Quarterly Journal of Economics*, **109**, 367-397.

Bernard, A. B. & Jones, C. (1996). 'Productivity across industries and countries: Time series theory and evidence.' *Review of Economics and Statistics*, **78**, 1, 135-146.

Bernstein, J. (1988). 'Costs of Production, Intra and Inter-Industry R&D Spillovers: Canadian Evidence.' *Canadian Journal of Economics*, **21**, No. 2, 324-49.

Bernstein, J. and Nadiri, M. (1988). 'Inter-Industry R&D Spillovers, Rates of Return and Production in High Tech Industries.' *American Economic Review*, **78**, No. 2, 429-34.

Bertschek, I. (1995). 'Product and Process Innovation as a Response to Increasing Imports and Foreign Direct Investment.' *Journal of Industrial Economics*, **43**, No. 4, 341-58.

Bester, H. and Petrakis, E. (1993). 'The incentives for cost reduction in a differentiated industry.' *International Journal of Industrial Organization*, **11**, 519-534.

Blanchard, O. (1985). 'Debt, deficits and finite horizons.' *Journal of Political Economy*, **93**, 223-247.

Blundell, R., Griffith, R. and Van Reenen, J. (1995). 'Dynamic count data models of technological innovation.' *Economic Journal*, **105**, 333-344.

Blundell, R. and Bond, S. (1999). 'GMM estimation with persistent panel data: An application to production functions.' *IFS Working Paper, No. W99/4*, Institute of Fiscal Studies, London.

Bozeman, B. and Link, A. (1983). Investments in Technology: Corporate Strategies and Public Policy Alternatives. NY, Praeger.

Brander, J. and Spencer, B. (1983). 'Strategic commitment with R&D: The symmetric case.' *Bell Journal of Economics*, **14**, 225-235.

Branstetter, L. (1996). 'Are Knowledge Spillovers Intra-National or International in Scope: Microeconometric Evidence from the US and Japan.' *NBER Working Paper*, No. 5800.

Brown, W. B. and Hagendorn, J. S. (1992). International Economics. Addison-Wesley, Reading MA.

Bureau of Economic Analysis (1997): Survey of Current Business, May, 1997.

Bureau of Economic Analysis (1989): Fixed Reproducible Tangible Wealth in the United States, 1925-1989.

Cameron, G. (1995). 'Innovation, Spillovers, and Growth: Evidence from a Panel of UK Manufacturing Industries.' Paper presented to the Royal Economic Society Conference, University of Kent, 1995.

Cameron, G. (1996). 'Innovation and economic growth.' *Centre for Economic Performance, Discussion Paper No. 277*, London School of Economics.

Cameron, G., Proudman, J. and Redding, S. (1998). 'Productivity growth in an open economy: The experience of the UK.' Bank of England, *Working Paper No. 77.*

Campbell, J.Y. and Mankiw, G. (1989). 'Consumption, income and interest rates: Reinterpreting the time series evidence.' *NBER Macroeconomics Annual*, Cambridge, MA, MIT Press.

Campbell, J.Y. and Perron, P. (1991). 'Pitfalls and opportunities: What macroeconomists should know about unit roots.' *NBER Macroeconomics Annual*, 1991, 141-201. MIT Press.

Cass, D. (1965). 'Optimum growth in an aggregative model of capital accumulation.' *Review of Economic Studies*, **32**, 233-240.

Cecchini (1988). The European Challenge. Gower.

Choi, J. (1993). 'Co-operative R&D with product market competition.' *International Journal of Industrial Organization,* **11**, 553-571.

Coe, D. and Helpman, E. (1995). 'International R&D spillovers.' *European Economic Review*, **39**, 859-887.

Coe, D., Helpman, E. and Hoffmaister, A.W. (1997). 'North-South R&D Spillovers.' *NBER Working Paper*, No. 5048.

Cohen, W. and Levin, R. (1989). 'Empirical studies of innovation and market Structure,' in Schmalensee, R. and Willig, R. (eds) *Handbook of Industrial Organization, Volume II.*

Danthine, J. and Hunt, J. (1994). 'Wage bargaining structure, employment and economic integration.' *Economic Journal*, **104**, 528-541.

Dasgupta, P. (1986). 'The theory of technological competition,' in Stiglitz, J. and Mathewson, F. (eds), *New Developments in the Analysis of Market Structure.* Cambridge, MIT Press.

Dasgupta, P. and Stiglitz, J. (1980). 'Industrial structure and the nature of innovative activity.' *Economic Journal*, **90**, 266-293.

D'Aspremont, C. and Jacquemin, A. (1988). 'Co-operative and non co-operative R&D in duopoly with spillovers.' *American Economic Review*, **78**, 1133-1137.

De Long, J. (1988). 'productivity growth, convergence and welfare: Comment.' *American Economic Review*, **78**, 5, 1138-1154.

Denison, E. (1962). 'Sources of growth in the United States and the alternative before us.' Supplement Paper 13, Committee for Economic Development, NY.

Denison, E. (1985). Trends in American Economic Growth. Washington, Brookings Institute.

Devereux, M.B. and Lapham, B.J. (1994). 'The stability of economic integration and endogenous growth.' *Quarterly Journal of Economics*, **109**, 1, 299-308.

Dixit, A. and Stiglitz, J. (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**, 137-147.

Dougherty, C. (1991). *A Comparison of Productivity and Economic Growth in the G7 Countries*, PhD Dissertation, Harvard University.

Driffill, E.J. and Van Der Ploeg, F. (1993). 'Monopoly unions and the liberalization of international trade.' *Economic Journal*, **103**, 379-385.

Driffill, E.J., Meschi, M. and Ulph, A. (1997). 'Product market integration and wages: Evidence from a cross-section of manufacturing establishments in the United Kingdom.' *CEPR Discussion Paper, No. 1818.*

Durlauf, S. N. and Johnson,P. (1995). 'Multiple regimes and cross-country growth behaviour.' *Journal of Applied Econometrics*, **10**, 4, 365-384.

Durlauf, S. N. and Quah, D. (1998). 'The new empirics of economic growth,' in Taylor, J. and Woodforde, M. (eds), *Handbook of Macroeconomics*, forthcoming.

Edwards, S. (1992). 'Trade orientation, distortions and growth in developing countries.' *Journal of Development Economics*, **39**, 31-57.

Eicher, T. S. and Turnovsky, S. J. (1999). 'Non-scale models of economic growth.' *Economic Journal*, **109**, 394-415.

Elias, V. J. (1990). Sources of Growth: A Study of Seven Latin American Economies. San Francisco, ICS Press.

Engelbrecht, H.J. (1997). 'International R&D spillovers amongst OECD economies.' *Applied Economics Letters*, **4**, 315-319.

Englander, A., Evenson, R. and Hanazaki, M. (1988). 'R&D, Innovation and the Total Factor Productivity Slowdown.' *OECD Economic Studies*, **No. 11**, Autumn 1988.

Engle, R.F., Hendry, D.F. and Richard, J.F. (1983). 'Exogeneity.' *Econometrica*, **51**, 277-304.

Evenson, R.E., Kortum, S. and Putnam, J. (1991). 'The Yale-Canada Concordance.' Unpublished manuscript, Economic Growth Centre, Yale University, 1991.

Fagerberg, J. (1994). 'Technology and international differences in growth rates.' *Journal of Economic Literature*, **32**, 1147-1175.

Feenstra, R.C. (1996). 'Trade and Uneven Growth.' *Journal of Development Economics,* **49**, 229-256.

Frankel, J.A. and Romer, D. (1999). 'Does trade cause growth,' *American Economic Review*, **89**, 3, 379-399.

Geroski, P.A., Van Reenen, J. and Walters, C.F. (1997). 'How persistently do firms innovate?' *Research Policy*, **26**, 33-48.

Gilbert, R. J. and Newberry, D. M. G. (1982). 'Pre-emptive patenting and the persistence of monopoly.' *American Economic Review*, **72**, 3, 514-526.

Greene, W.H. (1991). LIMDEP Version 6.0: User Manual and Reference Guide. Econometrics Software Inc. NY.

Greene, W.H. (1993). *Econometric Analysis*, 2nd Edition. MacMillan.

Greene, W.H. (1997). *Econometric Analysis*, 3rd Edition. Prentice-Hall International.

Griliches, Z. (1979). 'Issues in assessing the contribution of R&D to productivity growth.' *Bell Journal of Economics*, **10**, 1, 92-116.

Griliches, Z. (1980). 'R&D and the Productivity Slowdown.' *American Economic Review*, **70**, 343-48.

Griliches, Z. (1988). 'Productivity Puzzles and R&D: Another Non-Explanation.' *Journal of Economic Perspectives*, **2**, No. 4, 9-21.

Griliches, Z. (1992). 'The search for R&D spillovers.' *Scandinavian Journal of Economics*, **94**, 29-47.

Grossman, G. and Helpman, E. (1990). 'Trade, innovation and growth.' *American Economic Association, Papers and Proceedings*, **80**, 2, 86-91.

Grossman, G. and Helpman, E. (1991a). Innovation and Growth in the Global Economy. MIT Press, Cambridge, MA.

Grossman, G. and Helpman, E. (1991b). 'Trade, Knowledge Spillovers and Growth.' *European Economic Review*, **35**, 517-26.

Grossman, G. and Helpman, E. (1994). 'Endogenous innovation in the theory of growth.' *Journal of Economic Perspectives*, **8**, 1, 23-44.

Grossman, G. and Shapiro, C. (1986). 'Research joint ventures: An antitrust analysis.' *Journal of Law, Economics and Organization*, **1**, 315-337.

Grossman, G. and Shapiro, C. (1987). 'Dynamic R&D competition.' *Economic Journal*, **97**, 372-387.

Grout, P.A. (1984). 'Investment and wages in the absence of binding contracts: A Nash bargaining approach.' *Econometrica*, **52**, 449-460.

Hagendoorn, J. and Schakenraad, J. (1993). 'The effect of strategic technology alliances on company performance.' *Strategic Management Journal*, **15**, 4, 291-309.

Hall, B. (1992). 'Investment and R\&D at the firm level: Does the source of financing matter?' *NBER Working Paper, No. 4096.*

Hall, B. and Mairesse, J. (1995). 'Exploring the Relationship Between R&D and Productivity in French Manufacturing Firms.' *Journal of Econometrics*, **65**, 263-93.

Hall, B. and Cummins, C. (1997). *TSP User's Guide Version 4.4.* TSP International, Palo Alto, CA.

Hall, R.E. (1990). 'Invariance Properties of Solow's Productivity Residual', in *Growth/Productivity/Unemployment – Essays to Celebrate Bob Solow's Birthday*, Diamond, P. (ed), MIT Press, Cambridge, MA.

Harris, C. and Vickers, J. (1987). 'Racing with uncertainty.' *Review of Economic Studies*, **54**, 1-21.

Harrison, A. (1996). 'Openness and growth: A time-series, cross-country analysis for developing countries.' *Journal of Development Economics*, **48**, 419-447.

Harrod, R. (1939). 'An essay in dynamic theory.' *Economic Journal*, **49**, 14-33.

Hejazi, W. and Safarian, E. (1998). 'Trade, Investment and R&D Spillovers.' Mimeo, University of Toronto, Canada.

Helpman, E. and Krugman, P. (1985). Market Structure and Foreign Trade. Cambridge, MIT Press.

Henrekson, M., Torstensson, J. and Torstensson, R. (1996). 'Growth effects of European integration.' Forthcoming, *European Economic Review*.

Himmelberg, C.P. and Petersen, B.C. (1994). 'R&D and internal finance: A panel study of small firms in high-tech industries.' *Review of Economics and Statistics*, **76**, 38-51.

Horowitz, I. (1962). 'Firm size and research activity.' *Southern Economic Journal*, **28**, 298-301.

Im, K., Pesaran, M. and Shin, Y. (1996). 'Testing for unit roots in heterogeneous panels.' Mimeo, University of Cambridge.

Jacquemin, A. (1988). 'Co-operative agreements in R&D and European antitrust policy.' *European Economic Review*, **32**, 551-560.

Jones, R. (1995). 'R&D based models of economic growth.' *Journal of Political Economy*, **103**, 4, 759-784.

Jorgenson, D.W. (1990). 'Productivity and Economic Growth.' *Harvard Institute of Economic Research, Discussion Paper, No. 1487.*

Jorgenson, D.W. (1995). Productivity. Cambridge, MIT Press.

Jorgenson, D. W. and Griliches, Z. (1967). 'The explanation of productivity change.' *Review of Economic Studies*, **34**, 249-284.

Jorgenson, D.W. and Landau, R. (1993). 'Tax Reform and the Cost of Capital,' (eds), The Brookings Institution, Washington, D.C.

Judge, G., Griffiths, W.E., Hill, R. and Lee, T.C. (1985). The Theory and Practice of Econometrics, Chapter 13. New York: John Wiley & Sons.

Judson, R.A. and Owen, A.L. (1997). 'Estimating dynamic panel data models: A practical guide for macroeconomists.' *Finance and Economics Discussion Series, 1997-3*, Federal Reserve Board, Washington D.C.

Kaldor, N. (1963). 'Capital Accumulation and economic growth,' in F. Lutz and D. Hague (eds), *Proceedings of a Conference Held by The International Economic Association*, London, MacMillan.

Kamien, M. and Schwartz, N. (1970). 'Market structure, elasticity of demand, and incentive to invent.' *Journal of Law and Economics*, **13**, 241-252.

Kamien, M. and Schwartz, N. (1982). Market Structure and Innovation. Cambridge MA, CUP.

Kamien, M., Muller, E. and Zang, I. (1992). 'Research joint ventures and R&D cartels.' *American Economic Review*, **82**, 1293-1306.

Kao, C. (1999). 'Spurious regression and residual based tests for cointegration in panel data.' *Journal of Econometrics*, **90**, 1-44.

Katz, M. (1985). 'An analysis of co-operative research and development.' *Rand Journal of Economics*, Winter 1985, 527-543.

Keller, W. (1997). 'Trade and the Transmission of Technology.' *NBER Working Paper*, No. 6113.

Kiviet, J. (1997). 'Expectations of expansions for estimators in a dynamic panel data model: Some results for weakly exogenous regressors.' Mimeo, University of Amsterdam.

Koopmans, T.(1965). 'On the concept of optimal economic growth,' in *The Econometric Approach to Development Planning*. Amsterdam.

Krugman, P. (1995). 'Growing world trade: Causes and consequences.' *Brookings Papers on Economic Activity,* **1**, 327-362. Cambridge MA.

Krugman, P. and Venables, A. (1995). 'Globalization and the Inequality of Nations.' *NBER Working Paper*, No. 5098.

Lach, S. and Schankerman, M. (1989). 'Dynamics of R\&D and investment in the scientific sector.' *Journal of Political Economy*, **97**, 4, 880-904.

Leahy, D. and Neary, J. P. (1999). 'R&D spillovers and the case for industrial policy in an open economy.' *Oxford Economic Papers*, **51**, 40-59.

Leamer, E. (1988). 'Measures of openness,' in R. Baldwin (ed): *Trade Policy and Empirical Analysis.* University of Chicago Press, Chicago, IL.

Lee, J-W. (1993). 'International trade, distortions and long-run economic growth.' *IMF Staff Papers*, **40**, 2, 299-328.

Leibenstein, H. (1966). 'Allocative efficiency vs "x-efficiency".' *American Economic Review*, **56**, 392-415.

Levine, R. and Renelt, D. (1992). 'A sensitivity analysis of cross-country growth regressions.' *American Economic Review*, **82**, 4, 942-963.

Lichtenberg, F. (1992). 'R&D Investment and International Productivity Differences.' *NBER Working Paper*, No. 4161.

Lichtenberg, F. and Pottelsberghe de la Potterie, B. (1996). 'International R&D Spillovers: A Re-examination.' *NBER Working Paper*, No. 5668.

Lucas, R. (1988). 'On the mechanics of economic development.' *Journal of Monetary Economics*, **22**, 3-42.

Machin, S. and Van Reenen, J. (1998). 'Technology and changes in skill structure: Evidence from seven OECD countries.' *Quarterly Journal of Economics*, **113**, 1215-1244.

Maddison, A. (1987). 'Growth and slowdown in advanced capitalist economies.' *Journal of Economic Literature*, **25**, 649-698.

Mankiw, N., Romer, D. and Weil, D. (1992). 'A contribution to the empirics of economic growth.' *Quarterly Journal of Economics*, **107**, 407-437.

Mansfield, E. (1968). Industrial Research and Technological Innovation: An Econometric Analysis. NY, Norton.

Mansfield, E., Schwartz, M. and Wagner, S. (1981). 'Imitation costs and patents: An empirical study.' *Economic Journal*, **91**, 907-918.

Mansfield, E. (1985). 'How Rapidly Does New Industrial Technology Leak Out?' *Journal of Industrial Economics*, **34**, 2, 217-23.

Matyas, Laszlo (1996). 'Testing unit roots in heterogeneous panels: Provoking comments.' Presented to 5th Biennial Panel Data Conference, Amsterdam, 1996.

McVicar, D. (1998). 'Strategic R&D and Economic Integration.' Mimeo, University of Southampton, May 1998.

Meliciani, V. and Piermartini, R. (1998). 'Integration and international competitiveness: The changing role of price and non-price competitiveness.' Mimeo, University of Southampton.

Miyagiwa, K. and Ohno, Y. (1997). 'Foreign technology, spillovers and R&D policy.' *International Economic Review*, **38**, 2, 405-430.

Mountford, A. (1996). 'Convergence, overtaking and the welfare implications of international trade in a convex economy.' *Discussion Papers in Economics and Econometrics, No. 9624*, University of Southampton.

Muniagurna, M. and Singh, N. (1997). 'Foreign technology, spillovers and R&D policy.' *International Economic Review*, **38**, 2, 405-430.

Nadiri, M. and Kim, S. (1996). 'International R&D Spillovers, Trade and Productivity in Major OECD Countries.' *NBER Working Paper*, No. 5801.

Neven, D.J. and Roller, L.H. (1991). 'European integration and trade flows.' *European Economic Review*, **35**, 1295-1309.

Nickell, S. (1981). 'Biases in dynamic models with fixed effects.' *Econometrica*, **49**, 1417-1426.

Nickell, S. (1996). 'Competition and corporate performance.' *Journal of Political Economy*, 104, 4, 724-746.

Nickell, S. and Nicolitsas, D. (1997). 'Human capital, investment and innovation: What are the connections?' *CEP Discussion Paper, No. 370.*

OECD (1987). Recent Trends in International Direct Investment, Paris, OECD.

OECD (1996) Technology and Industrial Performance. Paris, OECD.

OECD, DSTI (ANBERD Database), 1997. Paris, OECD.

OECD, DSTI (STAN Industrial Database), 1996. Paris, OECD.

OECD National Accounts, various years. Paris, OECD.

Ordover, J. and Willig, R. (1985). 'Antitrust for high technology industries: Assessing research joint ventures and mergers.' *Journal of Law and Economics*, **28**, 311-333.

Pack, H. (1994). 'Endogenous growth theory: Intellectula appeal and empirical shortcomings.' *Journal of Economic Perspectives*, **8**, 1, 55-72.

Pesaran, M. and Smith, R. (1995). 'Estimating long run relationships from dynamic heterogeneous panels.' *Journal of Econometrics*, **68**, 79-113.

Phillips, A. (1966). 'Patents, potential competition and technical progress.' *American Economic Review*, **56**, 301-310.

Piermartini, R. and Ulph, A. (1996). 'Trade liberalization and innovative success with unionised labour markets.' *Discussion Papers in Economics and Econometrics, No. 9637,* University of Southampton.

Pritchett, L. (1996). 'Measuring outward orientation in LDCs: Can it be done?' *Journal of Development Economics*, **49**, 307-339.

Quah, D. (1994). 'Exploiting cross-section variation for unit root inference in dynamic data.' *Economics Letters*, **44**, 9-19.

Quah, D. (1996). 'Growth and convergence in models of distribution dynamics.' *Economic Journal*, **106**, 1045-1055.

Quah, D. (1997). 'Empirics for growth and distribution: Stratification, polarization and convergence clubs.' *Journal of Economic Growth*, **2**, 1, 27-59.

Qui, L. (1994). 'R&D investment and product differentiation: Cournot vs Bertrand competition.' Mimeo, Honk Kong University of Science and Technology.

Ramsey, F. (1928). 'A mathematical theory of saving.' *Economic Journal*, **38**, 543-559.

Rebelo, S. (1991). 'Long run policy analysis and long run growth.' *Journal of Political Economy*, **99**, 3, 500-521.

Rivera-Batiz, L. and Romer, P. (1991a). 'Economic Integration and Endogenous Growth.' *Quarterly Journal of Economics,* **106**, 531-555.

Rivera-Batiz, L. and Romer, P. (1991b). 'International trade with endogenous technological change.' *European Economic Review*, **35**, 971-1004.

Romer, P. (1986). 'Increasing returns and long-run growth.' *Journal of Political Economy*, **94**, 5, 1002-1037.

Romer, P. (1987). 'Growth based on increasing returns due to specialization.' *American Economic Review Papers and Proceedings*, **77**, 2, 56-72.

Romer, P. (1989). 'Human capital and growth: Theory and evidence.' *NBER Working paper No. 3173,* Cambridge MA.

Romer, P. (1990). 'Endogenous Technological Change.' *Journal of Political Economy*, **98**, No. 5, S71-S102.

Romer, P. (1993). 'Idea gaps and object gaps in economic development.' *Journal of Monetary Economics*, **32**, 543-573.

Romer, P. (1994). 'The origins of endogenous growth.' *Journal of Economic Perspectives*, **8**, 1, 3-22.

Rousseeuw, P. J. and Leroy, A. M. (1997). Robust Regression and Outlier Detection. New York, Wiley.

Scherer, F. (1967). 'Market structure and the employment of scientists and engineers.' *American Economic Review*, **57**, 524-531.

Scherer, F.M. (1984). 'Using Linked Patent and R&D Data to Measure Inter-industry Technology Flows', in Griliches, Z. (ed), *R&D, Patents and Productivity*, Chicago: The University of Chicago Press, 1984, pp 417-61.

Schmookler, J. (1966). Invention and Economic Growth. Cambridge MA, HUP.

Schumpeter, J. (1934). The Theory of Economic Development. Cambridge MA. HUP.

Sheshinski, E. (1967). 'Optimal accumulation with learning-by-doing,' in K. Shell (ed), *Essays on the Theory of Economic Growth.* Cambridge, MA, MIT Press.

Smith, A. (1776). An Inquiry into the Nature and Causes of the Wealth of Nations. Everyman's Library, London, 1981.

Solow, R. (1956). 'A contribution to the theory of economic growth.' *Quarterly Journal of Economics*, **70**, 65-94.

Solow, R. (1957). 'Technical Change and the Aggregate Production Function.' *Review of Economics and Statistics*, **39**, 312-20.

Solow, R. (1994). 'Perspectives on growth theory.' *Journal of Economic Perspectives*, **8**, 1, 45-54.

Spencer, B. and Brander, J. (1983). 'International R&D rivalry and industrial strategy.' *Review of Economic Studies*, **50**, 707-722.

Stoneman, P. and Kwon, M-J. (1997). 'Technological opportunity, technological diffusion and gross investment: An inter-industry approach.' Mimeo, University of Warwick.

Sutton, J. (1996). 'Tehnology and market structure.' *European Economic Review*, **40**, 511-530.

Takalo, T. and Kanniainen, V. (1997). 'Do patents slow down technological progress?' Mimeo, University of Warwick.

Temple, J. (1998). 'Appropriate technology: Five growth puzzles solved?' Mimeo, Institute of Economics and Statistics, Oxford.

Temple, J. (1999). 'The new growth evidence.' *Journal of Economic Literature*, **37**, 112-156.

Tirole, J. (1988). The Theory of Industrial Organization. Cambridge MA, MIT Press.

Uzawa, H. (1965). 'Optimal technical change in an aggregative model of economic growth.' *International Economic Review*, **6**, 18-31.

Uzawa, H. (1968). 'Time preference, the consumption function and optimum asset holdings,' in J.N. Wolfe (ed), *Value, Capital and Growth*. Chicago, Aldine.

Van Reenen, J. (1996). 'The creation and capture of rents: Wages and innovation in a panel of UK companies.' *Quarterly Journal of Economics*, **111**, 195-226.

Venables, A. and Smith, A. (1986). 'Trade and industry.' *Economic Policy*, 621-671.

Ventura, J. (1997). 'Growth and interdependence.' *Quarterly Journal of Economics*, **112**, 1, 57-84.

Williamson, O. (1965). 'Innovation and market structure.' *Journal of Political Economy*, **73**, 67-73.

Wood, A. (1995). 'How trade hurt unskilled workers.' *Journal of Economic Perspectives*, **9**, 3, 57-80.

World Bank, (1996). Global Economic Perspectives and the Developing Countries. Washington, D.C., The World Bank.

Wu, D. (1973). 'Alternative tests of independence between stochastic regressors and disturbances.' *Econometrica*, **41**, 733-750.

Young, A. (1992). 'A tale of two cities: Factor accumulation and technical change in Hong Kong and Singapore,' in O.J. Blanchard & S. Fischer (eds), *NBER Macroeconomics Annual, 1992*. Cambridge MA, MIT Press.

Young, A. (1995a). 'The tyranny of numbers: Confronting the statistical realities of the East Asian growth experience.' *Quarterly Journal of Economics*, **110**, 641-680.

Young, A. (1995b). 'Growth without scale effects.' *NBER Working Paper, No. 5211*.