ECONOMIC DEVELOPMENT AND STRUCTURAL CHANGE

Leif van Neuss

Ph.D. Thesis in Economics and Management

Jury
Lionel Artige (ULg, Thesis Director)
Lilas Demmou (OECD)
Axel Gautier (ULg)
Jean Imbs (Paris School of Economics)
Bernard Lejeune (ULg)

HEC Liège
Academic Year 2016-2017
# Contents

Acknowledgements

General Introduction

1 Why Did the Industrial Revolution Start in Britain
   1 Introduction ........................................................................ 16
   2 The Industrial Revolution: A Technological Revolution .... 18
      2.1 The Steam Power ................................................................ 19
      2.2 The Textile Industry ....................................................... 21
      2.3 The Metalworking Industry ............................................ 24
      2.4 The Transport Industry .................................................. 27
         2.4.1 Canals and Roads .................................................. 27
         2.4.2 Railway ................................................................. 30

2 Before the Industrial Revolution ........................................... 32
   3.1 Evolution of English GDP per Capita in the Early Modern Period ........................................................................ 32
   3.2 A Model of Transition towards Modern Industry: Proto-industrialization ................................................................. 37

4 The Causes of the Industrial Revolution ............................... 39
   4.1 Geography and Natural Resources ................................. 41
   4.2 Demography ................................................................. 44
   4.3 Agricultural Progress ................................................... 48
   4.4 Consumer Revolution and Urbanization ..................... 53
   4.5 Trade and Empire ......................................................... 57
   4.6 Institutions and Policy .................................................. 60
   4.7 Modern Science ........................................................... 65
   4.8 Human Capital ........................................................... 68

5 Conclusion ............................................................................ 71
## 2 The Drivers of Structural Change

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>76</td>
</tr>
<tr>
<td>2</td>
<td>The Empirical Facts of Structural Change</td>
<td>78</td>
</tr>
<tr>
<td>2.1</td>
<td>How to Decompose the Aggregate Economy into Sectors?</td>
<td>78</td>
</tr>
<tr>
<td>2.2</td>
<td>How to Measure Structural Change?</td>
<td>79</td>
</tr>
<tr>
<td>2.3</td>
<td>The Empirical Facts of Structural Change for Different Regions of the World</td>
<td>80</td>
</tr>
<tr>
<td>3</td>
<td>The Drivers of Structural Change</td>
<td>90</td>
</tr>
<tr>
<td>3.1</td>
<td>Income Effects, Relative Price Effects, and Structural Change</td>
<td>90</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Income effects and Structural Change</td>
<td>91</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Relative Price Effects and Structural Change</td>
<td>93</td>
</tr>
<tr>
<td>3.1.2.1</td>
<td>Sectoral Differences in Total Factor Productivity Growth</td>
<td>94</td>
</tr>
<tr>
<td>3.1.2.2</td>
<td>Sectoral Differences in Factor Intensity</td>
<td>97</td>
</tr>
<tr>
<td>3.1.2.3</td>
<td>Sectoral Differences in the Elasticity of Substitution between Factors</td>
<td>98</td>
</tr>
<tr>
<td>3.1.3</td>
<td>The Relative Importance of Income Effects (Preferences) and Relative Price Effects (Technology)</td>
<td>100</td>
</tr>
<tr>
<td>3.2</td>
<td>Input-Output linkages and Structural Change</td>
<td>105</td>
</tr>
<tr>
<td>3.3</td>
<td>International Trade and Structural Change</td>
<td>112</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Trade in Goods and Structural Change in Developing Countries</td>
<td>114</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Trade in Goods and Structural Change in Developed Countries</td>
<td>117</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Trade in Tasks and Structural Change</td>
<td>119</td>
</tr>
<tr>
<td>3.3.4</td>
<td>The Relative Importance of Globalization and Trade</td>
<td>121</td>
</tr>
<tr>
<td>4</td>
<td>Conclusion</td>
<td>122</td>
</tr>
<tr>
<td>5</td>
<td>Appendix A: Country and Variable Coverage in the GGDC 10-Sector Database</td>
<td>125</td>
</tr>
</tbody>
</table>
3 Globalization and Deindustrialization in Advanced Countries

1 Introduction ........................................ 128

2 Deindustrialization: Definition and Empirical Evidence .... 130
   2.1 What Is Deindustrialization? ...................... 130
   2.2 Evidence for Currently Rich Countries ........ 130

3 Existing Explanations for Deindustrialization ........... 133

4 Review of the Empirical Literature ........................ 137

5 Empirical Methodology .................................. 141
   5.1 Estimation strategy ................................. 141
   5.2 Measuring the Contribution of Globalization .... 143

6 Data .................................................. 146

7 Empirical Results ...................................... 149

8 Endogeneity of Trade Flows ............................ 156

9 Conclusion ............................................ 161

10 Appendix ............................................. 163
   10.1 Appendix A: Manufacturing Share of Nominal Value
       Added and Employment in Advanced Countries .... 163
   10.2 Appendix B: Growth of Value Added and Employment
       in Advanced Countries ............................. 165
   10.3 Appendix C: Direct Contribution of Manufacturing Trade
       to Deindustrialization in Advanced Countries .... 166

4 A New Shift-Share Method ................................ 169

1 Introduction ........................................ 170

2 On the Merits of Shift-Share Analysis ................... 172

3 Dunn's (1960) Shift-Share Method ........................ 173
   3.1 The Decomposition Method ......................... 173
   3.2 A Critical Assessment .............................. 175

4 Esteban-Marquillas's (1972) Shift-Share Method .............. 176
   4.1 The Decomposition Method ......................... 176
   4.2 A Critical Assessment .............................. 178

5 A Shift-Share Test .................................... 180
   5.1 Two Difficulties .................................... 180
Avenues for further research 191

Bibliography 193
Acknowledgements

I would like to express my special appreciation and thanks to my PhD supervisor, Professor Lionel Artige, for supporting me during these past five years and for allowing me to grow as a research scientist. Lionel has been a tremendous mentor for me. Extremely open-minded and exceptionally good at understanding the difficulties experienced by a PhD candidate throughout the doctoral thesis, he has always been highly supportive and encouraging, and has given me the freedom to pursue various projects without objection. He has also provided insightful discussions and suggestions about the research. I hope that I could be as lively, enthusiastic, and energetic as Lionel.

Besides my supervisor, I am also very grateful to Professor Bernard Lejeune for spending time reading and commenting drafts of several chapters of my PhD dissertation. His enlightened comments have always been challenging.

I would also like to thank Professor Lilas Demmou, Professor Axel Gautier, and Professor Jean Imbs for serving as my committee members and for providing very helpful comments and suggestions about the research.

Finally and most importantly, I would also like to thank my friends and the members of my close family, who encouraged and helped me at every stage of my personal and academic life.
General Introduction

Structural change or structural transformation, defined as the process of reallocation of activity across the three broad sectors agriculture, manufacturing and services (Herrendorf et al., 2014), has always been object of interest in the economic literature. While structural change is arguably a common feature of society development throughout the whole history of mankind, the historical analysis of structural change usually begins with the industrial revolution, sometimes viewed as the deepest mutation ever known to have affected men since Neolithic times (Cippola, 1976). Even though the “early modernists” have recently detected economic progress in Europe and, more particularly, in the North Sea Area - the UK and the Low Countries - as part of progressive trends going back to the late Middle Ages\(^1\), an evolution sometimes labelled as the “Rise of the West”, the industrial revolution, which first broke out in Britain before sweeping through Continental Europe and then the USA, is indeed traditionally considered as the gradual transition period towards modern industry and sustained economic growth, marking the onset of what Huntington (1996) has called the “Great Divergence” between the Western World and East Asia or, more generally, the rest of the world (e.g., Pomeranz, 2000; Goldstone, 2015; Clark, 2014, 2016). Thus, according to Clark’s (2014) estimates, the average rate of efficiency improvement through technological change in the world economy was close to 0 before 1760. Between 1760 and 1860, the average rate of efficiency advance in England increased to 0.5% per year, a rate still modest by later standards, i.e. in comparison to the rates of Total Factor Productivity (TFP) growth achieved by modern successful economies located at - or close to - the technological frontier in the 20\(^{th}\) century, but quite remarkable given its sustained character. As Clark (2014, p. 219) notes: “Before the British Industrial Revolution we find no sign of any equivalent efficiency advances. This is true globally all the way

from 10,000 BC to 1800”. For sure, the industrial revolution did not create structural change but eventually led to an acceleration of the process of reallocation of economic activity across the three broad sectors (agriculture, manufacturing, and services) in a movement narrowly linked to economic development. For this reason, Kuznets (1973) in his Nobel Prize lecture refers to structural change, a process that is characterized, among others, by the massive reallocation of labor out of agriculture, as one of the six main features of modern economic growth, described as the sustained rise in productivity and living standards. By way of illustration, the share of agriculture in total employment declined from 75% in 1800 to less than 3% in 2000 in the United States, while the share of agriculture in total output dropped from 40% in 1840 to around 1% in 2000 (Dennis and Iscan, 2009).

While structural change has always received attention from economists, its study has known a great revival over the last few decades. This surge of interest has been largely driven by the numerous issues and concerns raised by the deindustrialization process that has particularly affected the most advanced countries since the last third of the 20th century. While the reallocation of activity and resources away from agriculture, i.e. the industrialization, has long been viewed as growth-enhancing, characterizing the passage from predominantly agrarian societies to modern ones, the reallocation of economic activity and resources out of manufacturing, i.e. the deindustrialization, has commonly been associated with social troubles and relative economic decline. The literature is full of examples of regions or cities where deindustrialization, which materialized into massive shutdown and many job losses in the manufacturing sector, coincided with a period of socioeconomic dislocation. In line with these deindustrialization stories, a large number of scholars have emphasized the “socioeconomic costs” implied by economic restructuring. These costs can take a myriad of forms. Thus, deindustrialization has notably been blamed for rising poverty, which in turn can translate into higher inequalities (e.g., Doussard et al., 2009), lower economic aspirations, motivations and expectations (e.g., Beck, 2000; Meyer, 2009; Sissons, 2009), lower ability to finance public education and urban schools (e.g., Spayd and Dye, 1991; Bettis, 1994), poorer health conditions (e.g., Renner and Navarro, 1989; Wallace
and Wallace, 1999; Ostry et al., 2002; EIR, 2006; Holland et al., 2011), higher violence (e.g., Ousey, 2000; Matthews et al., 2001), as well as higher rates of suicide (e.g., Kubrin et al., 2006) and mortality (e.g., EIR, 2006; Hanlon et al., 2010).

While deindustrialization can truly have devastating socioeconomic effects on local communities in regions which are especially affected by the erosion of the manufacturing base, it is more difficult to assess the aggregate impact of deindustrialization at national level, as structural change involves a number of redistributive effects. Thus, despite a severe economic downturn in many industrial cities of the New England, Mid-Atlantic, and Midwest regions of the USA, Reardon (2005) notices that the period between 1970 and 1990 was, with the exception of a few years, a time of significant and sustained growth in the US economy taken as a whole. Adopting a more pessimistic view, Kollmeyer (2013) links the deindustrialization process to the persistently high unemployment rates experienced by most affluent countries since the mid-1970s, a relation influenced by the existence of mobility costs associated with the sectoral reallocation of labor. Focusing on the British economy, Kitson and Michie (2014) also blame deindustrialization for generating inferior growth and efficiency advance at the aggregate level during the postwar period. In line with this view, the alarming hypothesis of deindustrialization as the main factor responsible for current slower economic growth in the Western countries, which some economists believe to be the onset of a “secular stagnation”, has also gained in popularity in recent years as the manufacturing share of total workforce is reaching lower and lower levels.

In reality, the study of deindustrialization really took off during the 1980s as the scale and regional effects of economic restructuring and manufacturing job losses became more apparent and tangible, especially in the United Kingdom and the United States, thus contradicting Lawrence’s (1983) early claim that deindustrialization was a “myth”. As part of the research agenda, scholars have devoted rising efforts in understanding both the causes and consequences of deindustrialization and, more generally, of structural change, as well as the role of policy instruments in driving the allocation of activity across sectors of activity.
What are the key economic forces behind the process of structural change? What are the economic consequences of structural change, for instance in terms of employment performance and/or growth potential? Does structural change arise as a sort of efficient equilibrium outcome? Or does structural change justify any government intervention? As noted by Aghion and Howitt (1998) and Herrendorf et al. (2014), these kinds of questions require to extend the traditional one-sector growth model, which has long been used to capture and analyze the nature and dynamics of modern economic growth, and jump to multi-sector growth models. Recent literature has thus extensively used the neoclassical multi-sector growth framework to investigate structural change along with the behavior of macroeconomic aggregates. As part of this research, scholars have formalized a number of channels through which the reallocation of activity across sectors can take place. Accordingly, the factors potentially responsible for structural change range from preferences (e.g., Kongsamut et al., 2001; Foellmi and Zweimller, 2008) to technology (e.g., Ngai and Pissarides, 2007; Acemoglu and Guerrieri, 2008; Alvarez-Cuadrado et al., 2016), and also include input-output (sectoral) linkages (e.g., Berlingieri, 2014) and international trade (e.g., Matsuyama, 2009; Autor et al., 2013; Uy et al., 2013). Regarding the macroeconomic consequences of structural change, especially the growth effects, a major challenge of the recent multi-sector growth literature has been to solve the so-called “Kuznets-Kaldor-Puzzle”, that is to reconcile the Kuznets stylized facts on structural change with the Kaldor facts on aggregate dynamics, or less restrictively some aspects of the Kaldor facts which are well embodied in the one-sector balanced growth models (e.g., Kongsamut et al., 2001; Ngai and Pissarides, 2007; Stijepic and Wagner, 2011; Boppart, 2014).

From a theoretical point of view, the impact of structural change on economic growth depends on whether growth is “sector-indifferent”. Distinguishing between the concepts of “activity” and “sector”, Palma (2014) has recently proposed to classify the growth theories into three camps. The first camp includes the theories viewing growth as both sector-indifferent and activity-indifferent”. Examples are Solow-type models and the branch of endogenous theories that associate growth with increasing returns that are
“activity-indifferent”, that is not based on the use of R&D or the accumulation of human capital. The second camp includes the theories considering growth as “sector-indifferent” but “activity-specific”. In these models of growth, like the neo-Schumpeterian models, increasing returns, though generated by e.g. R&D activities, are explicitly not associated with one or several specific sectors as such. The third camp gathers the theories considering growth as both “sector-specific” and “activity-specific”. In these growth models, like the Post-Keynesian, Schumpeterian and structuralist models, the patterns and dynamics of growth are crucially dependent on the activities being developed, with the latter being specific to the nature of the sector(s) involved. As manufacturing is traditionally believed to have special properties as an engine of economic growth, Palma (2014, p. 21) notes that “the crucial difference between this camp and the previous two ones is that issues such as technological change, externalities, synergies, balance-of-payments sustainability, and the capacity of developing countries to catch up, are in fact directly linked to the size, strength and depth of the manufacturing sector”.

A number of empirical studies have also been devoted to quantifying the effects of deindustrialization on (productivity) growth (e.g., Pieper, 2000; Dasgupta and Singh, 2006; Szirmai and Verspagen, 2011; Szirmai, 2012; Naudé et al., 2013). Such exercise can actually be performed in a variety of ways, ranging from the use of growth accounting techniques, like the canonical shift-share method, to the use of econometric estimations. As recently pointed out by Tregenna (2015, p. 34), this empirical literature is however inconclusive: “The empirical evidence is mixed. A positive relationship between manufacturing and growth is found in several studies but there is variation in this relationship, including for different periods of time”. While acknowledging the uncleanness of the empirical results, Tregenna (2015) argues that the growth effects of deindustrialization are likely to be dependent inter alia on the level of income per capita and the degree of industrialization at which deindustrialization begins; on the nature of deindustrialization, including both its causes and dynamics; on the nature of the manufacturing activities which are in relative decline as well as of the activities, both the manufacturing and non-manufacturing ones, which are relatively growing.
This doctoral thesis is composed of 4 chapters that have a lot to do with structural change, defined as the reallocation of economic activity across the three broad sectors (agriculture, manufacturing and services) that accompanies the process of economic development. The first chapter of this thesis gets interested in the British industrial revolution, an event traditionally taken as the starting point of the processes that eventually brought about modern economic growth and accelerated structural change. The main goal of this chapter is to provide an integrated survey of the literature devoted to identifying the causes of the British industrial revolution. Why did the industrial revolution, the event that defines our lives, occurred in eighteenth-century Britain? This question has long animated a lot of discussions among scholars and is still nowadays heatedly debated, making the literature on the origins of British industrialization, and more generally on the origins of the modern economy, incredibly huge, fast growing and hard to keep track. This survey, which aims at updating and completing Mokyr’s (1999) previous contribution, also spends time emphasizing the main points of criticism that have been directed to the factors proposed as “causal” for the British industrial revolution. Besides the intrinsic interest of writing a survey on the theories aimed at explaining the origins of the British industrial revolution, this chapter is also beneficial in offering a historical perspective to a number of topical economic issues. By way of illustration, modern society is believed to be affected by the emergence of a new industrial revolution, sometimes labelled as the “Fourth Industrial Revolution” or “Industry 4.0”, which actually promotes the computerization and robotization of manufacturing, and the creation of “smart factories”, thus pushing even further the substitution of labor by capital.

The second chapter is dedicated to understanding the process of structural change and its causes. What is structural change? How does one measure it in practice? What are the economic forces behind the process of structural change? Using the GGDC 10-Sector Database (Timmer et al., 2015), Chapter 2 first presents an original comparative description of the empirical facts associated with the familiar sectoral trichotomy (agriculture, manufacturing, and services) in different regions of the world, namely Sub-Saharan and North Africa, Asia, Latin America, Europe and the USA. Among others,
the descriptive analysis reveals that the expansion of GDP per capita has been accompanied by a decline in the share of agriculture in both employment and nominal value added in all countries, while the share of services has risen. Manufacturing has moved on a quite different trajectory as its share follows a hump shape, that is, it is rising for lower levels of economic development (industrialization) and declining for higher levels of economic development (deindustrialization). The analysis also reveals that a certain number of developing economies in both Africa and Latin America have seemingly gone through a process of “premature deindustrialization” - defined by Tregenna (2015, p. 2) as “deindustrialization that begins at a lower level of GDP per capita and/or at a lower level of manufacturing as a share of total workforce and GDP, than is typically the case internationally” - since the 1980s.

Chapter 2 then makes a review of four theoretical determinants of structural change: (i) changes in income, (ii) changes in relative (sectoral) prices, (iii) changes in input-output (sectoral) linkages and (iv) changes in comparative advantages via international trade.

The third chapter deals more specifically with deindustrialization in advanced countries. As shown by Palma (2014), the advanced OECD countries began deindustrializing in the last third of the 20th century. By way of illustration, the share of manufacturing in total workforce declined from 28.2% to 15.6% in the EU15 between 1970 and 2007, while it decreased from 22.4% to 9.9% in the US over the same period (EU KLEMS Database - O’Mahony and Timmer, 2009). In 2015, the employment share of manufacturing is estimated to amount about 12.5% in the EU15 (Eurostat) and 8.9% in the United States (BLS). While the potential determinants of deindustrialization now seem to be relatively well understood on theoretical grounds, a number of empirical issues still remain largely open. For instance, the relative importance of the economic forces driving deindustrialization is not well established in the literature. Yet this question has important implications, for instance in terms of public policies and growth perspectives. In particular, the role of the “internal” factors and “external” factors, i.e. those linked to globalization, is not clear. The empirical results indeed appear to be highly fragile, arguably because the commonly used indicators of trade in manufactures are
not properly defined in most empirical literature using econometric estimations to capture the contribution of global exchanges to deindustrialization. Though the finding of our study, based on panel data for 15 OECD advanced countries from 1970 to 2007, do not necessarily contradict the widespread belief that internal factors are quantitatively more important in accounting for deindustrialization in the OECD taken as a whole, they however suggest that global exchanges have the potential to affect significantly and substantially a country’s sectoral patterns of employment, and that the contribution of trade, especially of trade with developing countries, to the deindustrialization observed in advanced countries may be revised upwards when using better-defined indicators of trade in manufactures.

The fourth chapter is about the shift-share method, a well-known decomposition technique widely used in various fields of research. As part of the study of structural change in economics, the shift-share analysis is largely performed to quantify an “industry-mix effect” and a “competitive effect” on the growth of any relevant macroeconomic variable (e.g. employment, output, consumption, and productivity) measured at territory level. More specifically, the “industry-mix effect” aims at computing the contribution of the evolution of a territory’s sectoral structure to the aggregate growth of the investigated variable at territory level. Thus, broadly speaking, the “industry-mix effect” seeks to determine to which extent structural change influences the macroeconomic performances of a territory. The shift-share method has always been subject to criticism for its lack of theoretical basis. Chapter 4 presents a critical assessment of the two founding shift-share methods by Dunn (1960) and Esteban-Marquillas (1972), then proposes a new shift-share accounting decomposition which separates out the two effects unambiguously. By way of illustration, it also gives an application to manufacturing employment in Belgian provinces between 1995 and 2007.
Chapter 1

Why Did the Industrial Revolution Start in Britain

Leif van Neuss

Abstract

The main goal of this paper is to provide an integrated overview of the literature devoted to identifying the causes of the British industrial revolution. Why did the industrial revolution, a fascinating and multifaceted event which eventually brought about modern economic growth, occur in eighteenth-century Britain? This research question has animated a lot of discussions among scholars and is still nowadays heatedly debated in the literature. This debate reflects in the large spectrum of theories that aim at explaining the origins of British industrialization. In addition to describing a large but not exhaustive list of these theories, this paper also spends time presenting the main points of criticism that have been directed to the factors proposed as “causal” for British industrialization. These factors are classified into eight broad categories: (I) geography and natural resources, (II) demography, (III) agricultural progress, (IV) consumer revolution and urbanization, (V) trade and empire, (VI) institutions and policy, (VII) modern science and (VIII) human capital.

Keywords: Economic History, Industrial Revolution, Manufacturing, Economic Growth
JEL classification: N, O10, O30
1 Introduction

The industrial revolution is often viewed as the deepest mutation ever known to have affected men since Neolithic times (Cippola, 1976). According to traditional historiography, the industrial revolution first started in Britain during the second half of the 18th century, then spread to Continental Europe and the USA. From an economic point of view, there seems to be a growing consensus on the idea that the industrial revolution corresponds to an increase in the rates of efficiency improvement. According to Clark (2007), the average rate of efficiency growth through technical change in the world economy was close to 0 before 1760. Between 1760 and 1860, it then increased to 0.5% per year in the English economy, a rate still modest by later standards, i.e. compared to the rates of total factor productivity growth achieved by modern successful economies located at or near the technological frontier in the 20th century, but quite remarkable given its sustained character. In Clark's (2014, 2016a) terms, this upsurge in efficiency was a singularity, a unique break in the world economic history, marking the onset of what Huntington (1996) has called the “Great Divergence” between the western countries and East Asia or, more generally, the rest of the world1.

The main goal of this paper is to carry out a survey of the abundant literature devoted to identifying the causes of the British industrial revolution. Which factors have been proposed to explain why Britain was the very first country to undergo a successful industrial revolution? This survey also spends time presenting the main points of criticism which have been directed towards the proposed causes of British industrialization. A major challenge facing the literature on the industrial revolution is the lack of universally accepted data for the period under investigation. Most data, especially economic data, are actually based on estimates and assumptions that are very far from being consensual. Highly conflictual estimates often make the debate hard to follow, at times supporting one specific theory and at times dismissing it. To keep things as clear as possible, we choose, as part of this survey, not to go too deeply into the statistical debate, though we will provide data

1See also Goldstone (2015).
with parsimony to illustrate the different theories. The paper is organized as follows. Section 2 aims to define the industrial revolution and describes some of the most important technological inventions which are commonly associated with that major event. Section 3 sheds light on a new rising debate on the evolution of the British economy in the centuries before the industrial revolution, as well as on the economic mechanisms that drove this evolution. Was the British economy trapped in Malthusian dynamics or precociously launched on a progressive path towards modernity? Section 4 is the core of this paper. It discusses a large but not exhaustive list of the factors proposed as “causal” for British industrialization. These factors are classified into eight broad categories: (I) geography and natural resources (II) demography (III) agricultural progress (IV) consumer revolution and urbanization (V) foreign trade and empire (VI) institutions and policy (VII) modern science and (VIII) human capital. A final section gives concluding thoughts and remarks.
2 The Industrial Revolution: A Technological Revolution

Defining the industrial revolution is not easy as the phenomenon eventually affected every aspect of people’s daily life in the societies that successfully went through economic modernization. In fact, there are different ways of studying the industrial revolution. This refers to what Mokyr (1999) calls the different “schools of thought”. Based on the selected approach, the definition of the industrial revolution can then objectively take various forms. However, it seems impossible to deliver a definition of the industrial revolution without using the concept of technological change, regardless of its exact historical role. As a consequence, the technology-based definitions of the industrial revolution are by far the most widespread in the literature. For instance, Mokyr (1999, p. 12) defines the industrial revolution as “an age of rapidly changing production technology propelled by technological creativity”. Distinguishing between the concepts of “macroinvention” and “microinvention”, Mokyr (1999, p. 16) also proposes to define this event as “a clustering of macroinventions leading to an acceleration in microinventions”. This definition actually looks like the traditional definition offered by the economic literature for the notion of “technological revolution”, viewed as the convergence or the combination, at one point, of macroinventions. In the case of the industrial revolution, most macroinventions were not only radical, with the real potential to trigger a technological rupture and define new ways of thinking about production and consumption, but also economically significant. They unlocked the production capacity of the economy and sowed the seeds of future microinventions that progressively diffused, adapted, extended, improved and made profitable all the techniques developed during the industrial revolution. In this section, we propose to briefly describe some of the main technological inventions that are usually connected to the industrial revolution. In order to limit the size of the section, we focus on the technological inventions that revolutionized (1) the use of energy, (2) the production processes in the textiles and the met-
allurgy\(^2\), and (3) the transport sector. Most of these inventions were British and prepared Britain’s long-lasting technological and economic superiorit y.

2.1 The Steam Power

According to many historians, the development of a large-scale industry capable of exploiting optimally the new technological inventions required the use of another source of energy. In line with this view, the steam engine was not only one of the most radical inventions ever made but also arguably one of the economically most important inventions connected to the industrial revolution. In these terms, it is no accident that the steam engine has become one of the technological symbols of the new industrial age. By breaking the separation between kinetic energy (work or motion) and thermal energy, the steam engine indeed paved the way for a new technical system.

The origins of the steam engine are usually located in the French scientist Papin’s invention (1690). It consisted of a piston moving up and down in a cylinder due to alternative heating and cooling. Papin’s invention, which extended Huygens’ previous work, actually marked the culmination of a long series of reflections on atmosphere pressure held throughout the second half of the 17\(^{th}\) century. It definitely proved the suspected potential of steam as a source of energy. In 1698, Thomas Savery patented the atmospheric engine. But the latter was not performant enough to receive any industrial application. It is the British mechanic Thomas Newcomen who developed the very first successful steam engine. The new machine was revolutionary as it used the power of steam to produce mechanical work. The first operational steam engine was commercially produced and installed in 1712 at the Conygree Coalworks near Dudley in the West Midlands. Initially, Newcomen’s engines were exclusively used to pump water out of mines. Due to a growing demand for coal as from the late 17\(^{th}\) century in Britain, drilling was deeper and deeper in mines, considerably increasing the risks of flooding. In order to fully exploit the huge available mass of underground coal, which provided

\(^2\)Rostow (1960) defines these two industrial sectors as the “leading sectors” of the industrial revolution, i.e. the sectors that give the impetus to the development of the whole economy.
much more energy than wood, it was necessary to find a new way to pump water out of the coal mines. The Newcomen’s steam engine brought the solution and finally unlocked the high potential of coal. Britain had henceforth seemingly inexhaustible quantities of cheap energy. But the Newcomen’s machine was highly inefficient. In particular, it burnt huge amounts of coal and its location was so limited to pithead (entrance to a mine) where coal was virtually free.

In 1769, James Watt came up with a decisive improvement which definitively gave to steam the power to support the upcoming industrialization. Aware of the main source of inefficiency of Newcomen’s machine, namely the heat loss, Watt created a separate condenser that dramatically improved the performances of the steam engine. Watt designed his new invention at the University of Glasgow, an establishment that was particularly known for its teaching of modern sciences and its cutting-edge research. Specialized in both the manufacturing and repairs of scientific tools and instruments, Watt was running a maintenance workshop on the campus. Prevented from going into business because of the very strict corporative rules in Glasgow city, Watt found at the university a favorable place to express his talent. He got the ingenious idea of the separate condenser during some repairs performed on a Newcomen’s steam engine. Associated with John Roebuck then Matthew Boulton, a great industrialist who was early aware of the advantages of not relying anymore on natural forces (water or wind) to power industrial activity, Watt developed the new version of the steam engine in Birmingham city. The first commercially produced Watt’s steam engines were installed in the mines of Bloomfield and in the Wilkinson’s blast furnaces in 1776. Watt continuously improved the steam engine over time to expand its industrial application and not limit its use to the world of mining. Among others, he invented the double acting piston, a mechanism necessary to transform the alternative movement into a rotative movement, and the ball regulator, a system intended to regulate the speed rotation of the steam engine. Thanks to all these incremental improvements, the steam engine could be adapted to the traditional machines and new inventions of industrial sectors like the textiles and metallurgy. In parallel, technological advances in the metallurgy
and the construction of machine-tools allowed to improve the performances of the steam engine. From 1800, date of expiration of Watt's patent, a number of new inventors contributed to further develop the steam engine. For instance, Trevithick invented the high-pressure steam engine. In 1804, he also developed the first full-scale working railway steam locomotive. This invention, which inspired the future Georges Stephenson's *locomotive rocket*, announced a great revolution in the transport sector. In 1805, the British engineer Arthur Woolf patented his high pressure compound machine. By the middle of the 19th century, the successive improvements and adjustments brought on the steam engine had given to the revolutionary machine its definitive characteristics.

### 2.2 The Textile Industry

In the textile industry, the first potential signs of mechanization appeared in the area of weaving with the invention of the *flying shuttle* by John Kay in 1733. This new machine, propelled by an innovative mechanical process, drastically reduced human intervention and enhanced the weavers' productivity. Moreover, it made possible the weaving of greater-width fabrics. Its progressive diffusion within the textile industry broke the traditional link between the weaving and spinning activities. The spinning methods indeed relied on the motion of spinners' fingers and so were highly unproductive. In order to solve this problem, described as a "bottleneck", and stimulate inventors' creativity, the English Society for the Encouragement of Arts, Manufactures and Commerce promised a fifty-pound reward to anyone able to develop a new machine capable of simultaneously spinning six yarns of cotton, linen or hemp. In practice, it took until the last third of the 18th century to observe the emergence of two inventions that decisively opened the way to mechanical spinning, namely the Hargreaves *spinning-jenny* (1765) and the Arkwright *water frame* (1769)\(^3\). After some incremental improvements, the *spinning*

\(^3\)Yet a mechanical spinning machine, the *roller spinning machine*, was patented by John Wyatt and Lewis Paul in 1738. The two men even opened a mill which used the new machine in Birmingham in 1742. But it was too early. The market was not ready to absorb the new technology and the mill finally went bankrupt. This historical example shows well the importance of the timing for a new technology to break successfully into the market.
*jenny*, only driven by human strength, allowed to produce up to 120 times more than the old spinning wheel. But its use was highly problematic for some delicate operations. Initially powered by horses or water, the *water frame* offered new potential for mass production. But this new invention could not be fully exploited in a protoindustrial system. That’s why Arkwright founded in the village of Cromford a new spinning mill, one of the first modern factories. In order to reduce the high dependence of his industrial activity on hydraulic energy, Arkwright then decided to implement a steam engine on his *water frame* and called on James Watt’s services for the technical installation. The first steam-powered spinning machine was born. In 1779, Crompton created a new spinning machine, namely the *mule-jenny*, a kind of hybrid of the *spinning jenny* and the *water frame* which combined their respective advantages. The new invention required water or steam energy to work and definitely turned the spinning activity into industrial operation. Between 1786 and 1801, the price of the cotton yarn was dramatically reduced in Britain. In 1811, 150,000 spindles were working on *jennies* in proto-industrial structures, 300,000 on *water frames* and 4.5 million on *mule jennies* in factories (Verley, 1997). The *mule-jenny*, in its principle, remained the engine of the British textile industry for more than one century.

The successful textile inventions, which transformed the textile industry, were perfected, completed and increasingly automated over time, leading to a gradual concentration of the textile production. The first textile machines were essentially designed for cotton work whose potential for mechanization was seemingly higher. Subsequently, the textile machines were progressively adapted to the large-scale production of linen and wool fabrics, and to the production of more luxurious goods. Thus, Philippe de Girard, probably encouraged by the monetary prize offered by Napoleon, brought significant improvements for the linen-based spinning activities. The Jacquard loom (1804) enhanced the production of decorated and more luxurious goods. Thimonnier invented the first sewing machine in 1829. One year later, he opened the first world’s machine-based clothing manufacturing company in order to honor a military order. The sewing machine was then improved by inventors such as Walter (1834) and Howe (1846).
The introduction of new high-impact machines greatly improved workers productivity in the textile industry and moreover opened the way to the industrial application of new discoveries in the field of chemistry. The chlorine-based bleaching process perfectly illustrates this last point. In 1774, Scheele isolated a new chemical element later called the chlorine. He also observed its destructive properties on vegetal colors. In 1785, the French chemist Berthollet published his famous study on the bleaching properties of chlorine. This scientist understood early the commercial potential of developing a new chlorine-based bleaching process which could support the rapid growth of the textile industry. Indeed, the textile fabrics were traditionally bleached through time-and-space-consuming (highly expensive) operations. In order to exploit commercially his discoveries on chlorine, Berthollet established a new chemical manufacture in the village of Javel, near Paris, and started the production of bleaching goods, great part of them were sold to the textile sector. Across the Channel, the Tennant’s factory, located in Glasgow (Scotland), also met economic success by commercializing bleaching powder. The bleaching products were, among others, composed of sulfuric acid and sodium carbonate or soda, two elements also used in the glass, soap and paper industries. The sulfuric acid could already be produced in large quantities thanks to the lead chamber process invented by Roebuck in 1746. By contrast, the soda was mainly extracted from the ashes of some natural plants such as the salicorne. Anticipating an upcoming shortage, the King Louis XIV of France, with the French Academy of Sciences, offered a monetary prize for a method able to produce sodium carbonate from sea salt. Nicolas Leblanc (1791) then came up with a solution, namely the Leblanc soda-making process that he patented and exploited commercially in his manufacture located in Saint-Denis. The Leblanc soda-making process contributed to creating a number of industrial giants like Saint-Gobain (1806) and Kuhlmann (1825). The sulfuric acid and soda remained the two basic products of the chemical industry in the first half of the 19th century.

The chronological evolution of the technological inventions in the textile and chemical industries during the first industrial revolution illustrates the notion of “technological interdependence” between economic sectors and
shows well the role potentially played by the “shortages” and “bottlenecks” in stimulating and guiding creativity and innovation. Thus, the successive textile inventions all proved to be important to restore some kind of equilibrium in the textile industry. But the economic importance of the textile inventions also relied on their potential impact on other industrial sectors. By way of illustration, mechanical spinning provided incentives to improve the steam engine, to extract more coal and iron and, hence, to enhance the extraction techniques, to produce better material (wrought iron, steel, etc.), to develop new machine-tools (lathes, milling machines, etc.), and to commercialize new products like the bleaching powder. Independently from its effects on other industries or manufacturing activities, which are particularly difficult to measure, mechanical spinning is nowadays considered as one of the most economically important inventions of the industrial revolution because it made the expansion of the cotton industry possible everywhere in Europe, especially in Britain.

2.3 The Metalworking Industry

Though Britain already imported pig iron and iron bars from countries like Sweden, Germany or Spain, the techniques used in the metalworking industry at the end of the 17th century were satisfying to meet the demand for iron in Britain (Verley, 1997). The cast iron was traditionally produced in small blast furnaces using charcoal, obtained from the heating of wood, and iron ore directly extracted from mines. Afterwards, the cast iron was transformed into iron through a set of different operations. The iron was then forged using a hydraulic hammer to remove the impurities. Some entrepreneurs early attempted to substitute charcoal and so timber by coal. In that direction, several processes were patented during the 17th century. One pioneer was the entrepreneur Clement Clerk who contributed to the introduction of coal in the metallurgy, notably by applying coal-fueled reverberatory furnaces (cupolas) to various metallurgy processes. But the first results were not successful. In particular, the cast iron was highly breakable. The problem was technically solved by Darby I who had the original idea to produce cast iron by mixing coke, a product resulting from the destructive distillation of coal,
with dust and turf. To exploit commercially the new production process, Darby I left Bristol in 1709 and set up the first coke-fueled blast furnace in Coalbrookdale where both iron ore and coal were available in large quantities. Interestingly, Darby I had previously worked as an artisan in other types of foundries (copper smelter and lead smelter) that already commonly used the coke. Although Darby I succeeded in making the new process economically profitable by applying it to some specific filling operations, the new cast iron making-process took time to spread within the British economy, essentially for two reasons. First, the new process required a huge upfront investment. Second, the cast iron made from coal was still of lower quality than the one made from charcoal, and was consequently cheaper. But the economic conditions seemingly changed around 1750 in Britain. The growing demand for metal products, in combination with the deforestation and the rarefaction of timber, caused a sharp rise in charcoal price in Britain. By contrast, the coal price was a downward trend as a result of technological advances in the field of extraction. These evolutions would have encouraged the forge owners and the metallurgists to adopt the new coke-based cast iron making-process which had meanwhile improved. If only 5% of the cast iron production was based upon the use of coke in 1750, this figure climbed to 55% in 1775 and 77% in 1785. Over the same period, between 1750 and 1790, the cast iron production was multiplied by around three (Verley, 1997). It was henceforth possible to produce cast iron in large quantities. But the large scale production of pure and resistant iron was not possible without progress in both refining and forging operations. In 1783, Henry Cort substituted the forging operation by the rolling operation and introduced the rolling machine in the iron production process. In 1784, the same Henri Cort patented the puddling process, consisting in decarburizing the cast iron in a reverberatory furnace. The pudding process, which was improved later by Crawshay, greatly enhanced the refining operations. Thanks to all these technological advances, wrought iron eventually turned into a common product that was increasingly produced in large steam-powered factories or furnaces migrating from woods and rivers to coal fields, and substituting the steam engine for the water mill. The new metallurgy was born.
By the late 18th century, the British industrialist Darby III, aware of the economic advantages of vertical integration, owned not only coal and iron mines but also blast furnaces and forges or foundries. The Wilkinson brothers controlled several factories both in Wales and in England. In the 1770s, Darby III and Wilkinson brothers had incidentally gone into partnership to build the first world arch bridge to be made of cast iron in England (West Midlands), namely the iron bridge that crosses the river Severn in Shropshire. The iron bridge was opened to traffic in 1781. The construction enthralled the whole Britain as it was previously unthinkable due to the size of the structure and the material (iron) cost. In order to expand their industrial empire, the brothers Wilkinson also participated in the development of the French metallurgy. In 1785, William Wilkinson thus carried out the first iron melting operation based on coke on the French territory in a factory located in Le Creusot.

During the first half of the 19th century, new processes and technological inventions contributed to further progress in metallurgy. For instance, the Scotsman Nasmyth and the Frenchman Bourdon came up together with a new invention, namely the steam hammer (1839). Based on the same principle as the drop hammer used to grind grains in agriculture, the steam hammer made possible the handling and forging of increasingly voluminous and diversified metal workpieces like the propeller shafts of steam boats. Iron material was seemingly satisfying to meet the needs of the industrial revolution until around the middle of the 19th century. Up to then, the use of steel was limited. Only produced in small quantities, steel was relatively expensive. But everything changed in a few decades with the introduction of new inventions that transformed the steel industry forever. Thus, the Bessemer process, patented in 1856, paved the way to the mass production of steel. However, it could only be applied on non-phosphorous cast iron. This reality led the French brothers Martin to build, in 1864, the first reverberatory furnace able to produce steel by using recycled scrap iron. This new process arguably constituted one of the first historical attempts to create a kind of circular economy. The solution was particularly valuable in France where non-phosphorous iron ore deposits were limited, so making the French metallurgy highly dependent on
imports. In 1877, Thomas and Gilschrist developed a new converter that finally unlocked the potential of phosphorous iron ore for the production of steel. The solution considerably enhanced the capacities of the steel industry in Britain and continental Europe. In France, the Thomas-Gilchrist process even shifted the gravity center of this industry, contributing to the economic development of regions like Lorraine. The continuous progress in the steel industry backed the large expansion of the rail sector during the second half of the 19th century.

2.4 The Transport Industry

2.4.1 Canals and Roads

Due to its privileged geographical situation, surrounded by sea, crisscrossed by rivers and penetrated by large estuaries, Britain naturally benefited from favorable conditions for water transport. At dawn of the 18th century, short sea shipping (coastal navigation) already connected the most important ports like Newcastle, Liverpool, London, Bristol or Hull, from where merchandise could be easily transported inside the country through rivers. The British roads contrastingly were poor until the mid-18th century. Between 1750 and 1800, Britain undertook massive investments to improve and expand its transport infrastructure. For instance, a large network of canals was constructed to link the main river basins of (essentially) the northern and central industrial regions, and enhanced the commercial exchanges between these industrial centers. These canals were almost entirely financed by the private sector. By way of illustration, from 1759 to 1776, Francis Egerton or the Duke of Bridgewater, probably inspired by the Canal du midi discovered on a trip round Europe, funded the construction of a new waterway, namely the Bridgewater canal, which connected his coal mine, located in Worsley, to Manchester. The project, supervised by the engineer James Brindley, largely contributed to the prosperity of the city of Manchester, especially as the cotton industry and the use of the steam really started to develop. The Duke of Bridgewater also provided funding for the Trent and Mersey canal, also called the Grand Trunk, which was opened to traffic in 1777. This colossal project required a
huge amount of money which was collected through the creation of an original joint stock company gathering both industrialists and mine owners. The entrepreneur Josiah Wedgwood, the grandfather of the naturalist Darwin, was part of this project. Sometimes considered as one of the fathers of marketing as he understood early the potential role of fashion on consumption, Josiah Wedgwood got interested in the construction work because his business - he held a pottery - was highly depended on the safe and smooth transport of his products. The canals thus proved to be important for the transportation of such fragile objects, but also for the transportation of heavy goods such as wood and coal.

The road infrastructure was also subject to investment during the 18th century Britain. In particular, an integrated network of turnpike roads, co-ordinated to a greater or lesser extent with the canals, was financed by local associations of landowners and/or businessmen. By 1800, the road infrastructure had already significantly improved but still remained limited. Nonetheless, it contributed to training a new generation of engineers, like Telford and McAdam, which dramatically enhanced the techniques used for public works in the beginning of the 19th century. In France, by contrast, the construction works were financed in large part by the public administration. The Royal roads, built in the last third of the 18th century, constitute a good illustration. The project consisted in building a network of roads connecting the capital, Paris, to the major French cities, and the latter together. These roads reduced greatly the travel time between the concerned cities. For instance, the travel time between Paris and Bordeaux went from fourteen days in 1764 to five days in 1780. But the road infrastructure showed important regional inequalities. In addition, the secondary roads were most of time terrible, limiting the access to rural areas and so the potential of the countryside market.

In spite of continuous improvements in the road infrastructure, the only dense transport network at dawn of the 19th century was the waterways. This network was systematically used and expanded over time. Thus, between 1800 and 1840, the total length of canals went from 2690 to 3470 miles in Britain. In 1822, France adopted a program intended to triple the total length of French canals before 1840. The start of the industrial revolution
really corresponded to the age of rivers and canals. Waterways were the arteries of the industrial revolution. And technical progress rapidly broke into this new sector of massive trade flows. By way of illustration, wooden sailing ships progressively turned into steamships. The idea of utilizing steam to propel boats can be traced back to Papin’s publications. But the first successful steam boat, the one which run up the Saône in Lyon from Saint John’s cathedral to Barbe island, was the work of Claude Jouffroy d’Abbans in 1783. Once its potential for commercial use was fully proved, the steam rapidly spread over the waterways. In 1818, a regular steam line was established between Glasgow and Belfast. In 1819, the Savannah came to be the first mixed boat, propelled with both steam and wind, to cross the Atlantic. By 1822, the line London-Paris was performed by the Aaron Manby, an iron-made steamboat. The same year, Savage patented the ship’s propeller which offered new opportunities to navigation by making boats more weather resistant and manageable. In 1837, the so-called Great Western, a 71-meter-length steamboat, started to ensure a regular transatlantic line. Navigation techniques substantially improved over time and the steamboats became increasingly performant, making possible or much easier the massive transport of people and merchandise. Large companies specialized in international trade, like the British India or the General Transatlantic Company, developed very fast. Ports were extended. And a decline of the transportation costs was observed on most of national and international lines. By way of illustration, the transport price for a ton of freight was divided by four between 1820 and 1850 on the transatlantic commercial line Anvers-New-York.

The road infrastructure also captured the attention. In Britain, many other turnpikes were built and the road infrastructure considerably expanded in the first half of the 19th century. In France, the Public Road Administration financed the construction of new roads and launched a great project designed to restore and improve the Royal roads, which had particularly suffered from the French Revolution and the first empire. In 1848, the travel time between Paris and Bordeaux had dropped to only two days. In 1824, among others, steam power had the advantage of making boats less dependent on natural conditions, especially on wind.

\[ \text{Among others, steam power had the advantage of making boats less dependent on natural conditions, especially on wind.} \]
a French law initiated the construction of vicinal roads which were initially supposed to open the countryside. A law of 1836 reinforced the development of both the regional and vicinal roads. The transport infrastructure investment (roads and canals) consumed a lot of capital in Britain and France. For instance, in Britain, it absorbed around 9% of total investment between 1821 and 1830 and around 7% between 1831 and 1840, decade during which the railway took over (Verley, 1997). The roads and canals played a key role on the capital market by making necessary the creation of new limited liability companies, which arguably were the only private structures able to collect the huge funds required by the construction works. The railway benefited later from these advances.

2.4.2 Railway

The earliest rudimentary signs of railway date from the 6th century B.C. with the construction of the Diolkos, a paved guided trackway that enabled boats to cross by land the Isthmus of Corinth and thus avoid the perilous circumnavigation of the Peloponnesian peninsula (Lewis, 2001). The first horse-drawn wagons, which travelled on tracks made of notched stones, appeared during the Antiquity in Ancient Greece and the Roman Empire. But this kind of transport was limited. The development of railway really began in Europe after the Middle Age. It was narrowly linked to the world of mining. The mine operators indeed rapidly realized that the wagons, loaded with heavy products like coal and ore iron, faced less rub resistance when they were towed along railway lines. The first railway tracks were made of wood and towed by men or horses. By the 17th century, wooden railway lines were thus common in Britain for carrying heavy coal from mines to canals where it could be loaded on boats. Thanks to the parallel development of metallurgy in the 18th and 19th centuries, the wooden rails were then gradually replaced with (cast) iron rails and later with steel rails. The Coalbrookdale Company is known to have experimented the first iron plate railway in its mines in 1768.

Steam broke into the rail sector in the beginning of the 19th century. The first steam locomotive was developed and tested by Richard Trevithick in Merthyr Tydfil (Wales) in 1804. The technology was then enhanced by differ-
ent skilled engineers to make it fully operational. In 1814, Robert Stephenson, chief mechanic at Killingworth mines, thus built a new model of locomotive, called the Blucher, which could tow 8 wagons loaded with 30 tons of coal at a speed of 6 km per hour. One year later, he considerably improved the technology to gain in traction power. This improvement turned the Blucher into the first effective steam locomotive. Initially, the Blucher was solely working in the experimental world of the collieries. But in 1821, Stephenson convinced an entrepreneur, who had first planned to construct a horse-drawn railway line between Stockton-on-Tees and Darlington, to utilize a steam locomotive. In 1825, the locomotion, which could transport until 450 people at a speed of 24 km per hour, thus became the first public passenger steam train in Britain. In 1827, the engineer Marc Seguin, involved in a professional relationship with Stephenson, patented the great tubular boiler. The invention was rapidly applied on the Lyon-Saint-Etienne line (1827).

As part of a new project intended to construct a new railway line between Manchester and Liverpool, and so facilitate the traffic flows between the two cities, an original and spectacular competition was organized in Rainhill (Rainhill trials) to select the means of traction for this line (locomotives or stationary engines). The steam locomotive designed by Stephenson, namely the Rocket, won the competition. By 1830, the rocket, an “iron horse” which could reach a speed of 58 km per hour, was operating on the Manchester-Liverpool line in Lancashire. The age of railway started. Indeed, the Manchester-Liverpool line proved to be so successful that a railway fever swept through Britain and Europe. Numerous railway companies saw the light of day and many railways lines emerged on both private and public initiatives. Built using British techniques, the first Belgian railway line, linking Bruxelles and Malines, was put on service in 1835 in the presence of the King Leopold I. The Malines-Anvers line and Malines-Termonde line followed rapidly. In Belgium, Malines was actually at the heart of the Belgian rail network that connected the main Belgian industrial centers to Anvers, which houses a commercial port, France and the Zollverein. In Germany, the first steam-powered railway, which linked Nuremberg and Fürth, was also opened in 1835. In France, Rothschild and Lafitte financed the Paris-Saint-Germain advertising
railway line in 1837. In 1845, Paris was connected to Lille. In the USA, the railroad era also began in the 1830s when Peter Cooper’s *Tom Thumb*, the first American-built steam locomotive, travelled for the first time between Baltimore and Ohio. In 1833, a new national railroad connected Charleston to Hamburg in South Carolina. In Britain, the total length of the British rail network was around 800 km in 1840. The figure climbed to 10500 km in 1850 and to 24500 km in 1870. By 1848, the steam locomotive developed by Crampton exceeded a speed of 100 km per hour.

The second half of the 19th century was a period of massive railway investment in both Europe and North America. The railway network expanded spectacularly in the Western countries. In addition to the principal profitable lines, a fabric of secondary lines, often subsidized by the public sector, emerged on national territories. The railway crossed the rivers (metal viaducts) and ran through the mountains (channels). It accelerated urbanization and modified the cities’ architecture. It boosted the economy and created lots of direct and indirect jobs through labor-intensive activities (construction, operation and maintenance). It also contributed to laying the foundations for modern credit (modern banking system) and opened bankers’ eyes on the profit opportunities offered by industrialization. Thanks to important ripple effects on the whole economy, the rail speeded up the industrial revolution and definitely placed it into orbit. It also redefined regional hierarchies by inducing some economic specialization in agriculture and industry. The rail development index computed by Bairoch (1974), which takes into account the countries’ length of the railway network, interestingly proves to be a good measure of countries’ industrial power at dawn of the 20th century.

3 Before the Industrial Revolution

3.1 Evolution of English GDP per Capita in the Early Modern Period

There is an increasing debate in the literature concerning the evolution of British income per head and British real wages in the centuries before the
industrial revolution. While the industrial revolution has traditionally been viewed as a key break in world economic history, characterizing the passage from stagnant societies to modern ones, and marking the starting point of the “Great Divergence” between Western countries and the rest of the world, a new view has recently emerged to support the concept of a “Little Divergence” taking place in the centuries before the industrial revolution. Clark (2011a, 2011b) makes reference to these two competing views as a new debate between “Malthus” and the “rebell of the early modernists”.

The Malthusian view, which is notably adhered to by the California School world historians, contends that all economies were trapped in Malthusian dynamics before the industrial revolution period. Income per head fluctuated around a subsistence level - defined as the level of income at which birth and death rates are equal (Nuvolari and Ricci, 2013) - but exhibited almost no upward trend, a situation described by Persson (2008) as “the Malthus delusion”. Thus, Clark (2007, 2011a, 2011b, 2013, 2014, 2016a) has repeatedly claimed that English income per capita in 1800 was not higher on average than in most of its history since 1200, and perhaps more surprisingly since as far back as the hunter-gatherer era. In line with the Malthusian view, many scholars have argued that Europe, by comparison with other civilizations, did not enjoy any economic advantage in terms of living standards and income per head when breaking into the 18th century. In parallel, a number of comparative studies have revealed that other parts of the world, especially East Asia, seemingly shared many characteristics with Britain and Europe’s economy and society in the 17th and 18th centuries, like long-distance trade, secure contracts and property rights, or consumption habits, and were even more advanced in certain areas (e.g., Wong, 1997; Li, 1998; Pomeranz, 2000; Vries, 2003; Parthasarathi, 2011; Goldstone, 2002, 2008, 2012).

The alternative view, which is defended by the “early modernists”, argues that Europe was already progressive, intellectually and technologically, since the late Middle Age, or at least the Renaissance, making successful advances in areas such as architecture, manufacturing, exploration, shipbuild-

---

5The expression “early modernists” was first used by de Vries (1994) to describe the challenging view that Europe was not stagnating in the preindustrial period.
ing, navigation, agriculture and science, a situation often described as the “Rise of the West” that would have contrasted with the stagnation of most other civilizations (Goldstone, 2015). From an accounting perspective, the “early modernists” have actually detected a small but steady acceleration of growth in Northwestern Europe, at least in two countries, England and the Netherlands, prior to the industrial revolution period (e.g., Wrigley, 1985; van Zanden, 2002; Broadberry and Bishnupriya, 2006; Allen, 2008; de Vries, 2008; Maddison, 2008; Persson, 2008; van Zanden and van Leeuwen, 2011, 2012; Broadberry et al., 2011, 2015; Nuvolari and Ricci, 2013; Broadberry, 2014; Broadberry, Guan and Li, 2014; Bolt and Van Zanden, 2014; Broadberry, Custodis and Gupta, 2015). For instance, Broadberry et al.’s (2011, 2015) estimates, based on an output-based approach, show a persistent upward trend in British GDP per capita that would have approximatively doubled between 1270 and 1700 (see figure 2). Using a demand-side approach, Nuvolari and Ricci (2013) find that the English economy was of “Malthusian type” between 1250 and 1580, then showed capacity to relax some of the Malthusian constraints between 1580 and 1780, achieving a positive rate of economic growth, though lower than those suggested by Maddison (2008) and Broadberry et al. (2015). According to the “early modernists”, the British industrial revolution would then root, at least to some extent, in progressive trends going back to the late Middle Ages (e.g., van Zanden, 2008; Broadberry, 2014; Studer, 2015; De Pleijt and van Zanden 2016). In line with this view, a number of scholars argue that initial increases in income per capita were crucial for the transition towards modern society and sustained economic growth (e.g., Galor and Weil, 2000; Hansen and Prescott, 2002).

Along with the construction of historical time series tracing the evolution of GDP per capita in Britain and Northwestern Europe, several studies have also been produced to provide estimates of preindustrial real wages and check the “Little Divergence” hypothesis (e.g., Allen, 2001, 2005a, 2009a; van Zanden, 2005; Broadberry and Bishnupriya, 2006; Clark, 2007b; Hersh and Voth, 2009; Panuk and van Zanden 2010).

The great difference in views between the California School World Historians like Clark (2010a) and the “early modernists” like Broadberry et al. (2015), for the British case, come from conflicting estimates of English income around 1400-1450. While Clark (2010a) find a relatively high income, leading to a Malthusian interpretation of preindustrial English economy, Broadberry et al. (2015) estimate a relatively low income, more consistent with the slow-but-modern view of English economic growth.
Figure 2: GDP per capita in England and the United Kingdom since 1270 (Adjusted for inflation and measured in British Pounds in 2013 prices)

Data source: Broadberry et al. (2015) and Bank of England. Note: Data refers to England until 1700 and the UK from then onwards.

If one is willing to acknowledge an expansion of GDP per capita and wages in preindustrial Europe, did this evolution really prefigure a process of modern economic growth in the Western World? While some scholars, especially the “early modernists”, have clearly positioned themselves in favor of a growth led, at least partly, by technological change and TFP improvements (e.g., van Zanden and van Leeuwen, 2011, 2012; Moller and sharp, 2014), other scholars have rather stressed Malthusian mechanisms, i.e. changes in birth and death schedules, to explain the rise in income per capita observed in preindustrial Northwestern Europe. In a study ironically entitled Malthusian Dynamism and the Rise of Europe: Make War, Not Love, Voigtlander and Voth (2009) thus contend that two European “inventions”, namely a peculiar marriage

In line with this view, many studies report evidence of (technology-led) productivity growth in economic sectors such as agriculture (e.g., Tarrow, 2007; Ang et al., 2010, 2013), shipping (e.g., van Zanden and van Tiellhof, 2009; Unger, 2011), road freight (e.g., Gerhold, 1996; Bogart, 2014), watchmaking (e.g., Kelly and rada, 2014b, 2016) or glassmaking (e.g., Barker, 1977) before the industrial revolution.
pattern - the European Marriage Pattern (EMP) - and a specific mortality regime, may well account for the rises in income observed between 1300 and 1800. In another study, reasserting that productivity growth played only a little role in the evolution of British and European incomes between 1350 and 1700, Voigtlander and Voth (2013b) design a neo-Malthusian economic model with two sectors and multiple steady states to explore the growth of income per capita and urbanization during this period in Europe. Among others, the authors show that a negative shock to the population size, as the Black Death was⁹, may have triggered some of the demographic changes observed in the preindustrial period as it rose not only the real wages, but also - because of the supposed Engel law - the size of the cities, the latter being characterized by higher death rates, where the urban goods, the superior ones, were produced. Higher wages moreover allowed higher levels of taxes, most of which were devoted to financing wars, whose effect on population size was clearly negative. In the model, the increase in the death schedule induced by the initial shock on population size and higher wages, a relation labelled as the “horsemens effect”¹⁰, contributes to reducing the population pressure, making higher equilibrium incomes sustainable.

Alternatively to the Malthusian interpretation of the evolution of preindustrial GDP per capita and real wages in Northwestern Europe, Goldstone (2015) proposes an explanation based, among others, on Smithian mechanisms, and finally rejects the concept of a “Little Divergence” as a launch pad for modern growth. According to Goldstone, a careful examination of the recent GDP per capita series reveals a pattern that is typical of pre-modern “efflorescences” in Britain and Holland, just as had occurred in earlier periods of “efflorescence” in Song China, Renaissance Italy, or the Roman Mediterranean world as recently documented by Endkamp (2016)¹¹.

⁹See e.g. Pamuk (2007) and Clark (2016b) for studies discussing the economic effects of the Black Death.

¹⁰In the standard Malthusian models, the death rates are contrastingly typically downward sloping in income, an evolution labelled as the “preventive check” hypothesis, while birth rates are either flat or upward sloping, an evolution labelled as the “positive check” hypothesis. See e.g. Nicolini (2007) and Kelly and Ó Grada (2012) for a test of these hypotheses in preindustrial England.

¹¹Goldstone (2015, p.1) notes: “According to the new data, at no point after 1600 and
3.2 A Model of Transition towards Modern Industry: Proto-industrialization

In *Capital*, Marx early wondered which specific form of preindustrial labor organization had been the most favorable for the transition towards modern industry and mechanized factory (Verley, 1997b). Without ignoring the diversity of the industrial structures and practices, historians have long been interested in the non-mechanized manufacture that concentrated an abundant labor force. In many cases, this manufacture indeed turned into modern factory at the age of industrialization. But another model of transition, namely the proto-industrialization, has been increasingly studied since the publication of Mendel’s thesis on Flanders in the 18th century. Mendel’s theory of proto-industrialization was actually an attempt to associate the movement of rural industrialization observed in Northwestern Europe in the 18th century with the creation of the modern industry. Mendels (1969, 1972) more particularly highlighted the fast-growing development of the *putting-out system*, a decentralized method of production for manufactures. It was traditionally characterized by a mixed form of labor organization combining rural and urban aspects. His principle was simple. A merchant or urban entrepreneur, called the *putter out*, provided the raw materials to home-based rural workers, who were then in charge of producing some goods in some predefined way, often with rudimentary tools. Once the goods were produced, the merchant collected and sold them on external markets. This point introduces the key distinction between the *putting-out system* and the *domestic system*. In most cases, rural workers were agricultural farmers - and their family - who viewed proto-industrial activities, essentially consisting of repetitive unskilled tasks, as the opportunity to earn an extra income during the off-peak agricultural periods. In Mendel’s terms, the development of this new rural industry allowed to destroy the traditional forms of industrial production and constituted a preliminary phase to industrial progress. According to Rioux

*before* 1780 *did any nation in Europe experience both significant population growth and significant per capita income growth, as would be necessary for modern economic growth to have emerged. The new GDP/capita data make it clear that in fact China and Europe, both in their leading regions and overall, were on very similar economic trajectories until after 1800 [*...*] Modern economic growth arose only in a late “Great Divergence” after 1800.*
proto-industrialization is the proof that preindustrial economies were not stagnant. In a few regions, some forces were acting in silence and finally led to great industrial changes. Mendels (1969, 1972) identified several reasons which can potentially explain the emergence of proto-industrial activities in Europe in a context of expanding trade. The advantages of using rural workforce to increase the industrial output were seemingly multiple. For instance, rural workers were cheaper than urban workers. Rural workers indeed accepted a lower wage as it was just an extra income. The proto-industrial system can thus be viewed, at least to some extent, as the result of managerial rationality, which consisted in “outsourcing” the unskilled tasks linked to some specific production (clothes, clocks, iron-made products, etc.) to rural workers, while entrusting the other tasks to more-skilled workers in urban areas. With respect to this point, the example of the textile industry is particularly illustrative. In this sector, the spinning activities were traditionally performed by rural workers, while the other operations, both the downstream and upstream operations, were usually performed in urban workshops. The proto-industrial system was also an opportunity to escape from the constraining corporative rules of the cities and/or to avoid the social problems related to urban growth. If the proto-industrial system contributed to deindustrializing a number of urban areas, it helped, on the other hand, to further develop the commercial, financial and consumption functions of cities. Mendels (1972) also formulated some assumptions regarding the regions which were more likely to see the emergence of proto-industrial activities. They concern, among others, demography, agriculture, urban crafts and the access to external markets. These conditions have been heavily criticized in the literature as the economic reality often contradicted them. The environment in which proto-industrial activities emerged and prospered was variable and so difficult to generalize.

The theory of protoindustrialization has met a great success. Following Mendel’s work, the existence of proto-industrial structures was documented almost everywhere in Western Europe, like in the German Ruhr area, Yorkshire and Lancashire, Alsace and Lombardy, Lyon and Calvados, Languedoc, North Italy, from Ireland to Switzerland and Poland to Andalusia. While the
model proposed by Mendel (1972) seems relevant to describe the transition toward modern industry for some European regions, it nevertheless fails for other regions. The geographical continuity between proto-industrial activity and modern industry was indeed not observed in every proto-industrial region. Thus, the examples of regional “deindustrialization” are numerous. For instance, in France, it was the case for Brittany, Languedoc and Calvados. In England, Coleman (1983) concludes that only four out of the ten identified proto-industrial regions, namely Yorkshire, Lancashire, Trent and Western Midlands, successfully went through modernization. Different factors can be advanced to explain these failures (Verley, 1997). They include the disruption of trade flows, the loss of position revenues due to growing markets, the lack of industrial motivation from local capital owners, and the excessive energy cost in a context of rising dependence of industry on steam power.

4 The Causes of the Industrial Revolution

The first waves of studies on the industrial revolution mainly seek to identify its social and economic effects. They highlight, among others, the growth of industrial production, the development of cities, the rise of the factory system, and the emergence of new social classes. They emphasize the key role of technological progress in destabilizing preindustrial societies but do not make a great deal of effort to analyze the conditions under which the industrial revolution could take place. The work of Mantoux (1906), one of the first historical syntheses of the industrial revolution, is very typical of this approach (Rioux, 1989). By contrast, the economic historian Ashton (1948) relegates the technological inventions to a lower level of priority, contending that they were obviously essential but could not emerge, bloom, and eventually launch a process of modernization and sustained efficiency growth out of a conducive environment. As Mokyr (1999, p. 12) notes: “Inventions do not rain down upon an economy like manna from heaven. They are stimulated by economic and social pre-existing conditions. They emerge in the minds of some people for some reason that may or may not be identified, are communicated, adapted, refined, implemented, and imitated”.

39
The main goal of this section is to provide an integrated survey of the main factors that have been advanced in the literature to explain the British industrial revolution. As a preamble, it is noteworthy that the notion that some factors were either necessary or sufficient for the British industrialization to take place has become increasingly hard to defend (Gerschenkron, 1962, pp. 31-51). As Mokyr (1999, p. 19) notes: "Some factors present in Britain truly facilitated the Industrial Revolution and in this sense can be said to be causal. Others impeded its progress, and the Industrial Revolution proceeded in spite of them. The term ‘facilitated’ does not mean, however, that there were any elements that were indispensable. After all, factors that were neither necessary nor sufficient for the outcome can still be thought of as causal. For instance, heart attacks cause deaths, though not all deaths are caused by them and not all heart attacks are fatal. Moreover, insofar as heart attacks are themselves caused by other factors, it is debatable to what extent they are ultimate causes or just transmission mechanisms. The causal explanation of the British industrial revolution runs into very similar quandaries”.

According to Clark (2003, p. 14), any convincing explanation of the British industrial revolution has to do with the following two things: “First explain why no society before 1800 - not ancient Babylon, Pharaonic Egypt, China through countless centuries, Classical Greece, Imperial Rome, Renaissance Italy, medieval Flanders, the Aztecs, Mogul India, the Dutch Republic - expanded the stock of knowledge by more than 10% a century. Then explain why within 50 years of 1800 the rate of growth of knowledge rose to modern rates in one small country on the margins of Europe, Britain. And of course explain why economies around the world have benefited from this knowledge expansion to such different degrees. Then we will understand the history of man”. Taking an inventory of the existing theories of the British industrial revolution, Clark (2003) proposes to distinguish among three different categories of theories: the exogenous growth theories, those arguing that exogenous events created the conditions needed for the industrial revolution to happen; the multiple equilibrium theories, those arguing that a specific shock (disease, war, and so on) led the economy to move from the Malthusian equilibrium to the new dynamic one; the endogenous growth theories, those
believing that industrialization was written into the humanity’s genetic code, i.e. dependent on the time evolution of a specific state variable, like the population size (e.g., Galor and Weil, 2000), the market size (e.g., Desmet and Parente, 2010) or the stock of useful knowledge (e.g., O’Rourke et al., 2013), which, beyond some critical threshold, triggers a process of sustained efficiency advance. Pointing out the huge complexity of the causal mechanisms at work in the emergence of the industrial revolution, Mokyr (1999, p.19) argues in favor of “positive feedback and interactive path-dependent” models, showing that monocausal, linear models based on concepts of equilibrium or steady states have difficulty doing justice to historical reality. Independently from the discussion on the type of model that is suitable to capture the reality of the industrial revolution, (economic) historians have increasingly accepted the idea that the industrial revolution first took place in the 18th century Britain because a package of technological inventions appeared in a conducive environment. For that reason, we need to observe the characteristics of the national and international environment in which Britain successfully underwent an industrial revolution.

4.1 Geography and Natural Resources

Britain’s geography and geological conditions have often been highlighted to explain British economic success in history. Surrounded by sea, crisscrossed by rivers and penetrated by large estuaries, Britain benefited from favorable conditions for water transportation which early stimulated the commercial exchanges between the British ports and the coastal regions. The sea moreover constituted a natural barrier which likely dissuaded enemies from invading the country, contributing to some relative peace and stability on British territory. Being an island also provided Britain with great incentives to develop and continuously improve the navigation techniques, which eventually translated into British maritime superiority. Maritime power helped Britain to defend its territory, to win commercial wars, and to play an important role in trade. Along with a privileged geographical situation, Britain is also often believed to have enjoyed favorable geological conditions. Indeed, the country had abundant natural resources like iron and non-ferrous metals and was
moreover “built on an underground mountain of coal” (Levine 1987, p. 97). Introduced on a large scale into the British economy, initially to overcome the expected shortage of timber and satisfy the fuel needs of improving technology of transporting, coal exploitation gave Britain access to large quantities of cheap and powerful energy.

The impact of natural factors on economic inequalities among countries has always been subject to debate. It’s widely acknowledged that Britain took advantage of both its geographical and its geological conditions to pave its own path to industrialization. Waterways were the main arteries of commercial flows until the 1800s, and coal truly propelled the industrial revolution. But did the British industrial revolution really depend on geographical and geological factors? Tackling this question, especially the role of waterways, Deane (1965, p. 76) moderately says: “If Britain had had to depend on her roads to carry her heavy goods traffic the effective impact of the industrial revolution may well have been delayed until the railway age”. Regarding the role of natural resources, especially coal, literature is highly controversial. Thus, Wrigley (1962, 2004, 2010) argues that the substitution of inorganic sources of supply in raw materials for inorganic ones was a sine qua non condition to industrial growth. Tracing all the history of the coal industry, Church (1986, p. 758) states that “it is difficult to exaggerate the importance of coal to the British economy”, thus summarizing the thought of several generations of scholars. In Great Divergence, Pomeranz (2000) also attributes a great part of British economic success to accessible reserves of mineral resources, and in particular coal, near population centers. In the same vein, Sieferle (2001) contends that the early use of fossil energy in Britain, mainly due to the rarefaction of wood, explains why this country was the very first to experience a successful industrialization. In line with this view, Eren and Garcia-Macia (2013) describe the British industrialization as a transition from wood to coal as the main source of energy. Recently, Ridley (2010a, 2010b) has also identified coal as the materialist account for British prosperity, claiming that “without coal, Britain would be just another flash in the pan, a golden age that truly produced some luxury and culture and science but no real transformation in living standards” (2010b). Yet a number of cliometric studies tend to
seriously qualify the role played, either directly or indirectly, by coal and the coal sector in the British industrial revolution (e.g., McCloskey, 1981; Crafts, 1985a, 2004; Mokyr, 1990; Crafts and Harley, 1992; Clark and Jacks, 2007; Borowieccky and Tepper, 2015). Based on these results, Ó Grada (2016, p. 230) concludes that “steams - and so coal’s - role at the height of the British industrial revolution, both as a source of energy relative to waterpower and in terms of its contribution to growth, was distinctly secondary”.

In reality, it’s immensely difficult to assess the exact role of coal in the British industrial revolution. In particular, it is not clear how resource availability plays on technological progress (Mokyr, 1999, p. 22). On the one hand, resource abundance is usually viewed as fortunate as it reduces production costs and, in the case of energy resource, makes (potentially) economically viable the expansion of a new “technical system” based on its large-scale industrial use. On the other hand, resource scarcity is likely to promote creativity and innovation by imposing challenges to the economy. As Mokyr (1990, p. 160) argues: “In the absence of coal, the ingenuity applied to using it would have been directed towards replacing it”. According to this view, coal was not a driver for technological progress. It truly shaped the British industrial revolution but did not create it. Resource endowments would then work as a “steering mechanism”, or what Rosenberg (1969) calls a “focusing device”, leading to a national bias in technical activity (Kuznets, 1965). In these conditions, Britain’s technological creativity, defined as the country’s ability to develop new inventions and adapt them - or those made elsewhere - to its highest benefit, would be more important.

Geographical and geological factors perfectly illustrate the notion that factors proposed as “causal” in the literature were hardly either necessary or sufficient. Like Britain, Ireland and Japan were islands. France, based on Verley’s (1997a, p. 219) calculations, had comparable amounts of naturally navigable rivers. The Netherlands also enjoyed good conditions for water transportation and had virtually access to cheap coal, the latter being not a British monopoly. China, Russia and the USA possessed important reserves of coal as well. And coal was anyway a traded commodity. Trade could absorb, at least to some extent, national disparities in natural endowments.
4.2 Demography

An original feature of the 18th century is the evolution of the number of men and women in Britain (e.g., Habakkuk, 1953; Flinn, 1970). Indeed, the British population, and more generally the European population, began to grow sustainably as from the 18th century. Between 1000 and 1700, the European population (including Russia) rose from 43 million to 125 million of individuals (Biraben, 2003). Over this period, phases of demographic crisis and demographic expansion followed each other. The phases of demographic growth were usually observed in times of relative peace, political stability, and modest economic development. The phases of demographic crises mainly resulted from the three misfortunes, also called the “three horsemen of the Apocalypse” or the “three mortal fates” (Sauvy, 1963), of those times, namely the epidemics, the wars and the famine which hit cyclically the European populations. The 17th century was characterized by an overall stagnation of the population in Europe. Between 1700 and 1800, the European population then increased from 125 to 195 million of individuals, with national disparities. In one single century, i.e. between 1710 and 1810, the English population almost doubled (Wrigley and Schofield, 1981; Clark 2010a).

Table 1: Population in England and her neighbors (millions)

<table>
<thead>
<tr>
<th>Country</th>
<th>1600</th>
<th>1700</th>
<th>1750</th>
<th>1800</th>
<th>1850</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>19.6</td>
<td>22.6</td>
<td>24.6</td>
<td>29.3</td>
<td>36.6</td>
</tr>
<tr>
<td>Germany</td>
<td>16.0</td>
<td>17.0</td>
<td>24.5</td>
<td>35.4</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>13.5</td>
<td>13.6</td>
<td>15.8</td>
<td>18.3</td>
<td>24.7</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>1.5</td>
<td>1.9</td>
<td>1.9</td>
<td>2.1</td>
<td>2.3</td>
</tr>
<tr>
<td>Spain</td>
<td>6.7</td>
<td>7.4</td>
<td>8.6</td>
<td>10.6</td>
<td>14.8</td>
</tr>
<tr>
<td>Sweden</td>
<td>1.4</td>
<td>1.8</td>
<td>2.4</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>England</td>
<td>4.2</td>
<td>5.2</td>
<td>5.9</td>
<td>8.7</td>
<td>17.0</td>
</tr>
</tbody>
</table>

Source: Shaw-Taylor and Wrigley, 2014, table 9

A number of scholars early viewed demographic expansion as a factor likely to explain the timing of the British industrialization (e.g., Deane and Cole, 1962; John, 1965; Eversley, 1967; Perkin, 1969; McKendrick et al.,
The Economics Nobelist John Hicks (1939, p. 302) even went as far as to claim that “perhaps the whole industrial revolution of the last two hundred years has been nothing else but a vast secular boom, largely induced by the unparalleled rise in the population”. The proponents of the demographic explanation have stressed both demand-side and supply-side factors to account for how population growth may have spurred a process of industrialization. On the demand side, demographic growth arguably increased the volume of desired industrial goods in the British economy. This would have made necessary a new growth based on technological progress. An alternative demand-side view contends that the rise in demand, which translated into larger consumption markets, would have fostered both investment and innovation by creating a much less risky business environment for entrepreneurs and by modifying their economic expectations. This refers to the typical “Keynesian” argument. On the supply side, demographic growth potentially rose the number of available workers, contributing to decreasing wages and economically harmful rigidities on the labor markets. This, in turn, would have encouraged new original forms of business venture and innovative activity. Furthermore, population growth would have guaranteed the presence of readily available workers for industrial development.

Although appealing at first glance, some major criticisms have arisen over this simple view of the relationship between population growth and industrialization. Europe was not the only region in the world to exhibit a population growth in 18th century. Other regions, such as China and Central America, were also growing in population. Why did not these regions take off? The current experience of the Third-World countries, though it takes place in a radically different world, also tends to show that an acceleration of population growth is not enough to set into motion a cumulative process of industrial progress. Demographic growth was not a sufficient condition, especially as the potential relationship between demographic growth and industrialization was, at best, an indirect one. Indeed, the rise in the total volume of desired products possibly induced by population growth could not translate into effective demand for manufactures and industrial growth without supply-side adjustments, notably in agriculture, able to provide the extra population with
new resources and to overcome the propensity for the marginal productivity of labor to decline (Wrigley and Schofield, 1981, pp. 443-9; O'Brien, 1985, p. 786; Hudson 1992, p. 160). But, as Anderson (1996, p. 267) argues: “There is no theoretical reason why the population pressure should induce these. Indeed, in the alternative way, it was only because they were taking place anyway that population growth could continue”.

The scholars who locate population growth at the heart of the industrial story have in fact to use a dynamic argument which explicitly links population dynamics to technological change and productivity (Boerup 1981). Yet this idea is far from being consensual (e.g., Jones, 1995; Voth, 2003; Mokyr, 2005c; Lagerlof, 2006; Clark, 2003, 2005a, 2014). Not only the effects of population dynamics on demand and technology, but also the idea that population growth favored labor mobility and supply industry with readily available workers has been challenged. In particular, the industrial use of the workforce was not determined by the overall demographic conditions but by the sectoral and regional characteristics of the labor markets. Thus, rural overpopulation in Britain in the first half of the 19th century did not prevent some entrepreneurs, at least in some regions and/or sectors, facing difficulties in recruiting skilled workers (Verley 1997b, p. 242). It may even be that British industry innovated to compensate for the shortage of skilled workers. This point introduces the two competing views of the role of labor in the British industrial revolution. The first view, closely related to Habakkuk’s (1962) work, considers technological progress as the firm’s response to labor scarcity. Accordingly, labor scarcity - and the resulting high wages - was friendly to industrial progress as it induced the new labor-saving technologies of the industrial revolution¹². The second view, based on what Mokyr (1976) calls the growing-up models, suggests that industrial progress was more likely to happen in areas where labor was abundant and cheap (e.g., Lewis 1954, Pollard 1978). In support to this view, Kelly et al. (2015) have recently found that British industrialization was seemingly restricted to low-wage areas.

¹²See e.g. Broadberry and Gupta (2009) and Allen (2009a, 2009b, 2010) for recent theories defending the idea of an industrial revolution induced by factor prices, i.e. by high wages and cheap coal.
Developing around the idea that modern economic growth emerged as the result of a very close interplay between demographic and economic changes, a growingly influential literature has tried to provide theories able to account for both the economic and demographic transitions within the same framework. The industrial revolution and the demographic transition are then seen as two different aspects of a single economic event (Lucas, 2002). In these models, much of the economic acceleration goes through a decline in fertility and a rise in educational standards, i.e. the accumulation of human capital. Possible mechanisms for generating a decline in fertility and a gradual switch to growth are technical progress (e.g., Greenwood and Seshadri, 2002; Galor, 2011), a variation in food price (e.g., Strulik and Weisdorf, 2008), the introduction of a new contraceptive technology (e.g., Strulik, 2016), structural transformation associated with an increasing share of population investing in education (e.g., Doepke, 2004), institutional changes like changing marriage institution (e.g., Jones, 2001; Gould et al., 2008), female empowerment (e.g., Prettner and Strulik, 2016) growing useful knowledge (e.g., O’Rourke et al., 2013), the natural selection in favor of parents with a bias towards low fertility (e.g., Galor and Moav, 2002; Galor and Klemp, 2016), growing time cost of raising children (e.g., Hazan and Berdugo, 2002; Lagerlof, 2003; Doepke and Zilibotti, 2005; Boldrin et al., 2005), population dynamics (e.g., Becker et al., 1990; Kremer, 1993; Galor and Weil, 2000; Kgel and Prskawetz, 2001; Galor, 2005, 2011), rising adult life expectancy (e.g., Soares, 2005; Cervellati and Sunde, 2005, 2007; Vogel, 2011) and declining mortality (e.g., Kalemli-Ozcan et al., 2000; Kalemli-Ozcan, 2002, 2003; Boucekkine et al., 2011).

Recently, the literature has stressed the potential role of the European Marriage Pattern (EMP) that would have early contributed to reducing fertility below the biological maximum, thus increasing the opportunities of investment in human capital and improving women’s participation in labor force (e.g., Greif, 2006a; De Moor, 2008; De Moor and Van Zanden, 2010; Foreman-Peck, 2011; Voigtländer and Voth, 2013a; Foreman-Peck and Zhou, 2016). The economic effects of the EMP are however under large discussion (e.g., Dennison and Ogilvie, 2014, 2016; Carmichael, De Pleijt et al., 2016). This discussion is actually part of a larger debate on the impact of the family systems, which define power structures at the micro level, on development outcomes (Carmichael, Dilli and van Zanden, 2016).

Clark (2007a) also offers a new theory of the industrial revolution based on natural selection, with the richest having more surviving children, which supposedly favored the spread of “middle class” or “bourgeois” values from the elite to the rest of society.
All these theories provide plausible explanations of the industrial revolution and the subsequent rise in living standards in the Western world. But as Bar and Leukhina (2010, p. 425) note, “the relative importance of each such mechanism for the case of a particular country remains unclear”. This joins Mokyr’s (2005c, p. 1147) statement that “the exact connection between the demographic changes and the economic changes in the post-1750 period are very far from being understood”. In a discussion on the potential of the unified growth theory to explain British take-off, Clark (2014, pp. 238-241) also identifies a set of empirical challenges associated with the theories modeling the actual world of the industrial revolution in the child quantity-quality framework. For instance, the decline in aggregate fertility rates, although the latter were lower in England compared to France on the eve of the industrial revolution, did not happen before the late 19th century in England\textsuperscript{15}. Several studies even provide evidence of a positive association between fertility and economic status in pre-industrial England (e.g., Clark and Hamilton, 2006; Boberg-Fazlic et al., 2011). The evidence of higher returns to human capital, which would have acted as a signal in favor of investment in education, also seems to be lacking for the period of the industrial revolution (e.g. Clark, 2003, 2005a; Clark and Cummins, 2015; compare Klemp and Weisdorf, 2012, 2016).

4.3 Agricultural Progress

According to converging estimates, agricultural productivity would have more than doubled between 1700 and 1850 in Britain, just at the time of the industrial revolution (e.g., Deane and Cole, 1962, Crafts, 1985a; Allen, 1994; Overton, 1996a). Agricultural productivity estimates are however controversial in the literature and, in fact, considerably vary in where exactly they

\textsuperscript{15}Besides these aggregate results, Clark (2014) however mentions recent studies showing that richer families may have reduced their fertility just at the time of the onset of the industrial revolution, thus offering some hope for models based on heterogeneous agents as opposed to a single representative agent.
place the real productivity growth, reflecting a huge discussion on the true timing of the British agricultural revolution. From the late 19th century, the standard view among economic historians has located the agricultural revolution in the late 18th century and early 19th century in a movement closely linked to the industrial revolution (e.g., Toynbee, 1884; Mantoux, 1906; Ernle, 1961; Mingay, 1963; Deane, 1965; Chambers and Mingay, 1966; Beckett, 1990; Campbell and Overton, 1993; Martins, 1993; Overton, 1996a, 1996b). Yet a number of authors have criticized the standard view, some of them arguing that the major agricultural changes and much productivity growth took place earlier (e.g., Havinden, 1961; Jones, 1965; Kerridge, 1967; Allen, 1992; Clark, 1998a, 1998b, 2002a). The industrial revolution is then interpreted as consecutive to a long prior phase of agricultural expansion. Allen (1999) suggests the existence of two agricultural revolutions. The first one, essentially accomplished by the yeomen, would have actually preceded the parliamentary enclosures period, i.e. the second half of the 18th century, while the second one would have occurred in the first half of the 19th century. By contrast, the period conventionally associated with the start of the British industrial revolution would be characterized by the stagnation of British agricultural output and productivity\(^\text{16}\), a result shared by some other scholars (e.g., Turner \textit{et al}., 2001; Clark 2002a) but dismissed by recent estimates that find a rise in British agricultural output per worker throughout the 18th century (e.g., Apostolides \textit{et al}., 2009; Banerjee, 2009; Broadberry \textit{et al}., 2013). Table 2 below reports the estimates of agricultural output per worker in a range of European countries from the influential study by Allen (2000).

In parallel to the studies aimed at calculating British agricultural productivity during the preindustrial period and the industrial revolution, a number of studies have also been devoted to estimating food supply measured in calories per head of population in England (e.g., Fogel, 2004; Allen, 2005b, Floud \textit{et al}., 2011; Muldrew, 2011; Kelly and Ó Grada, 2013; Meredith and Oxley, 2014; Broadberry \textit{et al}., 2015; Ó Grada, 2015; Harris \textit{et al}., 2015). Unsurpris-

\(^{16}\)In another paper, Allen (2000) even finds a drop in agricultural output per worker over the period 1750-1800.
Table 2: Output per worker in agriculture, 1300-1800 (England in 1500 = 1.00)

<table>
<thead>
<tr>
<th></th>
<th>1300</th>
<th>1400</th>
<th>1500</th>
<th>1600</th>
<th>1700</th>
<th>1750</th>
<th>1800</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>0.80</td>
<td>0.92</td>
<td>1.00</td>
<td>0.76</td>
<td>1.15</td>
<td>1.54</td>
<td>1.43</td>
</tr>
<tr>
<td>Germany</td>
<td>0.85</td>
<td>0.74</td>
<td>0.57</td>
<td>0.54</td>
<td>0.56</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>1.02</td>
<td>0.89</td>
<td>0.76</td>
<td>0.87</td>
<td>0.80</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>0.72</td>
<td>0.89</td>
<td>0.80</td>
<td>0.83</td>
<td>0.81</td>
<td>0.70</td>
<td>0.57</td>
</tr>
<tr>
<td>France</td>
<td>0.76</td>
<td>0.83</td>
<td>0.72</td>
<td>0.74</td>
<td>0.80</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>1.02</td>
<td>0.93</td>
<td>0.78</td>
<td>0.94</td>
<td>0.93</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>0.46</td>
<td>1.39</td>
<td>1.26</td>
<td>1.20</td>
<td>1.22</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td>1.07</td>
<td>1.06</td>
<td>1.24</td>
<td>1.48</td>
<td>1.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>1.00</td>
<td>0.91</td>
<td>0.57</td>
<td>0.74</td>
<td>0.91</td>
<td>0.81</td>
<td></td>
</tr>
</tbody>
</table>

Source: Allen, 2000, table 8

In 1300-1800, the estimates are subject to heterogeneity as well. While Fogel (2004) and Broadberry et al. (2015) find per capita supplies at around 2,200 kcals in 1750 and 1800, Kelly and Ó Grada (2013) estimate higher levels, more compatible with the disappearance of the “positive check” - in the sense of the short-run response of mortality to price and wage shocks - by the late 18th century England (Kelly and Ó Grada, 2014a) and the well-documented advantage of England over France in terms of life expectancy, mean adult height, wages and agricultural productivity (Kelly and Ó Grada 2014b).

Agricultural progress has long been at the heart of the discussion related the origins of the British industrial revolution. In Capital, Marx (18967) already stressed the dramatic role of the parliamentary enclosures that supposedly favored capital accumulation in the British agriculture and contributed to forming, through a movement of expropriation, the first industrial battalions. In line with the approach giving an essential role to agriculture, Rostow (1960) even raises agricultural progress to the level of “precondition for take-off”17. The underlying logic is simple. As agriculture captured a substantial part of economic resources in all preindustrial societies, the industrial revolution could seemingly not have taken place without prior improvements in

17In a later work, Bairoch (1974) extends Rostow’s model, notably highlighting the effects of changing agriculture on the demand for industrial goods.
agriculture. The precedence of an agricultural revolution, defined as great agricultural progress, over an industrial revolution is however not empirically verified in all the countries that experienced a successful industrial revolution, with the possible exception of Britain and the Netherlands. For instance, a number of historians conclude to the absence of any agricultural revolution in France before 1840 (e.g., Morineau, 1968, 1971, 1974). Turning the argument around, Richet (1968) even contends that it was rather the cities, new poles of growth, which stimulated the French agriculture, and not the other way round. Tackling this question, Mokyr (1985, p. 21) claims that it would be unreasonable to think in terms of a necessary sequence of agriculture first, industry next. Indeed, the industrial revolution largely fed itself as it provided agriculture with non-agricultural inputs which improved food production and distribution in the 18th and 19th centuries.

As part of the discussion on the role of agricultural progress, the origins of the resources, both labor and capital, used by the growing industry have also been heatedly debated. In particular, the supposedly important effect of the enclosures on employment and capital formation in British agriculture would not be statistically confirmed (e.g., Chambers, 1953; Gullickson, 2002). Thus, according to Griffin (2010, p. 65), agricultural workforce remained roughly constant between 1700 and 1850, at around 1.5 million workers. In line with this observation, Crouzet (1967) categorically rejects the idea that agriculture provided the British industry with many workers, at least in the first phase of the industrial revolution, especially as the process of labor reallocation was hampered by a low labor mobility, both geographically and sectorally (e.g., Williamson, 1990; compare Wallis, 2014). In this respect, Verley (1997b, p. 132) notices that the rural exodus, which led many rural people to join the cities, took place only after the take-off.

18The argument of a (urban)-demand-driven agricultural development, which stresses the role of the growth of cities in agricultural progress, has been resorted to by many scholars to explain why Northwestern Europe was the world’s most productive region in agriculture until 1800 (e.g., De Vries, 1974; Rose, 1981; Wrigley, 1987, 1988; Hoffman, 1996; van Zanden, 1999; Weisbrod, 2006; Kopsidis and Wolf, 2012).

19The idea of agricultural progress as a positive externality of industrial development is largely present in recent growth literature. See e.g. Matsuyama (1992, 2008) and Grossmann (2013).
The source of industrial capital has also animated many discussions. These discussions have actually developed around the argument formulated by Lewis and Rostow in the 1950s that a take-off could not occur without a sudden upward leap of the national gross investment rate. According to Rostow (1960), capital accumulation necessarily took place in agriculture. Yet other possible sources of capital accumulation have been highlighted in the literature, and in particular colonial and foreign trade (e.g., Williams, 1944; Frank, 1978; Wallerstein, 1974, 1980; Pomeranz, 2000). Due to the lack of reliable data, it is hard to come up with any formal conclusion about the ulterior use of the capital accumulated in agriculture and trade. But there is evidence that great part of this capital turned away from industry. Regarding the capital from trade, some authors contend that the profits from trade were too modest, or at least too modestly invested in industry, to have had any direct role in the causation of the British industrial revolution (e.g., Engerman, 1972; O’Brien, 1982; Mokyr 1985, Acemoglu et al., 2005a). With respect to the capital coming from agriculture, Allen (1994) estimates that the agricultural surplus did not contribute to financing massively the industrial capital formation, thus endorsing Postan’s (1935, p. 2) view that only little of the wealth of rural England “found its way into the new industrial enterprises”. Providing a critical assessment of the various channels through which the British industry may have financed its development, Crouzet (1972, p. 172) concludes that “the capital which made possible the creation of large-scale factory industries in Britain came mainly from industry itself”.

Did agricultural progress really contribute to British industrialization? According to Allen (1999), the answer is critically dependent on the true timing of the British agricultural revolution. Indeed, as Clark (2005b, p. 50) notes: “A diffuse revolution occurring precisely at the time of the Industrial Revolution implies that the gains of the Industrial Revolution period most likely stemmed from some economy wide social or institutional change”. Convinced by the precedence of an agricultural revolution over an industrial revolution, Allen (1994, p. 116) thus argues that “one reason why the British industrial revolution could proceed in the face of a largely static agriculture was that agriculture had already revolutionised itself between 1600 and 1750”.

52
which enabled it to feed the rising population from which the new industrial labor force was drawn. Crafts (1985b) also argues that agriculture played a key role as it did release a lot of labor between 1500 and 1750 when the agricultural share of British population declined. In line with this view, De Pleijt and van Zanden (2016) attribute great part of British economic success to the occupational change that occurred in the pre-industrial period 1300-1800, an evolution confirmed by Shaw-Taylor and Wrigley's (2014) recent estimates. Whatever the exact role of agricultural progress, Zangheri (1969) documents that this factor can hardly be viewed as sufficient for the British industrial revolution to have happened. Mokyr (1985, p. 21) even questions the necessary aspect of agricultural progress, claiming that “it would be unwarranted to infer that because agricultural expansion affected industrialization, the latter could not have occurred without the former”. Providing some support to this view, a number of studies show that the role of agricultural progress needs to be reconsidered when switching from the context of a closed economy to the context of an open economy (e.g., Matsuyama, 1992, 2008). Thus, Clark (2002b) defends the idea that the growing population of industrial revolution England was fed essentially through food imports and switching agricultural output towards food, not through an agricultural revolution.

4.4 Consumer Revolution and Urbanization

A facet of the literature suggests that the technological inventions of the late 18th century Britain were driven by some changes in consumption behavior. McKendrick (1982) was the first to introduce the concept of “consumer revolution” to describe the “large and rapid increase in the consumption of consumer goods like tableware, curtains, pictures, or cutlery, a lust for objects” (Clark 2010b, p. 1) that occurred in England over the period from around 1600 to 1750. The concept of “consumer revolution” has been empirically funded, at least partly, on probate inventories, the principal source of information on material life in England at those times, which seemingly reveal a marked increase in households’ material assets over the 17th and 18th centuries (e.g., Weatherill, 1988, 1993; Shammas, 1990; Styles, 1993; Overton et al., 2004). Yet the idea of a consumer revolution has been challenged, notably by those
who fail to observe a parallel rise in real wages in Britain\textsuperscript{20}. As an attempt to provide a solution to this apparent incompatibility, de Vries (1994, 2008) has then come up with a new theory of “industrious revolution”. According to this theory, British households, both male and female, started to work much harder to increase their level of consumption, becoming more industrious or market-oriented\textsuperscript{21}. In support to the consumer revolution, an the attendant industrious revolution, a number of studies have also documented the early existence in Britain of a large urban consumer class whose living standards were already comparatively high at dawn of the industrial revolution, partly because of agricultural and commercial progress (e.g., De Vries, 1981; Allen, 2001; Maddison, 2003, 2008; Shaw-Taylor and Wrigley, 2014), a situation in line with the “Little Divergence” hypothesis defended by the “early modernists”. In this respect, Allen and Weisdorf (2011) point out that a consumer revolution was more likely to occur in urban areas.

The high levels of British urbanization and real wages or living standards\textsuperscript{22}, which were reflected in the development of a new consumer society\textsuperscript{23}, have often been stressed to account for British economic success (e.g., Bairoch, 1974; Allen, 2010; Abramson and Boix, 2014; Brunt and Garcia-Penalosa, 2015). In particular, the advantages of cities for promoting growth and technological change have been largely discussed in literature. On the demand side, the cities provided large markets that favored entrepreneurial ventures, innovation, and the development of a large-scale industry characterized by a finer labor division (e.g., Murphy et al., 1989; Kelly, 1997; Desmet and Parente, 2010, 2012, 2014; Peretto, 2015; Desmet et al., 2016). In addition, the cities would have contributed to the development of a number of

\textsuperscript{20}For instance, using two wage-series for unskilled English women workers 1260-1850, Humphries and Weisdorf (2015) contest the idea of a “girl-powered” economic breakthrough.


\textsuperscript{22}Urbanization rates are widely used as a guide to preindustrial income and economic development. For recent studies on the determinants of the rise in British and European urbanization, see e.g. Rosenthal and Wong (2011), Stasavage (2011, 2014), Bosker et al. (2013), Voigtländer and Voth (2013b), Dincecco and Onorato (2016).

\textsuperscript{23}See e.g. McKendrick et al. (1982), Mui H.C. and Mui L.H. (1988), Hersh and Voth (2009).
institutions, like the public services and the intellectual property rights (e.g., Khan and Sokoloff, 1990, 2000, 2001; Khan, 2002; Abramson and Boix, 2014), as well as to the modernization of the transport infrastructure as cities were highly dependent on supplies from the countryside. It is often believed that the continuous improvement in the transportation technology, combined with the absence of (prohibitive) internal trade barriers, precociously resulted in more integrated and unified consumption markets in Britain compared to the rest of Europe, thus favoring industrial progress (e.g., Chartres, 2003; Shime and Keller, 2007).

| Table 3: Urban growth in England (populations in 000s) |
|---------------------------------|-----------|-----------|-----------|-----------|-----------|
| 1600   | 1700   | 1750   | 1801   | 1851   | 1871   |
| London | 200    | 575    | 675    | 971    | 2,362   | 3,267  |
| Other towns with 5,000 or more inhabitants | 135    | 275    | 540    | 1,590   | 5,054   | 8,918  |
| Total urban | 335    | 850    | 1,215   | 2,561   | 7,416   | 12,185 |
| Population of England | 4,162 | 5,211 | 5,922 | 8,671 | 17,031 | 21,488 |

| % of total population |
|------------------------|-----------|-----------|-----------|-----------|-----------|
| London | 4.8 | 11.0 | 11.4 | 11.2 | 13.9 | 15.2 |
| Other towns with 5,000 or more inhabitants | 3.2 | 5.3 | 9.1 | 18.3 | 29.7 | 41.5 |
| Total urban | 8.0 | 16.3 | 20.5 | 29.5 | 43.5 | 56.7 |
| Population of England | 100 | 100 | 100 | 100 | 100 | 100 |

Source: Shaw-Taylor and Wrigley, 2014, table 10

Increasingly popular, the demand-side factors, including population growth, are however far more difficult to integrate into the industrial story. As Mokyr (1977) argues, supply and demand are not symmetrical in long-term economic change, making necessary a careful examination of whether changes in demand preceded technological change. Thus, some scholars have criticized the demand-based approach, showing that most inventions of the industrial revolution arose in already existing markets like the textiles, the metallurgy and the paper industry (e.g., Mowery and Rosenberg 1979). Production growth on those markets may consequently be interpreted as technical progress shifting the supply curves to the right along the demand curves.
The consumer revolution, viewed as the key “causal” factor for the British industrialization, has also met a number of objections. For instance, due to the chronological gap between the estimated peak of the consumer revolution (1680-1720) and the latter industrial revolution, Weatherill (1988) concludes that the two events were largely independent from each other. If variations in household consumption behavior really triggered an industrialization process in Britain, other authors wonder why the Dutch Republic, which has often been considered as the first modern economy regarding its preindustrial levels of urbanization rates and income\(^24\), did not become the cradle of the industrial revolution. As van den Heuvel and Ogilvie (2013, p. 70) note: “The Dutch Republic - where income per capita was 50% higher than in Britain by 1700 (Maddison, 2003) - is universally viewed as the first economy to experience an explosive transformation in retailing that enabled broad masses of consumers to shift from household to market consumption and production”. Recently, Clark (2010b) has even gone as far as to refute the existence of a British consumer revolution over the period 1600-1750, assimilating the phenomenon to a statistical artifact that would result from a misinterpretation of the data.

Concerning the potential role of cities in the industrial revolution, Mokyr (1995) contends that the supply-side factors commonly associated with the urbanization process - which consisted in providing environments conducive to the positive externalities, economies of agglomeration and economies of scale highlighted by the growth literature, as well as to the diffusion, through human interactions, and preservation of information and (useful) knowledge - played a more important role than the demand-side factors. As part of a more general thinking on the idea of a demand-driven industrial revolution, Mokyr (1999, p. 25) considers that supply was by far the more important historically, claiming that “an autonomous and prior shift of the industry demand curve was not an essential part of the story”.

4.5 Trade and Empire

On the eve of the industrial revolution, Britain was a nation characterized by a relatively high degree of openness. The share of foreign trade in British national income amounted to around 9% in the mid-18th century, making Britain one of the world's greatest trading countries, then rose to around 16% by 1800 (Clark, 2010a). Enjoying a privileged access to the sea and the most powerful navy in the world, Britain, which was guided by a very pronounced colonial spirit, increased the level of its exports and imports all over the 17th and 18th centuries (e.g., Cain and Hopkins, 1993; Morgan, 2000; Harley, 2004; Daudin et al., 2010; Clark, 2015; Charles and Daudin, 2015). This growth of trade was accompanied by the expansion of British ports, like Liverpool, London, Bristol or Glasgow, and contributed to the economic development of harbors hinterlands and British proto-industrial regions, especially those active in textiles.

International trade has always received attention in the quest to identify the causes of the industrial revolution in Northwestern Europe, especially in Britain and France, two countries that possessed well-endowed colonies and were early involved in international exchanges. Since the publication of Tocqueville's lectures, several generations of scholars have thus stressed the role of trade as a powerful engine of industrial growth, thus joining many trade theorists' view. For instance, Allen (2003, p. 432) states that "intercontinental trade boom was a key development that propelled Northwestern

---

25See Unger (2011) for a book, composed of sixteen essays, that explores the dramatic rise in the efficiency of British and European shipping in the three centuries before the industrial revolution, and the dramatic role of that sector in the emergence of the West as the dominant force on the oceans of the world. See van Lottum and van Zanden (2014) for a study analyzing the role of human capital - the quality of the labor force employed on ships - in long term performance of European shipping in the pre-1800 period.

26This colonial spirit was fueled among others, by military superiority. For studies associating Britain's or Europe's military superiority with economic success and later global hegemony, see e.g. Findley and O'Rourke (2007), Pomeranz (2009), O'Brien (2011a), Voigtlander and Voth (2009, 2013b, 2013c), Rosenthal and Wong (2011), Ko et al. (2014), Hoffman (2011, 2012, 2015).

27Summarizing this view, Sachs and Warner (1995, p. 3) write down: "Trade promotes growth through a myriad of channels: increased specialization, efficient resource allocation according to comparative advantage, diffusion of international knowledge, and heightened domestic competition as a result of international competition."
Table 4: Exports and imports, 1699-1856

<table>
<thead>
<tr>
<th>Years</th>
<th>Exports m.</th>
<th>Imports m.</th>
<th>Ratio to NNI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1699-1701</td>
<td>4.43</td>
<td>3.86</td>
<td>0.072</td>
</tr>
<tr>
<td>1722-4</td>
<td>5.04</td>
<td>4.04</td>
<td>0.066</td>
</tr>
<tr>
<td>1752-4</td>
<td>8.42</td>
<td>4.71</td>
<td>0.086</td>
</tr>
<tr>
<td>1772-4</td>
<td>9.85</td>
<td>6.92</td>
<td>0.088</td>
</tr>
<tr>
<td>1784-6</td>
<td>11.4</td>
<td>16.1</td>
<td>0.125</td>
</tr>
<tr>
<td>1794-6</td>
<td>20.2</td>
<td>24.8</td>
<td>0.158</td>
</tr>
<tr>
<td>1804-6</td>
<td>34.6</td>
<td>38.4</td>
<td>0.164</td>
</tr>
<tr>
<td>1814-6</td>
<td>40.3</td>
<td>45.4</td>
<td>0.146</td>
</tr>
<tr>
<td>1824-6</td>
<td>33.5</td>
<td>47.7</td>
<td>0.133</td>
</tr>
<tr>
<td>1834-6</td>
<td>43.9</td>
<td>50.4</td>
<td>0.144</td>
</tr>
<tr>
<td>1844-6</td>
<td>55.5</td>
<td>59.8</td>
<td>0.148</td>
</tr>
<tr>
<td>1854-6</td>
<td>97.3</td>
<td>109.7</td>
<td>0.219</td>
</tr>
</tbody>
</table>

Source: Clark, 2010a, tables 25 and 26

Connecting the relationship between trade and growth to the British industrial revolution, Esteban (1997, p. 900) argues that “the overseas demand in general provided the opportunity and the stimulus for technological innovation as the industry reached the limits of growth within a protected domestic market”. Influenced by the work of Acemoglu et al. (2005a), Cordoba (2007) adheres to Esteban’s (2007) view, locating the roots of the British industrial revolution into the colonization era that provided Britain with an increasing set of trading and exploitation opportunities in Europe, Asia, and the New World. Investigating the origins of the “Great Divergence” between Europe and East Asia, Pomeranz (2000) also delivers a theory that considers Britain’s privileged access to coal and the raw materials of the New World as a factor able to explain why Britain, and not China, was the first country to experience a successful industrial revolution. In the same vein, De Zwart (2016a, 2016b) provides evidence of the integration of the global commodity markets before 180028, saying that this dimension of globalization, and the associated colonialism, played part in the rise of global economic inequality. Still recently, Allen (2009b, 2010) has proposed a so-

28 See also Sharp and Weisendorf (2013).
olution to the industrial revolution puzzle by developing a theory that places foreign trade at the heart of the story. Accordingly, Britain’s participation in international trade led to higher urbanization, and to the creation of a unique wage-price structure, characterized by high wages (international standards) and low-energy costs, which induced the labor-saving technical inventions of the industrial revolution. Additionally, through the expansion of markets, trade modified positively the internal rate of return for industrial investment and innovation, thus encouraging the “technological bet” in the British industry. Highlighting the dynamic effects of trade on the British economy, Palma (2016) also attributes a key role to intercontinental trade in the transition away from Malthusian dynamics.

While it is widely accepted that technological change boosted British exports as from the second half of the 18th century\(^{29}\), the role of trade in inducing technological change is far more controversial in the literature (Krugman, 1995). Thus, Flinn (1966) stresses the domestic market as the mainspring of industrial progress in Britain. In line with this view, Davis (1979, p. 62) claims that “overseas trade did not have an important direct role in bringing about the British Industrial Revolution”. In the same vein, Thomas and McCloskey (1981, p. 102) support the idea that “the strongest effect between commerce abroad and industry at home was from industrialization to commerce, and not the reverse. Trade was the child of industry”. Emphasizing the “home grown” nature of the British industrial revolution on the supply side, Mokyr (1985, p. 23) similarly argues that “the gains accruing to the British economy from foreign trade were not necessary to achieve the gains from efficiency growth”, a view adhered to by McCloskey (2010)\(^{30}\). According

\(^{29}\)As Clark et al. (2008, p. 523) note: “The magnitude, scale and transforming power of the Industrial Revolution lay in its unification of technical advance with the military power that generated easy British access to the markets of Europe, the Americas, the Near East and the Far East”. In the same vein, Ferreira et al. (2016) show by use of a calibrated overlapping generations model that trade specialization played a crucial role in supporting the industrial revolution and the transition to modernity.

\(^{30}\)See the Chapters 18 and 19 of McCloskey’s (2010) book. Considering that the logic of trade-as-an-engine is dubious and that even the dynamic effects of trade were small in pre-industrial and industrial Britain, McCloskey (2010) rejects foreign trade as the cause of the British industrial revolution, saying that “trade is anyway too old and too widespread to explain a uniquely European - even British - event”.

59
to the theories that dismiss international trade as the trigger of the British industrial revolution, the increasing dependence of the British industry on foreign markets over time\textsuperscript{31} was an ex-post economic reality, not something that initially drove technological change.

4.6 Institutions and Policy

The role of the institutional and political environment in the industrial revolution puzzle has been the object of rising interest over the last few decades. Thus, according to some scholars, the reason why the industrial revolution started in Britain is because the country had the right institutions, notably the right political institutions\textsuperscript{32}. In line with this view, some (economic) historians have located the roots of British industrialization in the Glorious Revolution that broke out in England in 1688 and finally led to the adoption of the \textit{Bill of Rights} (1689) just ten years after the \textit{Habeas Corpus Act} (e.g., North and Weingast, 1989; Acemoglu and Robinson, 2012). As Acemoglu and Robinson (2012, p. 4) note: “The reason that Britain is far richer than Egypt is because in 1688, Britain had a revolution that transformed the politics and thus the economics of the nation”. According to the “institutionalists”, the Glorious Revolution marked the starting point of a long period of relative peace and political stability in Britain. Moreover, this political event contributed to creating a business climate favorable to entrepreneurship, notably by consolidating the parliamentary monarchy, limiting the power of the King and action of the government, and reinforcing the property rights and contracting institutions. The new regime, which was headed by the Protestants William III of Orange and Mary II of England, also adopted a number of financial institutions, some had been previously developed on the Continent, and allowed the foundation of the first national bank, namely the \textit{Governor and Company of the bank of England}, which provided the public sector with large quantities of money. A substantial part of these funds was used to fund

\textsuperscript{31}see e.g. Crouzet (1980) and Clark \textit{et al.} (2014) for empirical evidence.

\textsuperscript{32}The idea that institutions matter for economic development has reflected an increasing consensus in economic literature (e.g., North, 1990; Acemoglu \textit{et al.}, 2005b; Greif, 2006; North \textit{et al.}, 2009; Greif and Mokyr, 2016).
the development of the Royal Navy which gradually became the Master of the oceans, favoring Britain's expansion and participation in international trade (e.g., O'Brien 1998, 2005, 2011b). North and Weingast (1989) directly attribute the rise of modern public credit in the 18th century England, and the attendant financial revolution, to the establishment of a new Constitutional monarchy that strongly defended creditors' property rights against the exercise of arbitrary state power. Thus, according to O'Brien (1991), British authorities increasingly grew in favor of property and against customary rights in the 18th century. This evolution, which contrasted with the situation in the rest of Europe, in particular during the French Revolution, led to a more efficient organization of the production in Britain and spurred economic development (e.g., Besley and Ghatak, 2010; Kishtainy, 2011; Bogart, 2011; Bogart and Richardson 2011). The institutionalists also believe that the property rights, including intellectual property rights, stimulated the creativity of British inventors, leading to an increase in innovative activity. The early existence of a British patent law, enacted as soon as 1624 in the Statute of Monopolies, would then account for the early emergence of market-oriented technological inventions in Britain (e.g., North and Thomas, 1973; Dutton, 1984; MacLeod, 1988). The likely achievements of the English Glorious Revolution for promoting entrepreneurship were actually echoed by the new liberal thought which gained in influence all over the 18th century.

Although attractive, the idea of an industrial revolution rooted into the English Glorious Revolution has faced major criticism. For instance, some

---

33 As North and Thomas (1973, pp. 155-156) note: "Innovation will be encouraged by modifying the institutional environment, so that the private rate of return approaches the social rate of return [...] By 1700 England had begun to protect private property in knowledge with its patent law. The stage was now set for the industrial revolution".  
34 See e.g. Sullivan (1989, 1990) for data on patents during the British industrial revolution.  
35 See Mokyr (2010b) for a discussion on the debated role (dilemma) of entrepreneurship in promoting the British industrial revolution.  
36 Liberal ideas - of both equal liberty and dignity for ordinary folk - are at the heart of McCloskey's (2006, 2010, 2016) trilogy The bourgeois Era, which indeed attributes great part of British economic success to liberalism, which arose from theological and political revolutions in Europe and enhanced social status for business entrepreneurs in the form of a unique respect for betterment and its practitioners.
authors have pointed out the absence of discontinuity in institutions after 1688 (see Ogilvie and Carus, 2014, pp. 426-428). In line with this view, the Glorious Revolution just reaffirmed a longstanding tradition of parliamentary control over the executives (e.g., van Zanden et al., 2011\(^{37}\)). Additionally, the Glorious Revolution does not appear to have produced any acceleration of growth (e.g., Van Zanden, 2001; Broadberry et al., 2015). Several analyses of financial data also fail to detect significant improvement in the business climate after 1688 (e.g., Clark, 1996; Epstein, 2000; Stasavage, 2002; Sussman and Yafeh, 2006). The absence of such kinds of discontinuity seems to provide some support to scholars who strongly contest the belief that the British contracts and property rights were insecure before the English Glorious Revolution (e.g., Mokyr, 1999; Harris, 2004; Clark, 2007; McCloskey, 2010). Whatever the origins of the “good” British institutions, it has often been said that the latter were anyway not peculiar to Britain. For instance, other European regions, such as the Dutch Republic and Venice, also had parliamentary institutions exercising a certain degree of control over the executives (Clark, 2014). Concerning the property rights, Allen (2003, p. 429) argues that “property was secure in all leading European nations, whatever their constitution”.

Regarding more specifically the intellectual property rights, the literature has long pointed out the difficulty to assess the exact role of the patent system in the British industrial revolution, especially as the results are critically dependent on the reliability of patent statistics as indicators of innovation (e.g., MacLeod, 1988, 2009; Mokyr, 2009; MacLeod and Nuvolari, 2016). If there seems to be some growing consensus on the idea that the rapid growth of British patents as from the 1760s reflected an upsurge of technological activity in Britain, it is not clear whether the patent law, as it existed, played a great incentive role in the processes that eventually brought about sustained growth (Mokyr, 2009). With respect to this point, Ó Grada (2016, p. 228) notices that “most of British goods and processes on show at the Great Ex-

\(^{37}\)Van Zanden et al. (2011) carry out a study on the evolution of the European parliaments over the period 1188-1789. The paper evidences an institutional “Little Divergence” - between, on one side, Britain and the Netherlands, and on the other side, Southern and central Europe - between 1500 and 1800 in Europe.
hibitation of 1851 had been developed without a patent⁴⁸. The patent was only one way of protecting and rewarding an invention. In many cases, secrecy proved to be a better solution (e.g., Meisenzahl and Mokyr 2012). Financial rewards were also given by public institutions such as the Royal Society and the Parliament. In any event, all inventors were not first and foremost motivated by the desire to maximize their income. For instance, honor and social recognition were two other possible incentives. In support to this view, Clark (2014, p. 18) documents that “the industrial revolution economy was spectacularly bad at rewarding innovation”, concluding that “there is no evidence that it was some institutional changes giving better rewards for British innovators that unleashed mankind’s creative potential”. In the same vein, MacLeod (1988) shows that the rapid growth of British patents in the 18th century was more closely linked to the phenomenon of “emergent capitalism” than “inventiveness”. In a provocative book, Boldrin and Levine (2008) even go as far as to claim that “intellectual monopoly was not necessary for innovation and as a practical matter was very damaging to growth, prosperity and liberty”. Though more nuanced, yet Mokyr (2009) paradoxically suggests that the English economy may have benefited from the failures of the English patent system⁴⁸.

The other appealing idea that British public authority, operating through strengthened democratic institutions, sought to implement a general policy favorable to progress over the 18th century, and the effects of which would have culminated with the appearance of the industrial revolution, has also been challenged⁴⁹. As Mokyr (1999, p. 31) notes: “any policy objective aimed deliberately at promoting long-run growth would be hard to find in Britain before and during the Industrial Revolution era. To be sure, certain statutes aimed at encouraging progress [...] But many of these acts were directed towards increasing the economic rents of a successful political lobby and their overall impact on technological progress at best ambiguous”. Thus, the slave trade, the mercantilist regulations and the Corn Laws are all examples of

---

⁴⁸See Zukerfeld (2014) for a similar argument.
⁴⁹See Hoffman (2015) for a discussion on the largely misunderstood role of politics in economic history.
policies maintained by the English parliament to defend the property rights and profits of special-interest groups. The experience of other countries, such as Poland, the Dutch Republic and Prussia, also tends to show that the early existence of powerful parliaments, exercising some control over the executives, was hardly enough to induce growth-friendly economic policies, especially when the parliaments did not represent a diversity of views (Ogilvie and Carus, 2014). In these conditions, the likely advantage of Britain, where the Central State arguably implemented less harmful economic policies and was less subject to the action of powerful pressure groups (e.g., Olson, 1982), was less the existence of parliamentary institutions than the absence, at least before the 1800s, of an effective local paid bureaucracy able or willing to enforce the possibly harmful economic policies (e.g., Brewer, 1989; Mokyr, 1999; Ogilvie and Carus 2014).

Besides the ambiguous role of English parliament in designing growth-friendly economic policies, other arguments have been advanced to contradict the idea that British industrialization was produced or simply favored by the action of a State committed to the long-run development of the whole British economy. For instance, O’Brien (1991) points out the flaws and weaknesses of the British legal system that failed to provide efficient solutions to trade disagreements and commercial disputes. With respect to this point, Mokyr (2008) sheds light on the informal private order institutions and “cultural beliefs”, as defined by Greif (1994, 2006b), that substituted for the formal institutions to keep a relatively safe environment in which entrepreneurs and innovators could operate and collaborate freely. Another point often brandished to discredit the role of public authorities is the sharp rise in public revenues from taxation between the English Glorious Revolution and 1815.\textsuperscript{40} Over this period, real gross national income rose by three while real peacetime taxation rose by around fifteen (O’Brien 2001). In principle, the enhanced

\textsuperscript{40}As Ogilvie and Carus (2014, p. 458) note: “This huge rise in government control over national resources after 1688 casts serious doubt on the view that 1688 marked an improvement in the security of ownership rights of British taxpayers”. By contrast to most research that views the explosion of sovereign debt in the 18\textsuperscript{th} century Britain as either detrimental or neutral for economic growth, Ventura and Voth (2015) have recently argued that Britain’s borrowing boom was beneficial to the British economy.
State capacity for collecting funds after 1688 may have favored the implementation of growth-friendly policies. But the British State rarely turned into public entrepreneur, and most of the projects which are now financed collectively because of their positive externalities, like education, transport or the promotion of useful arts, were often left to the individual initiative (Mokyr 1999). Public expenditures were primarily devoted to military goals (e.g., O’Brien, 1988, 2001).

4.7 Modern Science

Addressing the European aspect of economic development, i.e. the European miracle, that led to the emergence of a “convergence club” at dawn of the 20th century, a number of scholars have highlighted the role of modern science in producing the British and European industrial revolution (e.g., Musson and Robinson, 1969; Jacob and Stewart, 2004; Bekar et al., 2005; Jacob, 1988, 1997, 2007, 2014). In line with this view, the industrial revolution would owe a lot to the European scientific revolution. Although broader in scope, the term “scientific revolution” is nowadays traditionally used to describe the changes in the scientific thought which progressively took place in Europe over the 16th and 17th centuries, leading all the scientific disciplines to reorganize around new principles and axioms (Cohen, 2010). The starting point of the scientific revolution is often located in Copernicus’ heliocentric theory that was published in 1543 in a book entitled On the Revolutions of the Heavenly Bodies. Copernicus’ model placed the Sun, and not the Earth, at the center of the Solar system and hence heretically broke with the geocentric astronomy and the canonized Aristotelian tradition which was widely taught at universities and had been elevated to the level of religious dogma through the influence of Christian scholastic philosophy.

Regardless of whether Britain enjoyed scientific leadership on the eve of modernity, there is a quite old debate in the literature on the exact role of science in the British industrial revolution. Most debate has developed around the argument formulated by Musson and Robinson (1969, p. vii) that the British industrial revolution was not unrelated to the scientific revolution and required something more than the “unlettered empiricism” suggested by
traditional historiography (Ó Grada, 2016). Though a direct connection can be established between science and some of the inventions of the industrial revolution, such as the chemical inventions and arguably the steam engine, yet scholars have found it difficult to associate the main technological breakthroughs of the British industrial revolution with the scientific discoveries of its time (Mokyr, 2011). Thus, a number of economic historians have criticized Musson and Robinson’s emphasis on science, arguing that early British technological inventions were mostly empirical and owed very little to direct scientific guidance and/or knowledge (e.g., Mathias, 1972; Hall, 1974; Allen, 2006; Ó Grada, 2016). As Mokyr (1999, p. 51) notes: “In the development stage of the basic inventions, in which the engineers and technicians on the industrial shopfloor improved, modified, and debugged the revolutionary insights of inventors such as Cort, Cartwright, and Roberts to turn them into successful business propositions, pure science played only a modest role”. Far from dismissing any contribution of modern science to industrial progress, Mokyr (1950, p. 50) adds that “if science played a role in the Industrial Revolution, it was neither through the pure foundation of technology on scientific understanding nor through the role of scientists in invention but rather through the spillovers from the scientific endeavor”. In this respect, Mokyr (2000) proposes to make the distinction among three closely interrelated phenomena: scientific method, scientific mentality, and scientific culture.

Thus, according to Jacob (1988, 1997, 2007), new science contributed to British economic success by designing the cultural and intellectual background for British industrialists. In The first Knowledge Economy, Jacob (2014) reaffirms her position, arguing that the industrial revolution resulted from the emergence of a new class of entrepreneurs within a new industrial culture. As from the 18th century, the British industry would have been increasingly penetrated by Newton’s insights and modern science (e.g., Jacob and Stewart 2004, Bekar et al. 2005, Jones 2009). In line with this view, Goldstone (2009) attributes great part of British prosperity to the diffusion of a new “engineering culture”, which would be directly derived from

---

41 See e.g. Mousnier (1958), Burstall (1963) and Daumas (1963).
Galileo’s works on mechanics (Cardwell, 1972), into the world of production. Shedding light on the culture of the “Enlightenment”, Mokyr (2005a, 2010a, 2011) also proposes a new explanation of the British industrial revolution as a byproduct of the scientific revolution, which took a particular intellectual turn in the 18th century (Clark, 2012). Mokyr (2005a, p. 291) uses the expression “Industrial Enlightenment”, defined as “the belief in both the possibility and desirability of economic progress and growth through knowledge”, to define the slice of the “Enlightenment” movement that bridged the scientific and industrial revolutions.

But, as Clark (2012, p. 89) points out: “The British industrial revolution was largely made not by the Philosophes in the Salons or the professors in the Universities, but by Craftsmen with very limited formal education solving very basic technical problems”. This raises the following question: how did the culture that forged the link between science and the modern industry disseminate into the world of production? Addressing this question, Goldstone (2009) stresses the importance of the social supports that emerged over the 17th and 18th centuries and finally made possible the combination of some new approaches to knowledge and their commercial application by private entrepreneurs. In the same vein, Mokyr (2005a, 2010a) sheds light on the 18th century Enlightenment that carried out several aspects of the Baconian program through a number of institutional developments that increased not only the amount of (useful) knowledge but also its accessibility to those who could make the very best use of it. All these institutional changes, which would root into the works of what Mokyr (2013a) calls the “cultural entrepreneurs” such as Bacon and Newton, made Britain and Europe friendlier to innovation. In reality, Mokyr (2011) identifies four headings under which the

---

42See also McLeod and Nuvolari (2009).
43In line with Mokyr’s view, Persson (2010) considers that the European economic history before the industrial revolution was much less constrained by the population growth outstripping available resources than the lack of useful knowledge, defined as the limiting factor. See Hansen and Prescott (2002) and O’Rourke et al. (2013) for formal take-off models based on the accumulation of (useful) knowledge that gives gradually access to more and more skill-biased technologies.
44In a recent book entitled Culture of Growth, Mokyr (2016) emphasizes the key role of the “Republic of Letters”, a transnational community of thinkers, in distributing writings and ideas - for which a competitive market grew up between 1500 and 1700 in a politi-
effective application of the Baconian program contributed to increasing the amount of useful knowledge: the research agenda, the capabilities, like the mathematical tool and the scientific instruments, the free selection of ideas, and their diffusion. Regarding the last point, literature has proposed various channels through which the new scientific thought may have permeated the entrepreneurial culture. They include the scientific societies like the Royal Society, the publication of scientific books, and the provincial scientific society meeting places, such as the masonic lodges and coffeehouses, where lectures on scientific topics were organized. Thus, the multiplication of the provincial scientific societies over the 18th century seemingly reveal an increasing connection between new science and industry (e.g., Thackray, 1974; Inkster, 1991; Dowey, 2014; compare Allen, 2006, 2009c; Ó Grada, 2016). Highly present in the northern regions of England and the Midlands, these societies were actually places at which industrialists, scientists, and enlightened philosophers met together and discussed. The Lunar Society of Birmingham is a famous example (e.g., Schofield, 1957, 1963). It housed prestigious figures like the savant Joseph Priestley, the physician and botanist Jonathan Stokes, the philosopher Erasmus Darwin, the industrialists Matthew Boulton and Josiah Wedgwood, and the great inventor James Watt. Member of the Lunar Society, the industrial chemist James Keir claims in Dictionary of Chemistry (1789) that “the diffusion of a general knowledge, and of a taste for science, over all classes of men, in every nation of Europe, or of European origin, seems to be the characteristic feature of the present age”.

4.8 Human Capital

The discussion on the level of scientific knowledge deployed by the British inventors has been part of a larger debate on the role of human capital in
cally fragmented Europe - that will form the “Enlightenment” and provide the supportive intellectual environment, i.e. the culture of growth, from which the industrial revolution emerged.

Mokyr (2005a, 2010a) documents a rise in the absolute and relative number of published books on scientific topics and technology over the period 1700-1800. See also Feather (1985), Baten and van Zanden (2008), Buring and van Zanden (2009) for data on book publication/consumption - that are growingly used as a measure of the long run development of human capital (Plopeanu et al., 2014) - in preindustrial and industrial England/Europe.
the British industrial revolution. If the recent literature largely stresses human capital as a key determinant of both modern economic growth and per capita income, the real contribution of human capital to the British industrial revolution is more controversial. Observing an increase in literacy and numeracy\textsuperscript{46}, book production and consumption, and educational attainment in pre-industrial Britain, a number of scholars have considered human capital accumulation as an important source of the British industrialization and take-off. This is the case of, among others, those scholars who have contributed to the “unified growth theory” and the “early modernists” who find a significant effect of human capital on economic expansion in preindustrial England and Europe (e.g., Baten and Van Zanden, 2008; De Pleijt and van Zanden, 2016; Madsen, 2016a). Yet international comparisons have shown that Britain, by comparison with other European countries, did not have any leadership in, for instance, literacy and numeracy on the eve and during the industrial revolution (e.g., Reis, 2005; Clark, 2014). Several studies even report a stagnation or decline in the levels of education in Britain during the period commonly associated with the industrial revolution (e.g., De Pleijt, 2016), thus giving credit to the “deskilling hypothesis” which supports the idea that the industrial revolution increased the demand for unskilled workers (e.g., Nicholas and Nicholas, 1992; Goldin and Katz, 1998; Mitch, 1999, 2004; Kirby, 2005; Humphries, 2010, 2013; de Pleijt and Weisdorf, 2016; compare De Pleijt et al., 2016). But, as Sanderson (2013, p. 31) says: “Those who still regard education as important could fairly point to the increase in [some indicators like] literacy in preindustrial period as establishing a threshold for industrialization, whatever dip ensued subsequently”. In any event, it is by now widely accepted that Britain did not enjoy any advantage in schooling and formal institutions which provided human capital on the eve and during the industrial revolution (Mokyr, 2013b). Anyway, thinking about the motivation of public education expenditures at that times, Galor (2011, p. 30) notes: “In the first phase of the industrial revolution, human capital had a

\textsuperscript{46}See e.g. Hippe (2012) for a discussion on the complementary relationship between numeracy and literacy, two indicators widely used to assess human capital at times of the industrial revolution.
limited role in the production process. Education was motivated by a variety of reasons such as religion, enlightenment, social control, moral conformity, sociopolitical stability, social and national cohesion or military efficiency’.

Attempting to rehabilitate human capital as a key factor likely to explain why Britain came to take the lead of the industrial revolution, a number of scholars have been highly critical of the studies that use only aggregate data on education to assess the level of human capital of the average worker in the British economy, highlighting the role that British engineers and entrepreneurs at the top of the skill or knowledge distribution may have played (e.g., Mokyr, 2005b, 2010a; Kelly et al., 2014; van der Beek, 2014; Squicciarini and Voigtlander, 2015). In line with this view, Mokyr and Voth (2010, p. 35) stress that “the British Industrial Revolution was carried not by the skills of the average (modal) worker but rather by the ingenuity and technical ability of a minority”. In this regard, the literature has increasingly devoted attention to the British education outside the schools, shedding light, among others, on the British apprenticeship system as an alternative mean of transmitting knowledge, such as the tinkering abilities and the mysteries and secrets of trade, and compensating for the numerous flaws of scientific texts or patents in disseminating technological knowledge (e.g., Epstein, 1998, 2004; Humphries, 2003, 2011; Meisenzahl and Mokyr, 2012; Minns and Wallis, 2012; Mokyr, 2010a, 2013b). Often considered as a hindrance to innovative activity (e.g., Desmet and Parente 2014, Desmet et al. 2016), the British apprenticeship system may actually have helped human capital formation, i.e. the formation of skilled engineers and technicians able to provide an empirical counterpart to the intellectual changes of the 18th century Enlightenment, especially as it was relatively open, accessible and flexible, and seemingly not as restrictive as in other European nations (e.g., Kaplan, 1993; Mitch, 2004; Crowston, 2005; Wallis 2008; Leunig et al., 2011; Minns and Wallis, 2013). Thus, Ó Grada (2016, pp. 227-8) observes that “most of the foremost inventor-entrepreneurs of the British Industrial Revolution were of rather modest, artisanal origins [...] The artisans-made-good were actually the most talented and ambitious products of a system that combined basic schooling in literacy and arithmetic with apprenticeships based mainly
on learning-by-doing”. Arguing that Britain’s technological lead on the eve
and during the industrial revolution was based upon “the supply of highly
skilled, mechanically-able craftsmen”, Meisenzahl and Mokyr (2012, p. 452)
also stress the importance of the apprenticeship system as it was the domi-
nant form of skill formation at that time. Based on the statistical exploita-
tion of a comprehensive dataset for Europe from around 1200 to 1900, Abramson
and Boix (2014, p. 4) similarly emphasize the key role of the British appren-
ticeship system as an information exchange channel, showing that “economic
growth was only possible when there was a population of craftsmen who em-
body a given stock of technological know-how/knowledge that enabled them
to take fully advantage of the technological breakthroughs of the 18th century”.
The basic idea behind this result is that the presence of high-skilled workers
facilitates both technology adoption and adaptation\textsuperscript{47}, and so the production
of inventions, especially of microinventions\textsuperscript{48}.

Besides the relatively important presence of highly skilled - in terms of
ability and dexterity - mechanics and engineers in Britain, and more partic-
ularly in the technologically dynamic places (de Pleijt \textit{et al.}, 2016), Kelly \textit{et al.}
(2014, 2015) highlight the physical condition of the average British worker
as a key determinant of the quality of British workforce. Accordingly, better
nutrition made British workers healthier and taller, which can be interpreted
as a higher level of health human capital (Schultz, 2002; Madsen, 2016b) that
enhanced their cognitive ability and productivity, with potentially important
implications in terms of economic development (e.g., Jones and Schneider,

5 Conclusion

The British industrial revolution is still home to great mysteries. In partic-
ular, the causes of the event are still heatedly debated in the literature. As
\textsuperscript{47}See also Becker \textit{et al.} (2011) for a study showing that human capital facilitated the
adoption of the new British industrial technologies in a number of European countries over
the 18\textsuperscript{th} and 19\textsuperscript{th} centuries.
\textsuperscript{48}In addition to the presence of highly-skilled workers, some scholars like Kuhn (1977)
and Jacob (1988, 1997) argue that the production of microinventions was favored by a more
pragmatic and applied science in Britain.
Clark (2012, p. 1) notes: “The British Industrial Revolution is the key break in world history, the event that defines our lives. No episode is more important. Yet the timing, the location, and the cause of this Industrial Revolution are unsolved puzzles”. This survey provides an overview of the literature devoted to identifying the main causes of the British industrial revolution. This literature is huge, fast-growing, and has moreover considerably evolved over the last twenty years, notably with the development of the “idealist approach” adopted by very influential scholars like Mokyr (2010, 2016) and McCloskey (2006, 2010, 2016), which attributes a crucial role to the arrival of a particular culture or ideology in the industrial revolution. Thus, while Mokyr views the industrial revolution as a byproduct of the European scientific revolution and the attendant “Industrial Enlightenment”, with the application of enhanced rationality to economic interests, McCloskey views the industrial revolution as the outcome of liberal ideas and enhanced social status for business entrepreneurs in the form of a unique respect for betterment and its practitioners. Though highly criticized by the proponents of the cliometric approach because they are hardly testable, the idealist theories appear promising in terms of defining the features of the new intellectual environment that was seemingly conducive to the emergence of continuous technological progress in Britain and, more generally, Europe. They however have a lot of difficulty explaining why Britain came to take the lead of the industrial revolution. The “Enlightenment” was indeed a movement across large parts of Europe. Similarly, the “Bourgeois Revaluation” first took place in Holland. In the absence of an idealist revolution that would be peculiar to Britain, the “incentives approach” then appears to offer a valuable avenue for understanding the reasons why Britain was the first country to undergo a successful industrial revolution, the latter relying on a new intellectual substratum.

In line with this view, the slow modernization process observed by the “early modernists” in Britain during the preindustrial era, which was reflected in increases in e.g. real incomes and urbanization, may potentially account for the early existence of relatively mature British markets that provided British entrepreneurs with material incentives to develop new technological inventions. Accounting for the “Little Divergence” would then become essential
for identifying the factors that led Britain to play an important role in the processes that brought about modernity. In a synthetic study, Broadberry (2014) summarizes the main potential forces behind Europe’s “Little Divergence”. They are closely related to the “causal” factors described throughout our survey essay. Thus, Broadberry (2014) stresses the role of, among others, long-distance trade, especially through its interaction with the institutions, pastoral farming and, more generally, agricultural progress, consumer preferences, attitudes to work, and demographic factors such as the European Marriage Pattern (EMP) which would have contributed to decreasing fertility below the biological maximum, thus increasing the opportunities of investment in human capital and improving women’s participation in the labor force. Testing various hypotheses explaining the process of differential growth in early modern Europe, De Pleijt and van Zanden (2016) find that institutional changes, in particular the rise of active Parliaments, and human capital formation are the main drivers of the economic growth that took place in the North Sea Area of Europe between 1300 and 1800. Moreover, the study suggests that religion, i.e. the spread of Protestantism, played a key role as it affected human capital formation. Broadly speaking, the role of institutions and human capital has gained a renewed interest over the recent years in the literature dealing with the two most successful European economies in the early modern period, namely Britain and the Netherlands. Both countries rejected political absolutism in the 17th century, with an increasing control of mercantile interests over the state through parliament, and experienced a growth of the traditional indicators measuring human capital. With respect to human capital, Britain is moreover believed to have benefited from the relatively important presence of highly-skilled mechanics and engineers. British technical creativity, fueled by a large stock of technical know-how, would then potentially explain why Britain could early take fully advantage of the new market opportunities offered by the technological breakthroughs of the 18th century. The British industrial revolution could then be interpreted as the

49 These results corroborate, at least to some extent, the findings from the study by Baten and van Zanden (2008), which also suggest that human capital, measured with data on book production, is a significant determinant of economic performance in the centuries before the industrial revolution in the North Sea Area of Europe.
technological reply, by a population of highly skilled and mechanically able craftsmen, to the expanding set of economic opportunities that emerged in a new institutional and intellectual environment more favorable to material progress.

An interplay of various factors best explains the early industrialization of Britain, and more generally the gradual rise of Europe towards modernity. While the multiplicity of theories is hopeful as it would be very dangerous to overestimate the explanatory power of any single factor, it has led Glaeser (2010) to conclude that the industrial revolution would never be fully understood: "While the reader craves a simple explanation, there is none to be had. The entire question of why the industrial revolution started in England will never be definitively answered. The event was sui generis a bolt of lightning; and there is a myriad of possible explanations for it". Every new plausible theory adds complexity to the industrial revolution story, but also does better justice to the historical reality, deepening our understanding of that major event, which improves as we gather more reliable data and information on the empirics of the industrial revolution and early modern period.
Chapter 2

The Drivers of Structural Change

Leif van Neuss

Abstract

The main goal of this paper is to provide an integrated survey of the literature devoted to identifying the driving forces behind the process of structural change, defined as the reallocation of economic activity across sectors. First presenting the main empirical facts associated with the familiar sectoral trichotomy agriculture, manufacturing and services in different regions of the world - Europe and the USA, Asia, Latin America and Africa - by use of the GGDC 10-Sector Database, the paper then reviews four theoretical determinants of structural change: (i) changes in income, (ii) changes in relative (sectoral) prices, (iii) changes in input-output (sectoral) linkages and (iv) changes in comparative advantages via trade.

Keywords: Growth, Structural change, Industrialization, Deindustrialization

JEL classification: E20, O40
1 Introduction

Following seminal contributions by scholars like Fisher (1939), Clark (1940) and Fourastié (1949) in the first half of the 20th century, structural transformation or structural change, often defined as the reallocation of economic activity across the broad sectors agriculture, manufacturing and services, has been the object of growing interest in the literature over the last few decades. In his Nobel Prize lecture, Kuznets (1973) refers to structural change, a phenomenon characterized, among others, by the massive shift of labor away from agriculture, as one of the six main features of modern economic growth. By way of illustration, the agricultural share of employment declined from 75% to less than 3% between 1800 and 2000 in the United States, while the agricultural share of output fell from 40% to 1% between 1840 and 2000 (Dennis and Iscan, 2009). The surge of interest in structural change has relied, at least partly, on the numerous concerns and issues raised by deindustrialization that has affected the world’s most economically successful countries since the last third of the 20th century, but also a number of low and middle income countries, especially in Africa and Latin America, since the 1980s. These concerns and issues have indeed fed a large debate on both the causes and consequences of structural change and on the role of policy instruments in driving the sectoral allocation of activity.

What are the main driving forces behind structural change? As noted by Herrendorf et al. (2013), the answer to this question has important implications. In particular, understanding the role played by each of these forces is essential to ensure the appropriate policy response. In view of the relative decline of manufacturing in a large range of countries, there is for instance a large discussion on how public policies should react. This depends crucially on the drivers of this decline. If deindustrialization is considered as a “natural” phenomenon arising from an efficient equilibrium outcome, there is no room for public intervention. If deindustrialization is rather viewed as taking place out of an efficient equilibrium outcome or resulting from a large and persistent structural disequilibrium in the macroeconomy, then appropriate public policies can potentially lead to improvement by modifying the pace of
deindustrialization and/or accompanying the process through mitigating its associated costs. These costs could, for example, result from imperfect resource (labor and capital) mobility across sectors of activity, and materialize into higher unemployment and inequality. The drivers of structural change also have important implications regarding growth perspectives. For instance, if structural change takes place as part of Baumol’s (1967) “cost disease”, then the economic activity is being continually reallocated from the technologically dynamic sectors, i.e. those with a relatively high rate of productivity growth, towards the sectors with low productivity growth. In line with this theory, the alarming hypothesis of deindustrialization as the principal factor responsible for inferior economic performance in the Western countries, which some economists believe to be the onset of a “secular stagnation”, has gained in popularity over the recent years as the manufacturing share of total workforce is reaching lower and lower levels.

The principal goal of this paper is to provide an integrated overview of the fast-growing literature devoted to identifying the forces behind structural change. Multi-sector growth theories have primarily focused on two mechanisms through which the process of structural change can take place in market economies: changes in aggregate (real) income and changes in relative sectoral prices. Income effects result from non-homothetic preferences. This family of preferences is associated with cross-sector differences in the elasticity of income. Relative sectoral price effects result from sectoral differences in technology. In addition to presenting these two mechanisms, notably discussing the recent papers applying them simultaneously into the same analytical framework, in particular the papers integrating more flexible production structures and demand systems than the prior literature, this survey essay highlights two additional mechanisms through which structural change can take place: changes in input-output or sectoral linkages and changes in comparative advantage via international trade. While these two channels have been largely overlooked in the multi-sector growth literature dealing with structural change, this survey shows that the recent contributions taking explicitly into account the role of input-output (sectoral) linkages and trade are promising in terms of accounting for structural change.
Our review can therefore be considered as a useful complement to the study by Herrendorf et al. (2014). In their work, the authors mainly investigate the income and relative (sectoral) price effects, discussing extensively the influential studies by Kongsamut et al. (2001) and Ngai and Pissarides (2007). The trade channel is solely summarily addressed as one possible extension of the benchmark model. With respect to sectoral linkages, the authors do not include intermediates into their analysis and, consequently, can hardly explain the evolution of the input-output structure of the economy in both developed and developing countries - which has in fact largely reflected the large increase in the service intensity of production processes, not the least within manufacturing - and the substantial expansion of services for which final demand plays a relatively small role, like Professional and Business Services (PBS) or financial services. Our review has the hope of filling these gaps, leaving the readers with a comprehensive view of the recent literature on structural change. The rest of the paper is organized as follows. Section 2 discusses certain issues related to the measure of structural change, then graphically illustrates the empirical facts of structural change for different regions of the world, namely Europe and the USA, Asia, Latin America and Africa. Section 3 reviews the literature on the drivers of structural change. Section 4 concludes.

2 The Empirical Facts of Structural Change

2.1 How to Decompose the Aggregate Economy into Sectors?

In every economy, a wide range of goods and services are produced for consumption, intermediate and final consumption, and investment. In macroeconomics, it is usual to distinguish between these goods and services according to some specific attribute(s). In line with this approach, the products sharing similar attribute(s) are aggregated into a common group or sector, which is the object under study in the literature on structural change (Stijepic, 2011). The best known and most basic sector-division of the aggregate economy is the familiar sectoral trichotomy of agriculture (primary sector), manufactur-
ing (secondary sector), and services (tertiary sector). However, there is a myriad of ways to divide the economy into sectors of activity. Among others, the sectors can be defined according to technology (e.g., Caselli and Coleman, 2001; Peneder, 2007; Ngai and Pissarides, 2007; Acemoglu and Guerrieri, 2008), demand characteristics (e.g., Kongsamut et al., 2001; Foellmi and Zweimüller, 2008), the nature of output, the nature of intermediates, or the specificities of innovation (e.g., Pavitt, 1984; Castellacci, 2008; Castaldi, 2009). While some part of the theoretical literature on structural change explicitly seeks to explain the evolution of agriculture, manufacturing, and services over the course of economic development, so identifying some attribute(s) that are seemingly peculiar to these sectors, another part of the literature defines the sectors without referring to this sectoral trichotomy.

2.2 How to Measure Structural Change?

Structural change is typically investigated by looking at the evolution of the economic shares of the identified sectors of the economy. Such exercise requires, first and foremost, splitting the aggregate economy into relevant sectors, and then measuring the economic activity at sectoral level. The most common measures of activity at sectoral level are employment and value added, two production-side measures, and consumption. Regarding employment, two widely-used indicators are “total workers” and “total hours worked”. If the choice of one indicator over the other is very often dictated by data availability, it is important to point out that these two indicators may deliver different results, partly because the indicator “total workers” is a less precise measure of employment. Value added is often expressed in nominal terms, i.e. in current prices, but can also be given in real terms, i.e. in constant prices.

1While the structural change literature widely uses the term “manufacturing”, instead of “industry”, to refer to the secondary sector, i.e. to all activity that falls outside of agriculture (primary sector) and services (tertiary sector), it is noteworthy that manufacturing is just one - the largest one - component of industry in most statistical classifications of economic activity. Industry then includes the mining and quarrying activities, manufacturing, and the construction sector. Except for the section dedicated to presenting the empirical facts associated with structural change, we follow most of the prior literature and use the term “manufacturing” to refer to the secondary sector.

2If the introduction of the chained indices has considerably improved the accuracy of the measure of economic growth, the use of the value added series in chained prices does
Differences between the employment and value added shares reflect sectoral
differences in apparent labor productivity. While the literature on structural
change has long considered the different measures of structural change as eas-
ily interchangeable, Herrendorf et al. (2014) have recently cautioned against
this too simple view, arguing that the different measures may display different
behaviors and lead to differing conclusions with respect to the driving forces
behind structural change. For instance, based on the fact that the concept of
“value added” is different from that of “final output”, Herrendorf et al. (2013)
show that the choice of the commodity space, i.e. whether it is in terms
of final consumption expenditure or consumption value added, can have an
impact on the causes of structural change.

2.3 The Empirical Facts of Structural Change for Different
Regions of the World

We now examine the empirical facts associated with structural change by
use of the GGDC 10-Sector Database (Timmer et al., 2015). This database
provides internationally comparable data on economic activity at the sec-
tor level, in particular on employment and nominal value added, for a set
of countries that are categorized into six geographic regions (Sub-Saharan
Africa, North Africa, Asia, Latin America, North America and Europe) over
the period from 1947 onwards. This database offers the opportunity to com-
not appear suitable for the analysis of structural change, as these statistical series, by
construction, do not feature sectoral additivity, meaning that the total sum of the value
added of each sector is not equal to the value added of the aggregate economy, with the
exceptions of the base year and the year following the base year.

3In order to understand the distinction between value added and final gross output and,

hence, to understand why the production and consumption measures of structural change
may display different behaviors, let's take an example. In particular, let's consider a man-
ufactured good put up for sale in a retail establishment. In fact, the price paid by the final
consumer covers the compensation of all the production factors involved in the production
and distribution process of this good. From the production (value-added) perspective of
the national accounts, the final price should then be split into different sectors, including
manufacturing and services, and possibly agriculture in case the production process of the
good requires agricultural inputs. Even in the extreme scenario of an economy in which
households would consume only manufactured goods, all the sectors other than manufac-
turing, and in particular services, could account for a substantial part of total value added.

This shows well the importance of taking into account the input-output (sectoral) linkages
when attempting to match the production and consumption measures of structural change.
pare the patterns of structural change in a large range of countries, including developed and developing countries. By contrast to the Historical National Accounts Database (Smits et al., 2009), which gives data allowing the study of structural change in a few countries over the very long period as from the 19th century, the GGDC 10-Sector Database (Smits et al., 2009) also provides researchers with the possibility to isolate the manufacturing sector, viewed as the largest component of industry in most of official statistical classifications of economic activity, from the rest of industry. When it is possible, making the distinction between manufacturing and industry is relevant as manufacturing may behave differently from industry, especially in developing countries which are well-endowed with natural resources and highly active in mining and quarrying activities.

Figure 1 shows the evolution of the shares of agriculture, industry, manufacturing, and services in employment and nominal value added for the USA and a set of European countries over the course of economic development. Data are described in detail in Appendix A. The level of economic development is measured as the log of GDP per capita expressed in 1990 international dollar. Data on GDP per capita is retrieved from the “Maddison-Project”. They are presented and discussed in Bolt and van Zanden (2014). Though sectoral data on employment and nominal value added for the USA and the selected European countries is, at best, available from 1947 onwards, figure 1 spotlights what the research, based on historical time series, views as the stylized facts of structural change in currently rich countries. Up to the 18th century, agriculture was the most important sector of the economy, absorbing most of labour force and contributing most of production. Its importance was emphasized by influential economists like Frans Quesnay who considered agriculture as the only productive sector able to create profits, i.e. a produit net, relegating the other activities like manufacturing and trade, viewed as “sterile”, to the sidelines (Kongsamut et al., 2001). As from the second half of the 18th century, a number of nations experienced a successful industrialization. According to traditional historiography, the industrial revolution started in Britain, then spread to Continental Europe and the USA, before reaching Japan and Russia by the end of the 19th century. If the causes of
the industrial revolution are heatedly debated in the literature, it is widely accepted that the industrial revolution, which witnessed an unprecedented increase in the rates of efficiency advance, marked the transition period towards modernity and sustained economic growth, and accelerated the process of structural change. Over the last two centuries, the expansion of GDP per capita has thus been accompanied by a decline in the share of agriculture in both employment and nominal value added. By contrast, the share of services has risen. Industry, largely influenced by manufacturing, has moved on a quite different trajectory as its share follows a “hump shape”. It is rising for lower levels of economic development (industrialization) and decreasing for higher levels of economic development (deindustrialization). Due to the later economic development of European countries like Italy and Spain, this hump-shaped trajectory is clearly visible on the employment part of figure 1. As data on nominal value added at the sectoral level is missing for Italy and Spain between 1950 and 1970, the hump-shaped trajectory contrastingly does not clearly appear on the value added part of figure 1. The data, taken as a whole, is however consistent with a hump shape.

Now switching to the Asian countries, the plots in figure 2 seemingly confirm the structural patterns observed in currently rich countries. This is true even when removing Japan and Korea, two countries that traditionally belong to the list of the currently rich countries. In particular, the share of agriculture in employment and nominal value added has fallen with economic development in all selected Asian countries, while the share of services has risen. The data, taken as a whole, is consistent with a hump shape for the employment and nominal value added shares of industry and manufacturing.

The picture for Latin American countries, depicted in figure 3, looks somewhat harder to interpret. Regarding the employment shares, several facts are confirmed. In particular, the agricultural share has declined with the level of economic development in all selected countries, while the share of services has increased. The story for industry and manufacturing is more nuanced. If the process of development of Bolivia, the poorest of the selected Latin American countries, has been characterized by a rise in the employment share of manufacturing over the last few decades, Brazil and Colombia have experienced
increases in GDP per capita along with relative manufacturing workforce remaining roughly stable. In these two countries, the expansion of the industrial share of total workforce has been primarily driven by the construction sector.

One striking fact related to Latin America is that the region as a whole has gone through a “premature deindustrialization” - defined by Tregenna (2015, p. 2) as “deindustrialization that begins at a lower level of GDP per capita and/or at a lower level of manufacturing as a share of total employment and GDP, than is typically the case internationally” - since the 1980s. This average trend has been well embodied by individual countries like Argentina, Chile, Costa Rica or Venezuela. In line with the “premature deindustrialization” thesis, the value added share of manufacturing has declined in virtually all Latin American countries since the 1980s. The growth of the industrial share of nominal value added in countries like Chile, Colombia and Venezuela over recent decades, which translated into a decline in the share of services in the whole economy, has actually resulted from the evolution of the mining and quarrying activities that account for a substantial part of total production in these countries.

Africa is by far the poorest region in the world. As shown in figure 4, agriculture absorbed more than 70% of total employment in 1960 in countries like Botswana, Ethiopia, Kenya, Malawi, Nigeria, Senegal and Tanzania. Although the employment share of agriculture has fallen in all selected African countries over the last few decades, agriculture still accounted for great part of total employment in 2010, thus providing support to scholars who argue that the poor economic situation of developing countries, and especially of those located in Africa, is largely attributable to the large size and/or the low performance of the agricultural sector (e.g., Caselli, 2005; Gollin et al., 2002, 2007; Restuccia et al., 2008; Gollin, Lagakos et al., 2014). In parallel to the decline in the agricultural share of workforce, African countries have experienced an expansion of relative employment in services. With respect to manufacturing, the share of this sector in total workforce is still relatively small in the poorest African countries, i.e. in Botswana, Ethiopia, Nigeria, Tanzania and Zambia. In most other countries, relative manufacturing workforce has been on a declining trend since the 1980s, revealing that Africa has also been
affected by some “premature deindustrialization”. Looking more specifically at figure 4, a country, namely Mauritius, stands out among African countries with development patterns which are broadly similar to those of the currently rich countries. In Mauritius, the most successful economy in Africa, relative manufacturing employment grew rapidly during the 1970s and 1980s, peaked in the 1990s, then started falling in favor of services. Switching to the nominal value added shares, one can also observe the relative decline of agriculture in all selected African countries over the last few decades. Despite this trend, agriculture still absorbed a substantial part of nominal value added in a number of African nations in 2010. For instance, agriculture accounted for around 50% of total nominal value added in Ethiopia. The share of services in total nominal value added has clearly grown along with economic development in countries like Mauritius and Morocco. In other countries, the trend is less obvious. In Zambia, the observed decline in GDP per capita between 1960 and 2010 was even associated with an increase in the value added share of manufacturing and services. The country actually experienced a large decline of its industry, and especially of its mining and quarrying sector. While these activities accounted for almost 60% of total nominal value added in 1960, this figure declined to around 14% in 2010.

It appears difficult to come up with any formal conclusion about structural change in developing countries. While the selected Asian countries and Mauritius seem to follow a development path that looks very similar to the one carved out by the currently rich countries, a number of developing countries in both Latin America and Africa seem to have precociously run out of industrialization opportunities through experiencing a process of premature deindustrialization as from the 1980s. As noted by Rodrik (2016), premature deindustrialization can potentially have significant economic - and political - implications, including lower economic growth, especially as the drop in

\footnote{For a detailed analysis of the structural change among low and lo-middle-income countries in sub-Saharan Africa, see e.g. De Vries \textit{et al.} (2015) and Fox \textit{et al.} (2017).}

relative manufacturing employment in developing countries has been accompanied by the expansion of low productivity and non-tradeable services, as well as by the expansion of the informal sector.

Focusing on the large heterogeneity in economic performance across developing nations, Szirmai (2012) argues that the growth disparities reflect cross-country differences in the trajectory of manufacturing, a sector that would accordingly act as the main engine of economic development. In line with this view, the performance of manufacturing would explain why some developing countries have met with economic success, especially in Asia, while other developing countries have remained stagnant or only characterized by very low rates of economic progress. As Szirmai (2012, p. 417) notes: “There are no important examples of success in economic development in developing countries since 1950, which have not been driven by industrialization. All the Asian success stories are indeed stories of industrialization. Neither tourism, nor primary exports, nor services have played a similar role, with the possible exception of software services in India since 2000”.

Will the currently poorest countries follow a development path similar to the one carved out by the currently rich countries? Will the future economic development of these countries be based upon the large expansion of the manufacturing sector, whose effects on economic growth and importance for catching-up is largely emphasized in the economic literature? If so, what will be the industrialization strategy of these countries, some of them being already characterized by a relatively high degree of tertiarisation? As noted by Pollard (1990), the industrialization strategy of a specific country is indeed not independent from its initial conditions and its moment of entry into the race. Will the low and middle-income countries experiencing a “premature deindustrialization” be able to reverse the trend in the near future? If not, can the services emerge as the main engine of economic development? All these fascinating questions remain largely open and deserve further research.
Figure 1: Sectoral shares of employment and nominal value added - selected European countries and the USA from GGDC 10-Sector, 1947-2011

Data source: GGDC 10-Sector Database [Timmer et al., 2015]
Figure 2: Sectoral shares of employment and nominal value added - selected Asian countries from GGDC 10-Sector, 1952-2012

**Data source:** GGDC 10-Sector Database (Timmer *et al.*, 2015)
Figure 3: Sectoral shares of employment and nominal value added - selected Latin American countries from GGDC 10-Sector, 1950-2012.

Data source: GGDC 10-Sector Database (Timmer et al., 2015)
Figure 4: Sectoral shares of employment and nominal value added - selected African countries from GGDC 10-Sector, 1960-2010.

Data source: GGDC 10-Sector Database [Timmer et al., 2015]
3 The Drivers of Structural Change

As presented in the previous section, empirical evidence largely suggests that structural change takes place along with economic development. It is therefore quite naturally that the literature, especially the growth literature, has increasingly sought to identify, within a rigorous analytical framework, the forces that drive the process of structural change, taking over from the pioneering works which were mainly descriptive. The main goal of this section is to provide an integrated overview of this rapidly growing literature. In particular, this section reviews four determinants of structural change: (1) changes in income, (2) changes in relative (sectoral) prices, (3) changes in input-output linkages and (4) changes in comparative advantage via globalization and international trade. In order to keep things as clear as possible, we primarily focus on the employment-based measure of structural change, which is by far the most investigated in the literature (Nordhaus, 2008). While the different theories will principally be assessed in the light of the experience of advanced countries, a specific section will be dedicated to developing countries when exploring the contribution of globalization and trade, as a few contributions have recently pointed out the important role played by globalization and trade in driving the main patterns of structural change in the developing countries since they liberalized.

3.1 Income Effects, Relative Price Effects, and Structural Change

The recent neoclassical multi-sector growth literature has primarily focused on two mechanisms through which structural change can take place in market economies: changes in real income and changes in relative (sectoral) prices. Income effects on structural change result from the introduction of non-homothetic preferences into the multi-sector models of growth. Relative (sectoral) price effects result from the introduction of heterogeneous sectoral production functions. While the various channels of structural change have been initially explored individually, recent work has devoted increasing efforts to incorporate both non-homothetic preferences and sectoral differences
in technology into the same analytical growth framework. A part of this work has been directed towards quantifying the relative importance of the income effects and relative price effects in accounting for the patterns of structural change observed in the data.

3.1.1 Income effects and Structural Change

A first facet of the literature on structural change shows that the process of reallocation of economic activity across sectors can be due to changes in - the structure of - demand resulting from changes in real income (e.g., Pasinetti, 1981; Falkinger, 1994; Echevarria, 1997; Laitner, 2000; Zweimüller, 2000, Caselli and Coleman II, 2001; Kongsamut et al., 2001; Gollin et al., 2002 and 2007; Greenwood and Seshadri, 2002; Meckl, 2002; Steger, 2006; Foellmi and Zweimüller, 2008, Bonatti and Felice, 2008; Duarte and Restuccia, 2010; Boppart, 2014). Structural change is thence driven by non-homothetic preferences, such as the Stone-Geary preferences, that generate non-linear Engel curves. As income increases, the marginal rate of substitution between the different goods varies, thus inducing some reallocation of activity towards the sectors providing goods that meet relatively higher hierarchical needs. In line with this theoretical approach, the three broad sectors agriculture, manufacturing and services are supposed to satisfy the most urgent needs, the less urgent needs and the most luxurious needs respectively (Foellmi and Zweimüller, 2008). If scholars like Fisher (1935) and Clark (1940) were pioneers in introducing this approach, Kongsamut et al. (2001) were the first to address explicitly the issue of how to reconcile the Kuznets facts on structural change - driven by income effects - with the Kaldor facts on aggregate dynamics, or

6Broadly speaking, the Kuznets facts state that massive structural change (labor reallocation across sectors) takes place along with economic development. In particular, structural change during the growth process is characterized by the massive shift of labor out of agriculture into manufacturing and services in the early stages of economic development, and away from manufacturing towards services in later stages of economic development.

7The Kaldor facts are a set of six statistical properties of long-term economic growth proposed by Kaldor (1957) in an influential study. Conducting empirical investigations with national income accounts, Kaldor (1957) evidences that some key aggregate measures of the economy behave in a quite stable manner when they are considered in a long-run perspective. In particular, the capital-output ratio, the rate of return on investment and the shares of labor and capital in national income are roughly constant over long periods.
less restrictively some aspects of the Kaldor facts (Kurose, 2015).8

Theoretically identified as one potential driver of structural change, the existence of differences in income elasticity across sectors has been empirically investigated in the literature. For instance, based on a data set comprising 94 countries with a wide range of income levels, 56 broad sectors including manufacturing and services, and 5 factors of production including the disaggregation of skilled and unskilled labor, Caron et al. (2014) show that the income elasticity varies considerably across goods from different sectors, with the production of income-elastic goods being (on average) more skilled-labor intensive.9 In line with this empirical evidence of a correlation between income elasticity and skill intensity, Buera et al. (2015) document for a broad panel of advanced countries that rises in GDP per capita are associated with a systematic shift in the sectoral composition of the economy towards sectors that are intensive in high-skill labor, an evolution labelled as “skill-biased structural change”, which potentially has significant implications for the level of the “skill premium”, defined as the wage differential between skilled labor and unskilled labor.

Dealing more specifically with the sectoral trichotomy agriculture, manufacturing and services, Houthakker (1987, p. 143) notes that “of all empirical regularities observed in economic data, Engel’s law - which states that the budget share for food products declines as per capita income grows - is probably the best established”. Extending the Engel’s law to manufacturing, Clark (1940) uses cross-national statistics to show that the share of income spent on manufactured goods increases during the early stages of economic development, then stabilizes, and eventually declines beyond a certain threshold of time, whereas output per capita and capital per capita grow at a fairly constant rate.

8As possible extensions of the existing models of structural change based on income effects, it is noteworthy that demand patterns may also vary both across countries and over time depending on many factors other than the level of GDP per capita, such as changing preferences (e.g., Phelan and Trejos, 2000; Addessi and Busato, 2011; Addessi, 2014; Addessi et al., 2017), the age structure of the population (e.g., Meijdam et al., 2005; Stijepic, 2011), the income distribution (e.g., Hein and Vogel, 2008, 2009; Boppart, 2014; Hartwig, 2013, 2015), the level of public intervention (e.g., Stijepic and Wagner, 2015; Felice, 2016) and the total amount of time devoted to leisure (e.g., Cruz and Raurich, 2016).

9See also Caron et al. (2017)
income per capita, an evolution sometimes labelled as the “Bell’s (1976) Law”. While it is widely accepted among scholars that agricultural goods exhibit a relatively low elasticity of income (e.g., Clark et al., 1995), the evidence for Bell’s law, which implicitly assumes an income elasticity lesser (higher) than unity for manufactures (services) in the later stages of economic development, is however more controversial in the empirical literature (e.g., Summers, 1985; Baumol et al., 1985; Falvey and Gemmell, 1996; Baumol, 2001; compare Appelbaum and Schettkat, 1999; Schettkat, 2004; Schettkat and Yocarini, 2006). If the nominal share of total income devoted to manufactures has been declining over recent decades in advanced countries, this trend is not necessarily confirmed for the real share of income, revealing the importance of taking into account the role of (relative) prices in the study of structural change. As Rowthorn and Ramaswamy (1999, p. 20) note: “a purely demand-based account of deindustrialization is incomplete because it neglects the influence of productivity and prices on the structure of demand, and so on output and employment”.

In a recent study, taking explicitly into account both income and relative price effects, Comin et al. (2015) estimate a full demand system derived from non-homothetic CES preferences, using historical data on sectoral shares from 25 different countries and household survey statistics for the postwar period. They find that the difference in the elasticities of income between agriculture and manufacturing is negative, while the difference between services and manufacturing is positive. They moreover find that these differences are remarkably stable over time, a result which contrasts with some previous empirical studies and does not give support to the evolution of sectoral income elasticities observed in the models of e.g. Kongsamut et al. (2001) and Foellmi and Zweimüller (2008).

3.1.2 Relative Price Effects and Structural Change

A second facet of the literature on structural change associates the process of reallocation of activity across sectors to changes in relative sectoral prices, the latter being driven by sectoral differences in technology. As noted by Stijepic (2011, p. 14), the existence of these differences is closely related to
the nature of the final product, with important implications in terms of innovation, rationalization or labor division. In what follows, we briefly discuss the potential effects on structural change of cross-sector differences in productivity growth, cross-sector differences in factor intensity, and cross-sector differences in the elasticity of substitution between factors.

3.1.2.1 Sectoral Differences in Total Factor Productivity Growth

Pioneer in studying structural change driven by cross-sector differences in technology, Baumol (1967) early formulates the “cost disease” hypothesis according to which economic resources, in particular labor, move from the dynamic or “progressive” sectors, those characterized by a relatively high rate of technical progress, to the stagnant or “non-progressive” sectors, those characterized by a relatively low rate of technical progress, leading the economy’s growth rate on a declining trend - the so-called “cost disease”. In a very influential paper, Ngai and Pissarides (2007) generalize Baumol’s theory and provide it with stronger foundations. In particular, they derive the conditions under which non-balanced technical progress - unequal (exogenous) sectoral rates of productivity growth - is able to generate a balanced growth path along with structural change in a multi-sector model of growth with homothetic preferences and two production factors, namely labor and capital. While structural change requires an elasticity of substitution across final goods different from one, a necessary and sufficient condition for the economy to exhibit a balanced growth path is a logarithmic intertemporal utility function, i.e. a unit intertemporal elasticity of substitution. Along the balanced growth path, labor employed in producing the consumption goods gradually moves to the sectors with relatively low TFP growth rates provided that the elasticity of substitution between final goods is relatively small, that is lesser than one\textsuperscript{10}. If the elasticity of substitution is greater than one, labor then gradually moves to the sectors with a relatively high TFP growth rate. If the elasticity of substitution is equal to unity, the sectoral structure of em-

\textsuperscript{10}This condition is empirically verified for various sets of consumption sectors (e.g., Ngai and Pissarides, 2004; Buera and Kaboski, 2009, Herrendorf et al., 2013; Boppart, 2014; Comin et al., 2015; Swiecki, 2017).
ployement remains constant, as well as the nominal value added shares and nominal consumption shares. In the model designed by Ngai and Pissarides (2007), structural change is the result of changes in relative (sectoral) prices resulting from differences in productivity growth across sectors. For the model to be able to replicate the trajectory of the employment shares observed in advanced countries, a few assumptions need to be made. More particularly, in case the elasticity of substitution across the broad sectors (agriculture, manufacturing and services) is relatively small, one needs to assume that productivity is growing at the fastest rate in agriculture and at the slowest rate in services.

Baumol’s intuitions, formalized by Ngai and Pissarides (2007) in a neoclassical growth framework, have fed many empirical works. Empirical evidence thus reveals important differences in productivity growth across sectors. For instance, based on US data, Baumol et al. (1985) find that employment growth was absorbed predominantly by the stagnant subsectors of services between 1947 and 1976. Relying on the EU KLEMS Database (O’Mahony and Timmer, 2009), Herrendorf et al. (2014) show that agriculture exhibited the highest TPF growth rate in the period 1970-2007 for a set of advanced countries, including Australia, Canada, the EU10 and the USA, while TFP in services grew at the slowest rate, thus joining the conclusions of Bernard and Jones (1996) for the period between 1970 and 1987. These findings are consistent with the conditions needed in the model of Ngai and Pissarides (2007) for labor to move away from agriculture and manufacturing and towards services in the postwar period.  

Focusing on the very long period, Alvarez-Cuadrado and Poschke (2011), based on data in a range of advanced countries, estimate that the relative price of agriculture to non-agriculture rose before World War II, the period which witnessed the most massive reallocation of labor out of agriculture, then declined in the postwar period. Under the assumption that relative prices rose before World War II, the period which witnessed the most massive reallocation of labor out of agriculture, then declined in the postwar period. Under the assumption that relative prices

---

11Focusing on the differences in the trajectory of manufacturing employment between Asian and Latin American countries over the period 1963-2010, Üngör (2017) shows that a multi-sector general equilibrium model of growth with sectoral differences in (exogenous) productivity growth rates can account for some of the differing sectoral allocations of employment.
between sectors only reflect relative productivities, as it is the case in the model of Ngai and Pissarides (2007), the relative price evolution estimated by Alvarez-Cuadrado and Poschke (2011) suggests that sectoral differences in productivity growth are not stable over time. Accordingly, productivity growth in agriculture would have been lower than productivity growth in non-agriculture before World War II, and faster after World War II. While the “push effect” of faster productivity growth in agriculture, suggested by scholars like e.g. Ngai and Pissarides (2007) and Üngör (2013), is consistent with this result for the postwar period, the relative price evolution of agricultural goods before World War II is rather consistent with the “pull effect” of faster productivity growth in non-agriculture during the early stages of economic development (e.g., Lewis, 1954; Hansen and Prescott, 2002; Doepke, 2004; Bar and Leukhina, 2010). This shows well the importance of introducing more flexible production structures into the analysis, and in particular production structures allowing non-constant biased technological change, if the objective is to account for structural change over the very long period. For instance, a few multi-sector models of growth have interestingly considered productivity improvements in agriculture as a positive externality of technological progress in non-agriculture (e.g., Johnson and Evenson, 1999; Matsuyama, 1992, 2008; Guillo et al., 2011; Grossmann, 2013), thus rooting the “push-effect” of faster productivity growth in agriculture during the post-war period in the prior development of non-agriculture (modern industry). Thus, in the study by Guillo et al. (2011), cross-sector knowledge spillovers - agriculture benefits from spillovers from the rest of the economy - are required for the model to be able to replicate the historical pattern of sectoral TFP growth.

12The “push effect” refers to the situation where the sectors characterized by relatively high rates of productivity growth push resources, especially labor, to the sectors characterized by relatively low rates of productivity growth. In the model of Ngai and Pissarides (2007), it is the case when the elasticity of substitution between sectors is relatively small, that is lesser than one

13The “pull effect” refers to the situation where the sectors characterized by relatively high rates of productivity growth pull resources, especially labor, from the sectors characterized by relatively low rates of productivity growth. In the model of Ngai and Pissarides (2007), it is the case when the elasticity of substitution between sectors is relatively high, that is higher than one
3.1.2.2 Sectoral Differences in Factor Intensity

Also focusing on structural change resulting from changes in relative sectoral prices, other scholars have developed models in which the changes in relative prices, which lead the households to modify the (sectoral) allocation key for their nominal income, are not driven by sectoral differences in TFP growth rates. For instance, assuming that technical progress is neutral across sectors of activity but differentiating between skilled and unskilled workers, Caselli and Coleman (2001) present a model in which the decline in effective education cost, a situation observed in the first half of the 20th century, increases the relative supply of skilled workers and decreases the relative price of non-agricultural products which are more skill intensive, thus contributing to a movement of labor out of agriculture and toward modern industries. Acemoglu and Guerrieri (2008) present a two-sector model of growth in which the sectors differ in capital intensity. In this context, changes in relative prices are driven by the relative supplies of capital and labor. The increase in the aggregate “capital-to-labor” ratio which accompanies the process of (neutral) TFP-led economic growth then contributes to decreasing the relative price of the good from the capital-intensive sector. Under the assumption that the elasticity of substitution between the two sectors is lesser than one, i.e. that the goods exhibit a relatively high degree of complementarity, labor moves away from the capital-intensive sector for similar reasons as in Ngai and Pissarides (2007). The economy reaches a balanced growth path asymptotically, with the long-run rate of growth being determined by the asymptotically dominant sector, namely the labor-intensive sector. For the model to show ability to replicate the trajectory of the employment shares observed in advanced countries for the sectoral trichotomy agriculture, manufacturing and services, one needs to make the assumption that agriculture is more capital-intensive than manufacturing and that manufacturing is more capital-intensive than services.

In line with the work of Acemoglu and Guerrieri (2008), empirical evidence shows large disparities in capital intensity and, hence, in factor income shares across sectors. For instance, based on US data for the period 1950-1994,
Kongsamut et al. (1997) document that the labor income share is relatively high in sectors of activity like manufacturing and construction, at around 70%, whereas it is relatively low in sectors like agriculture, finance, insurance and real estate, at around 20%. These findings are broadly similar to the ones of Close and Shulenburger (1971) for the period 1948-1965 and Acemoglu and Guerrieri (2008) for the period 1987-2004. In the context of the sectoral trichotomy agriculture, manufacturing and services, Herrendorf et al. (2015) find, from the estimations of sectoral production functions using US data for 1947-2010, that the agricultural sector is more capital intensive than services and that services are more capital intensive than manufacturing. As these findings contrast with the conditions needed in the model of Acemoglu and Guerrieri (2008) for labor to move away from manufacturing and towards services, Herrendorf et al. (2015) conclude that cross-sector differences in capital intensity are not the main quantitative force on the technology side behind the postwar US structural change.

3.1.2.3 Sectoral Differences in the Elasticity of Substitution between Factors

Within an encompassing framework, which includes the growth models of Ngai and Pissarides (2007) and Acemoglu and Guerrieri (2008) as particular cases, Alvarez-Cuadrado et al. (2017) have recently investigated another source of structural change, namely cross-sector differences in the elasticity of substitution between capital and labor. In addition to affecting relative sectoral prices, the existence of sectoral differences in the substitutability between capital and labor leads to the emergence of a specific effect, labelled as the “factor rebalancing effect”, which actually competes with the “relative price effect” in determining how factor allocations react to capital accumulation and technological change. Intuitively, the “factor rebalancing effect” captures the idea that the most flexible sectors, i.e. those characterized by a relatively high elasticity of substitution between capital and labor, can better

14These results are qualitatively similar to those previously derived by Herrendorf and Valentiný (2008) when the capital shares refer to final expenditure and not to value added.
take advantage of rises in the “wage to rental rate” ratio as the aggregate capital-labor ratio grows, i.e., as capital becomes relatively more abundant, by substituting away from labor. As a result, sectoral capital-labor ratios can grow at different rates, allowing the factor income shares, which are not structural parameters in the presence of CES production functions\textsuperscript{15}, to evolve with the passage of time, an evidence recently documented by e.g., Elsby et al. (2013), Karabarbounis and Neiman (2014) and Alvarez-Cuadrado et al. (2015). By contrast to the models of Ngai and Pissarides (2007) and Acemoglu and Guerrieri (2008), the model of Alvarez-Cuadrado et al. (2017) even makes possible that the fractions of capital and labor used in any sector move in the opposite direction. As pointed out by the authors, the results that they obtain may be particularly useful for the study of structural change in terms of other factors of production for which substitutability is different across sectors, like the substitutability between skilled and unskilled labor (see e.g., Reshef, 2013; Buera et al., 2015; Wingender, 2015).

The existence of sectoral differences in the elasticity of substitution between capital and labor has also been evidenced in the empirical literature. For instance, in the context of the sectoral trichotomy agriculture, manufacturing and services, Herrendorf et al. (2015) find that capital and labor are more easily substitutable in agriculture than in manufacturing and more easily substitutable in manufacturing than in services. More particularly, they estimate that the elasticity of substitution between capital and labor is equal to 1.58 for agriculture, while it is equal to 0.8 and 0.75 for manufacturing and services respectively. By use of these estimates, Alvarez-Cuadrado et al. (2017) conduct a quantitative analysis of the US experience and show that their multi-sector model is able to reproduce almost exactly the process of reallocation of labor away from agriculture, as well as the faster growth of the capital-labor ratio in agriculture compared to the one in the economy.

\textsuperscript{15}This contrasts with the Cobb-Douglas sectoral production functions used by Ngai and Pissarides (2007) and Acemoglu and Guerrieri (2008).
3.1.3 The Relative Importance of Income Effects (Preferences) and Relative Price effects (Technology)

There seems to be an increasing consensus in the literature on the idea that income and relative price effects together are needed to account for the existing empirical regularities associated with structural change. By way of illustration, Dietrich and Krger (2010) show that the model developed by Ngai and Pissarides (2007), when it is calibrated using the commonly accepted values employed in economic growth theory, can hardly replicate the trajectories of the employment shares of the three broad sectors agriculture, manufacturing and services, and in particular the hump-shaped development for manufacturing, in the absence of non-homothetic preferences (income effects). In the same vein, Herrendorf et al. (2014) and Comin et al. (2015) point out the importance of introducing non-homothetic preferences along with heterogeneous sectoral production functions into the models of structural change if the objective is to generate a positive correlation between the nominal and real measures of economic activity - which is a robust feature of US data - when the consumption goods are assumed to be complement.

In line with the recommendations of Acemoglu (2009), recent literature has thus made a great deal of effort to incorporate non-homothetic preferences and cross-sector differences in technology within the same analytical framework (e.g., Bonatti and Felice, 2008; Dennis and Iscan, 2009; Guillo et al., 2011; Buera and Kaboski, 2009, 2012a, 2012b; Boppart, 2014; Alonso-Carrera and Raurich, 2015; Comin et al., 2015; Alvarez-Cuadrado et al., 2017; Swiecki, 2017). A strand of this literature has been directed towards quantifying the relative importance of the income and relative price effects in driving the observed patterns of structural change. Though most studies are based on US data, the results of this literature are heterogenous. This heterogeneity has different sources. First, the investigating period can differ across studies. While some studies focus on the long period starting from the 19th century, other studies only focus on the postwar period. Secondly, the number and, hence, the definition of the consumption sectors can differ across studies. While some studies only define two consumption sectors, putting for-
ward the massive shift of activity away from agriculture or towards services, other studies include the three broad sectors (agriculture, manufacturing, and services) into the analysis. Thirdly, the studies do not all include the same determinants of structural change. This is particularly true for the economic forces located on the technological side. Fourthly, the studies do not all use the same measure of structural change to test the ability of their model to fit the data on structural change. In order to help the reader to understand some potential sources of heterogeneity of the results in the studies mentioned hereinafter, table 1 summarizes some of the specificities of these studies.

Pioneer in dealing with income and relative price effects within the same accounting framework, Dennis and Iscan (2009) simulate a model with non-homothetic preferences and sectoral differences in both TFP growth and factor income shares, using two centuries of US data. They find that that the “Engel effect” - income effect resulting from non-homothetic preferences - accounts for almost all labor reallocation until the 1950s, after which the “Baumol effect” - relative (sectoral) price effect resulting from cross-sector differences in productivity growth (sector-biased technological change) - becomes a key determinant.16 This finding is corroborated, at least to some extent, by the recent quantitative exercise of Swiecik (2017). Based on data for a set of 45 diverse countries over the period between 1970 and 2005, Swiecik (2017) indeed shows that the relative importance of these two classic mechanisms depends on how far ahead a country is in the process of economic development and structural change.

By contrast with Dennis and Iscan (2009), Guillo et al. (2011) provide support to the sector-biased technological change hypothesis as an important driver of structural change in both the 19th and 20th centuries. Interestingly, the model of Guillo et al. (2011) is able to replicate not solely the patterns of structural change out of agriculture in the US economy but also the patterns of relative productivity - supported by estimated data on the evolution of the relative price of agriculture to non-agriculture (e.g. Caselli

---

16 In another work, Iscan (2010) concludes from the main results of his calibrated three-sector model of structural change that jointly “Engel’s law” and “Baumol’s disease” account for approximately two-thirds of the reallocation of labor into services in the US over the 20th century.
and Coleman II, 2001; Kongsamuth et al., 2001; Johnson, 2002; Dennis and Iscan, 2009; Alvarez-Cuadrado and Poshke, 2011) - by developing a two-sector overlapping-generations model of endogenous technical change and growth which considers the possibility that the agricultural sector benefits from spillovers from the rest of the economy. In the same vein, Cruz (2015) shows that a model with solely relative price effects, driven by sectoral differences in TFP growth, performs well in terms of fitting the US data on structural change throughout the whole process of labor reallocation away from the agricultural sector when allowing for non-constant (time-variant) biased technological change and assuming a technological backwardness of agriculture. The studies of Guillo et al. (2011) and Cruz (2015) perfectly illustrate how using more flexible production structures can help to better capture the role of technology in driving the historical pattern of structural change over the very long period in advanced countries\(^{17}\).

Focusing on the postwar period using data on US experience, Boppart (2014) exploits the functional form of his multi-sector growth model\(^{18}\) and finds that the income and relative price effects have roughly been of equal importance in accounting for structural change. Confirming the significant contribution of both mechanisms to postwar US structural change, yet Comin et al. (2015) find that the income effects have been quantitatively more important in driving the observed patterns. This result lies on the simulation of a structural change model with independent relative price and income effects, which is made possible by using non-homothetic CES preferences instead of the Stone-Geary preferences - or the Price-Independent Generalized Linear (PIGL) preferences employed by Boppart (2014). By Contrast to the widely-used Stone-Geary preferences, non-homothetic CES preferences are moreover able to generate log-linear Engel curves with stable slopes\(^{19}\), thus allowing

\(^{17}\)As an attempt to consider more flexible production structures, some recent studies analyze structural change in the context of endogenous technological progress at the sector level (e.g., Boppart and Weiss, 2013; Struck, 2014; Hori et al., 2015; Herrendorf et al., 2015).

\(^{18}\)Boppart (2014) was the first to present a neoclassical growth model with intertemporal optimization that reconciles the Kaldor facts with structural change driven by income and relative price effects.

\(^{19}\)This empirical evidence has been shown by Young (2012) and Aguiar and Bils (2015).
for non-vanishing income effects as real income increases and restoring the potential for non-homothetic preferences to explain structural change in the long run.

In addition to the choice of a specific class of utility functions, Herrendorf et al. (2013) show that the results of the studies aimed at quantifying the relative importance of the forces behind structural change critically depend on the specification of the commodity space, i.e. whether it is in terms of final consumption expenditure or consumption value added. Using US data for the period 1947-2010, Herrendorf et al. (2013) indeed find that the income effects are the dominant force behind structural change when the preferences and, hence, the sectoral production functions are defined over the final expenditure categories, whereas the relative price effects are more important with the value added categories. Herrendorf et al. (2013) show how the input-output structure of the United States can reconcile these findings.

Broadly speaking, the calibration of multi-sector models of growth with income and relative price effects yields good results for the monotone labor dynamics between agriculture and non-agriculture on the one hand, and between goods and services on the other hand. By contrast, the reproduction of the patterns of structural change among the three broad sectors (agriculture, manufacturing and services) as from the early stages of economic development, including the hump-shaped trajectory for manufacturing, appears to be more challenging. By way of illustration, evaluating the ability of the traditional theories of structural change to fit US data, Buera and Kaboski (2009) show that a combination of the two driving forces proposed by Kongsamut et al. (2001) - income effect - and Ngai and Pissarides (2007) - Baumol effect - misses several stylized facts of structural change, in particular the steep decline in manufacturing and the sharp increase in services in the later data, i.e. after 1970. The following sections of this paper show that amending the traditional models of structural change by introducing the role of input-output (sectoral) linkages and international trade, two mechanisms largely overlooked in the recent multi-sector growth literature, appears to be promising in terms of improving the fit to the actual data in developed countries, but also in developing countries.
Table 1: Characteristics of studies quantifying the relative importance of preferences and technology

<table>
<thead>
<tr>
<th></th>
<th>Sectors¹</th>
<th>Structural Change Measure(s)²</th>
<th>Prewar Period</th>
<th>Postwar Period</th>
<th>Non-Homothetic Preferences</th>
<th># TFP Growth</th>
<th># Capital Intensity</th>
<th># Elasticity of Substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dennis and Iscan (2009)</td>
<td>A-NA</td>
<td>E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guillo et al. (2011)</td>
<td>A-NA</td>
<td>E, VA, C</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herrendorf et al. (2013)³</td>
<td>A-M-S</td>
<td>C</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boppart (2014)</td>
<td>G-S</td>
<td>E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swiecki (2017)³</td>
<td>A-M-S</td>
<td>E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
¹ A=Agriculture; NA=Non-Agriculture; M=Manufacturing; S=Services; G=Goods.
² E=Employment; VA=Value Added; C=Consumption; X=Capital.
³ Though we select both boxes, Herrendorf et al. (2013) do not specify the source(s) of relative price effects in their exercise (estimation of preferences, i.e. of utility functions).
⁴ The model of Swiecki (2014) also includes other determinants of structural change, namely international trade and changing wedges between factor costs across sectors.
3.2 Input-Output linkages and Structural Change

While the recent multi-sector growth literature has primarily investigated the final demand channels, especially the income and relative price effects, through which the process of reallocation of economic activity across sectors can take place in market economies, a number of studies stress the importance of taking explicitly into account the input-output (sectoral) linkages, as firms offering final goods and services are in turn “consumers” of intermediate goods (e.g. Pasinetti, 1981; Fixler and Siegel, 1999; Berlingieri, 2014). As documented by Jones (2013) using input-output data, the share of intermediates in gross output is about half across a large range of countries, with large disparities across sectors. For instance, services typically have a lower share. As a consequence, changes in the composition of intermediates, i.e. changes in the input-output linkages, have the potential to dramatically influence the allocation of labor across sectors and structural change.

A key driver of the evolution of the input-output structure of an economy is technology. Thus, Verspagen (2004) shows that the major changes observed in the input-output structure of most currently rich countries over their course of economic development have mirrored the broad history of technological change from an economic perspective. This observation is largely in line with the Schumpeterian theory according to which modern growth has been spurred by successive waves of technological breakthroughs, each of which “revolutionizing” the economic structure from within, and modifying the sectoral composition of production as part of a process of “creative destruction” characterized by the disappearance of old activities and the emergence of new modern ones. These waves have very commonly been labelled as “industrial revolutions”. For instance, the First Industrial Revolution is traditionally linked to the introduction of major technical breakthroughs in the textiles and the metallurgy, the growing use of steam power in both manufacturing and transportation, and the rise of the factory system. The Second Industrial Revolution is viewed as the “age of electricity and steel”. Under the impulse of individual manufacturing sectors, like the automotive sector, it is followed by a phase of generalization of the mass production sys-
tem, which relies on the application of economies of scale and scope, and the availability of cheap energy and new materials. The Third Industrial Revolution is closely associated with the introduction and gradual diffusion of Information and Communication Technologies (ICT) in the economy, with great implications in terms of dynamic performance for the industries able to take advantage of these new technologies (e.g. Antonelli, 1990; Castellaci, 2010). It is characterized by a relatively rapid growth of ICT services, leading to the vision of a society in deep transformation towards an “information age”.

While the Second Industrial Revolution gave rise to the development of large, vertically integrated companies, the ICT revolution has rather tended to re-shape firms’ boundaries, fostering a movement of vertical disintegration and favoring the implementation of “outsourcing” strategies, defined as “the process by which in-house provided activities are replaced by the provision from external agents” (Merino and Rodriguez, 2007, p. 1148). Over the last few decades, firms have thus increasingly attempted to specialize in their core competencies, outsourcing non-core activities to external suppliers (Pavitt, 2005). Facilitated by the progressive diffusion of ICT in the whole economy, this movement of vertical disintegration has been highly influenced by rising globalization (e.g., Windrum and Tomlinson, 1999; Franke and Kalmbach, 2005), the rising complexity of business operations (e.g., Holweg and Pil, 2012; Berlingieri, 2015), the rising cost of monitoring workers, and the fast rate of technological progress in ever more specialized service activities.

The movement of vertical disintegration observed over recent decades in advanced countries has been accompanied by the parallel substantial expansion of services for which final demand plays only a relatively small role\textsuperscript{20}, in particular of Professional and Business Services (PBS). PBS have been the fastest growing and the most dynamic component of the services sector in most advanced countries. For instance, Berlingieri (2014) shows that

\textsuperscript{20}These services are said to exhibit a high degree of forward linkage (Acemoglu \textit{et al}., 2012). Inter-sectoral linkages are commonly categorized into two types: backward and forward linkages. While backward linkages indicate the extent to which the output of a specific sector generates positive externalities to the rest of the economy, in particular though the demand for intermediate inputs, forward linkages indicate the extent to which the output of a specific sector is used as an input in production processes.
the employment share of PBS increased by 9.2 percentage points in the US between 1948 and 2007, accounting for around 40% of service employment growth. Using statistics from the EU KLEMS Database (Jäger, 2016), table 2 shows that the employment share of PBS has also grown dramatically in the EU15 and countries such as Australia and Japan. The rapid expansion of PBS has actually revealed their increasing integration in production processes and their growing strategic importance as a source of flexibility (e.g., Coffey and Bailly, 1991), efficiency gains (e.g., Di Cagno and Meliciani, 2005) and innovation. In a context of highly-sophisticated markets in which firms have to mobilise ever expanding ranges of diversified skills, which are sometimes beyond their internal capabilities, the literature has thus emphasized the key role of Knowledge Intensive Business Services (KIBS) as producers of innovation but also as intermediaries and nodes in innovation systems by which they contribute to transforming and diffusing knowledge and technology (e.g., Tomlinson, 1997; Windrum and Tomlinson, 1999; Muller and Zenker, 2001; Strambach, 2001, 2008; Den Hertog, 2001; Baker, 2007; Consoli and Elche-Hortelano, 2010; Shearmur and Doloreux, 2013; Ciriaci and Palma, 2015; Ciriaci et al., 2015).

Table 2: Employment share of professional and business services in 1970 and 2007 (in %)

<table>
<thead>
<tr>
<th></th>
<th>1970</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>3.2</td>
<td>10.9</td>
</tr>
<tr>
<td>EU15</td>
<td>3.4</td>
<td>12.7</td>
</tr>
<tr>
<td>Japan</td>
<td>2.2</td>
<td>10.5</td>
</tr>
</tbody>
</table>

Source: EU KLEMS Database (Jäger, 2016), author’s calculations.

While explicitly considering sectoral linkages may not be of high relevance for the study of structural change when the phenomenon is defined in terms of the broad sectors agriculture, manufacturing, and services, the rationale for incorporating sectoral linkages in the multi-sector models of structural change is given by the substantial increase of the service sector’s position as
a downstream provider of intermediates, an evolution mirrored by the parallel modification of the input-output structure of advanced countries since the 1970s. As noted by Lind (2014, p. 43), there is indeed strong empirical evidence of production processes becoming more dependent on service inputs, not the least within manufacturing\(^{21}\). The growing service intensity of manufacturing has deep implications for the study of structural change. Identified as a key source of the changes in the composition of intermediates in manufacturing, outsourcing, in particular PBS outsourcing, has indeed led to the gradual emergence of What Bryson and Daniels (2010) have called the “manuservice economy” in which the boundaries between manufacturing and services are becoming ever more blurred, especially as there is a rising trend for manufacturing firms to combine goods with services to offer more differentiated and more customized products to their clients (e.g. Crozet and Milet, 2014). Thus, many service-type activities, like catering, cleaning, accounting, transport, design and research, which were previously performed in-house by manufacturing firms have been increasingly outsourced to specialized service providers, inducing a reclassification of economic activity in favor of services. This part of the service sector’s growth at the expense of manufacturing has sometimes been referred to as a “statistical illusion” or “statistical artefact” (e.g., Schön, 2010, p. 397; Tregenna, 2015, p. 30).

As pointed out by Rowthorn and Coutts (2013), a wider definition of manufacturing, including all service inputs that are embodied in the final output of manufacturing, would not solely increase the size of manufacturing but also reduce the extent of deindustrialization, and hence of service growth, observed in most advanced countries since the last third of the 20th century. This sort of argument has motivated a number of scholars to study the movement of deindustrialization outside the sectoral trichotomy (agriculture, manufacturing, and services) as it is defined by most standard statistical classifications of economic activities such as NAICS and NACE (see Schettkat and Yocarini, 2006, pp. 133-139). For instance, based on the works of Katouzian (1970) and Singelmann (1978), some scholars propose to disaggregate the broad service sector and to make the distinction between “distributive services”, “producer

\(^{21}\)See e.g. Falk and Peng (2013) for further evidence.
services”, “social services”, and “personal services”. The first two categories are then usually interpreted as being related to goods production, though “producer services” provide intermediates for service provision as well. Utilizing this fourfold classification of services, several studies find that the shifts in intermediate demand in the manufacturing sector have largely contributed to fueling the growth of services (e.g., Elfring, 1988, 1989; Castells, 1996). Other scholars propose to look at occupational employment, which makes the distinction between “service-type” activities and “manufacturing-type” activities, instead of industry employment (e.g., Freeman and Schettkat, 1999; Tregenna, 2014). The advantage of this approach is that it identifies all service-type activities, irrespective of the industry (manufacturing or services) in which they are performed. Finally, a set of papers also resort to input-output statistics to assess the extent of outsourcing from manufacturing to services (e.g., Petit, 1986; Russo and Schettkat, 1999 and 2001; Greenhalgh and Gregory, 2001; Peneder et al., 2003; Gregory and Russo, 2006; Demmou, 2010). For instance, using French input-output data, Demmou (2010) finds that services outsourcing contributed around 20% of employment losses in manufacturing between 1980 and 2007.

While numerous empirical studies evidence the key role of input-output linkages, in particular of PBS outsourcing, in contributing to the expansion of services over the last few decades, the recent multi-sector growth literature has largely overlooked this channel of structural change. In their influential work, Ngai and Pissarides (2007) do generalize their baseline growth model by explicitly considering the possibility that sectors produce both final and intermediate goods. Making the assumption that all the intermediate goods are used as inputs into an aggregate CES production function producing a single intermediate good, Ngai and Pissarides (2007) notably show that the principal results that they derived from their baseline model remain unchanged when the aggregator is “Cobb-Douglas”, which appears to be a necessary and sufficient condition for the existence of a balanced growth path along with structural change. The fraction of total employment devoted to producing the intermediates is constant in the long run, both at the sectoral level and at the aggregate level, and structural change takes place in the form of a sec-
toral reallocation of labor devoted to producing the final consumption goods, with labor going out of the high-TFP-growth sectors and towards the low-TFP-growth sectors. One major drawback associated with the treatment of intermediate inputs in the model of Ngai and Pissarides (2007) is the absence of sectoral differences in the use intensity of intermediates. Indeed, the share of intermediates in gross output is the same in every sector. Furthermore, the aggregate intermediate good is produced identically across sectors based on a Cobb-Douglas aggregator that does not allow for a process of reallocation of labor devoted to producing the intermediates across sectors. These functional restrictions highly contrast with empirical evidence, and in particular with the increasing use of service intermediates in production processes.

To the best of my knowledge, the study conducted by Berlingieri (2014) is the first attempt to investigate, within the analytical framework of a multi-sector growth model, the role played by changes in the composition of intermediates in shaping both the sectoral allocation of labour and structural change. Using his calibrated model - with intermediate inputs and full sectoral linkages - to predict the trajectory of the employment shares of manufacturing and services in the USA over the period 1948-2002, Berlingieri (2014) finds that the evolution of the input-output structure, which is mostly due to PBS outsourcing, can account for 36% of the growth in services employment and 25% of the decline in manufacturing.

Using an open-economy setting, Sposi (2016) also explores the role of sectoral linkages in driving the allocation of economic activity across the three sectors agriculture, manufacturing and services. Documenting that input-output linkages systematically differ across levels of economic development, in particular the service intensity of manufacturing, Sposi (2016) argues there are two main channels through which input-output linkages matter for structural change. The first channel is related to how input-output linkages influence the impact of productivity shocks on relative prices across sectors of activity. Thus, the impact of a rise in the relative productivity of manufacturing on the

---

22 As part of the literature accounting for the growth of services for which final demand plays only a relatively small role, a number of studies also use a general equilibrium framework to provide an explanation of structural change towards finance (e.g., Philippon, 2014; Gennaioli et al., 2014; Falkinger et al., 2015).
relative price of manufactures depends on the use intensity of services as intermediate goods in the manufacturing production processes. Accordingly, a same rise in manufacturing productivity translates into a larger decline in the price of manufactures in poor countries compared to rich countries, as poor countries typically use manufactured goods more intensively. The second channel is related to how changes in final demand patterns maps into changes in the sectoral structure of value added and employment, which in fact depend exclusively on sectoral linkages. Based on a three-sector (multi-country) model in which structural change is driven by income effects, the Baumol effect (cross-sector differences in productivity growth rates), changes in comparative advantage via trade, and sectoral linkages, Sposi (2016) studies, more specifically, the extent to which cross-country differences in intermediate-input intensities can explain cross-country differences in the composition of economic activity. As a result of a quantitative exercise consisting of a set of counterfactuals, he finds that cross-country differences in sectoral linkages account for 74 percent of the curvature in the hump shape in industry’s share in value added across levels of economic development, which is twice as much as can be explained by variations in the composition of final demand (final domestic expenditures plus net exports). These results clearly give support for considering explicitly input-output (sectoral) linkages in the analysis of structural change in future work.

Regarding this last point, it is noteworthy that the investigation of the role of intermediate inputs in driving structural change in growth models requires a specific interpretation of commodities in the utility function. By way of reminder, Herrendorf et al. (2013) highlight that the utility function can be defined over the consumption of either sectoral “value added” or “final expenditure” categories, with important implications in terms of defining the corresponding sectoral production functions. Though the “value-added” specification implicitly takes into account, at least to some extent, the input-output structure of the economy, it does not offer the possibility to introduce explicitly intermediates, both domestic and foreign intermediates, into the sectoral production functions, making it hard to analyze the impact of changes in the composition of intermediates on structural change. As a consequence, the
studies aim at investigating the role of changes in input-output linkages and, more specifically, the role of increased outsourcing on structural change should opt for the more flexible final expenditure specification

3.3 International Trade and Structural Change

The theoretical literature on structural change has overwhelmingly focused on the factors that are likely to generate a process of reallocation of activity across sectors within a single economy, thus implicitly considering that this economy is not interacting with the rest of the world. One important implication of this restrictive assumption is that the sectoral productions must necessarily mirror the corresponding household decisions, either of consumption or of investment (Herrendorf et al., 2014). Yet a simple look at the data of any country reveals that such conditions are virtually never satisfied. As Matsuyama (2009, p. 486) notes: “We need to keep reminding ourselves of the simple truth: we live in an interdependent global economy and our planet, the world economy, is the only closed economy we know of”. In line with this economic reality, relaxing the assumption of closed borders opens up new perspectives to explore structural change.

According to trade theory, opening up the borders brings about a process of structural change in every economy involved in international exchanges.

The idea that outsourcing is part of the determinants of structural change can actually be traced back to the influential book of Fuchs (1968). Identifying outsourcing as one potential explanation for the large expansion of services, Fuchs (1968) indeed argues that contracting out services that were once produced in the firm or the household becomes more efficient as real income rises (Kongsamut et al., 2001). In line with this view, a strand of the literature on structural change has explicitly introduced home production into analysis as a factor likely to account for part of the labor market shift towards services (e.g., Freeman and Schettkat, 2005; Ngai and Pissarides, 2008; Rogerson, 2008; Buera and Kaboski, 2012a, 2012b; Barany and Siegel, 2016). As part of this literature, a number of studies have used multi-sector models to investigate the interactions between structural change, the marketization of home production and female work, and explain the evolution of gender outcomes in working hours and wages (e.g., Ngai and Pissarides, 2008; Rendall, 2015; Olivetti and Petrongolo, 2016; Ngai and Petrongolo, 2017). As noted by Ngai and Pissarides (2008), these papers have been motivated by a few stylized facts revealing, among others, a growth of female work - which has taken place entirely (in net terms) in services - and women’s relative wages, as well as a decline in women’s working hours in the household (see e.g. Bridgman et al. (2015) for empirical evidence on household and market production).
This sectoral reallocation of activity largely takes place as a result of comparative advantage(s) being driven by, among others, technology and factor endowments. A country’s economic structure is therefore directly affected by the specialization patterns induced by trade (e.g., Rowthorn and Wells, 1987; Imbs et al., 2012). In addition, trade has the potential to boost productivity, especially in the sectors exposed to foreign competition, and to foster income growth, thus affecting the sectoral expenditures shares and, hence, the economic structure through the classical channels that we presented previously, i.e. the income and substitution (relative price) effects. This refers to the dynamic effects of trade on structural change.

Matsuyama (2009) has been pioneer in studying explicitly the effects of international trade on the process of structural change, especially on the decline in relative manufacturing employment. Taking explicitly into consideration the global perspective of structural change in a rigorous analytical framework assuming some degree of interdependence across countries, Matsuyama (2009) designs a simple two-country Ricardian model in which the representative household maximizes a Stone-Geary utility function over three consumption goods (food, manufactures, and services) that are produced with technologies linear in labor, the only production factor. Matsuyama (2009) shows that a country characterized by comparatively faster productivity gains in manufacturing can temporarily experience a rise in the share of labor employed in manufacturing and delay or slow down its deindustrialization process, thus formalizing the argument that the same underlying forces can have different implications for structural change in a closed economy and in an interdependent world. As pieces of evidence, Obstfeld and Rogoff (1996) provide the examples of Germany and Japan, two countries with a strong competitive position on the world market for manufactures, which have seemingly experienced a much slower decline in relative manufacturing employment than countries like the US and the UK. The British industrial revolution is another historical example of how foreign trade can affect the trajectory of manufacturing workforce as the increasing British population was, at those times, partly fed through food imports, the latter being financed with the exports of manufactured products (e.g., Hicks, 1999; Clark, 2002; Teignier, 2016).
As an attempt to generalize Matsuyama’s (2009) intuitions, Uy et al. (2013) also propose a two-country Ricardian trade model in which the reallocation of activity across agriculture, manufacturing and services can take place through income and relative price effects, and through trade according to international disparities in relative productivity across sectors. Uy et al. (2013) notably show that an economy having or reinforcing a comparative advantage in manufacturing, for instance as a result of lower trade barriers like lower transportation costs, can exhibit the hump-shaped pattern of relative manufacturing employment, even if manufacturing is the most technologically dynamic sector. In this model, the assumption of being an open economy interestingly allows the manufacturing sector to have both the largest productivity gains and a rising labor share, a situation that highly contrasts with the implications of the (closed economy) model presented by Ngai and Pissarides (2007) when the elasticity of substitution across consumption sectors is relatively small. Using their model to analyze Korea’s episode of rapid industrialization during the postwar period, Uy et al. (2013) conclude from a set of simulations that trade really matters for explaining countries patterns of structural change in an ever more globalized world\textsuperscript{24}.

3.3.1 Trade in Goods and Structural Change in Developing Countries

Despite the relative lack, at least until recently, of a theoretical framework devoted to analyzing the dynamics of structural change under the more realistic assumption of open borders, the impact of international trade on structural change, especially on the trajectory of the manufacturing sector, in the postwar period - which has witnessed a large boom in the volume of world exchanges due in part to declining transportation costs and invisible transaction costs in trade (Krugman, 1996) - has long been investigated in the empirical literature for different regions of the world. For instance, the econometric results of Dodzin and Vamvakidis (2004) suggest that trade led to industrialization in most developing countries over the period 1960-2000. Using a

\textsuperscript{24}See e.g. Betts et al. (2016) and Teignier (2016) for further evidence on the role of trade in driving structural change in South Korea.
three-sector model of structural change, Bah (2014) estimates the sectoral productivity (TFP) paths that are consistent with the process of reallocation of workforce between the three broad sectors (agriculture, manufacturing and services) and the growth of GDP per capita in a range of developing countries over the period from 1960 to 2005. Bah (2014) finds that, relative to the US, developing countries are the least productive in the agricultural sector followed by services, then manufacturing. This result is interestingly in line with the idea of a sectoral specialization of developing countries in manufacturing according to Ricardian comparative advantages\textsuperscript{25}.

Giving a more nuanced picture for developing countries, Szirmai (2012) documents that, after an episode of very modest manufacturing development during the 1950s and the 1960s, a large number of developing countries, especially in Africa and Latin America, have experienced a “premature deindustrialization” since the 1980s, with falling manufacturing shares in both employment and real value added. As an attempt to explain the differences in the trajectory of manufacturing between Asian countries and many developing countries from Africa and Latin America, McMillan et al. (2014) point out the heterogeneous conditions under which the developing countries have gone through growing globalization. In particular, they stress the role of the industrial policy conducted by Asian countries like China, Taiwan and South Korea, which seems to have been conducive to their export industries, as well as the context of overvalued currencies, driven by disinflationary monetary policies and large foreign aid inflows, in which many Latin American and African economies liberalized\textsuperscript{26}. McMillan et al. (2014) also argue that the comparative advantage in natural resources and primary products held by many developing countries in Africa and Latin America may have limited the development of their manufacturing sector and hence the growth contribution of structural change, a situation sometimes labelled as the “resource curse”. In McMillan et al.’s (2014) terms, the “resource curse” hypothesis lies on the low capacity of the mining and quarrying activities to create a lot of jobs and absorb the excess of employment from agriculture and low-productivity

\textsuperscript{25}See e.g. Duarte and Restuccia (2010) for related research.
\textsuperscript{26}See e.g. Bogliacini (2013) for some further evidence.
services.

Providing some support to this view, Gollin et al. (2016), by use of a sample of 116 developing countries for the period 1960-2010, document that the sectoral composition of urban workforce differs between resource-exporters and non-exporters. In particular, developing countries that are highly dependent on resource exports tend to experience a process of urbanization - which is a function of real income across all countries - concentrated in “consumption cities” where economies, by comparison with “production cities”, are characterized by a larger share of workers in non-tradeable services, such as commerce and transportation or personal and government services, and a lower share of workers in manufacturing or tradeable services such as finance. Interestingly, these findings challenge the expected relationship between urbanization and industrialization that has commonly been observed throughout economic history. In addition to revealing differences in employment patterns between the “consumption cities” and “production cities”, Gollin et al. (2016) show that cities in resource-exporters tend to have higher poverty rates for given values of income per capita levels and urbanization rates, a result interpreted as a piece of evidence that “the income boost from resource export makes cities richer, but it does not appear to translate into improved quality of life to the same degree that a real income boost from industrialization would provide” (p. 23). These findings are seen by the authors as an example of “premature urbanization”27.

In a recent work, Rodrik (2016) also contends that trade and globalization have been the main economic force in shaping the industrialization-deindustrialization patterns within the developing world, with the latter being largely composed of countries which are small in world markets for manufactures, where they are essentially price takers. In particular, international trade helped Asian countries with a strong comparative advantage in manufacturing to experience a phase of fast industrialization and become net importers of manufactured products, with significant effects on productivity.

27Documenting, by use of historical data at both the country level and city level over 1500-2010, an upward shift in the urbanization rates at every level of income per capita, especially for richer and poorer countries, Jedwab and Vollrath (2015) provide a survey of the literature on the drivers of the urbanization process over time.
advance and real wages (e.g., Vu, 2017), while it rapidly triggered a process of “premature deindustrialization" in Africa and Latin America, which were lagging behind technologically. It is noteworthy that the socioeconomic consequences of premature deindustrialization are being increasingly discussed among scholars, especially as the manufacturing sector is traditionally believed to exhibit some very “special” properties that make it instrumental in the processes of catching up and development (e.g., Botta, 2009; Szirmai, 2012; Rodrik, 2013; Szirmai and Verspagen, 2015; Tregenna, 2015; Foster-McGregor and Verspagen, 2016; Marconi et al., 2016; Weiss and Tribe, 2016; Romano and Tra, 2017).

3.3.2 Trade in Goods and Structural Change in Developed Countries

In parallel to the fast industrialization of emerging countries in Asia during the second half of the 20th century, the first signs of deindustrialization have been detected in most advanced countries. As a consequence, global integration between rich and poor countries has commonly been blamed for the observed deindustrialization of rich countries, what Krugman (1996) has called the “deindustrialization hypothesis”.

For instance, relying on the main specification of their two-country model based on monopolistic competition and national variations in industry relative productivity, Autor et al. (2013) conclude that increasing Chinese import competition may account for up to one-quarter of the sharp decline in US manufacturing workforce between 1990 and 2007, a period that witnessed an important increase in the trade exposure of the US labor markets. Interestingly, the Ricardian model of Autor et al. (2013) shows that international trade can influence the allocation of labor between traded and non-traded sectors if bilateral trade is imbalanced. Given the assumptions of the model, balanced trade indeed implies that the job-content of exports is just exactly equal to the job-content of imports. Imbalanced trade is one way of breaking this symmetry. Accordingly, the severe deterioration of the

---

28See e.g. Pierce and Schott (2016) for related research.
manufactured trade balance recorded by countries like the UK and the USA since the 1980s, an evolution which has coincided with booming manufactured imports from the developing world, in particular from Asia, may have contributed significantly to the steep fall in relative manufacturing employment in these countries.\textsuperscript{29} Intuitively, another way to break the symmetry between the job-content of exports and imports is to take a Heckscher-Ohlin or specific-factors trade model with heterogeneous sectoral production functions in terms of factor intensities. Such kind of trade model indeed allows for a country's comparative advantage to be also driven by national disparities in relative factor (resource) endowments. The job-content of exports can then be different from the job-content of imports even if bilateral trade is roughly balanced. In line with this view, the results of the studies conducted by Wood (1994, 1995) and Kucera and Milberg (2003) suggest that international trade in manufactures has had a large negative effect on manufacturing workforce in most advanced countries over the last few decades, with the labor-intensive manufacturing sectors having experienced relatively large job losses due to rising North-South trade.

While traded sectors have often been assimilated to manufacturing sectors in trade models, with the implication that economies tend to specialize within the broad manufacturing sector, Herrendorf \textit{et al.} (2014) have recently drawn attention to fast growing trade in services, as this trend is likely to have a dramatic impact on both the nature and speed of structural change in advanced countries like the USA that seemingly have a comparative advantage in certain tradeable market services. In line with this view, a number of countries are believed to have suffered from the “Dutch disease”, a term first coined in 1977 by \textit{The Economist} to describe the harmful effects on Dutch manufacturing of booming Dutch exports in natural resources after the discovery of large natural gas reserves in the North Sea, in particular the Groningen natural gas field in 1959, leading to a sharp inflow of foreign currency in the Netherlands and the appreciation of Dutch currency. While the term “Dutch

\textsuperscript{29}In line with this view, Craighead and Hineline (2015) link part of the compositional changes (structural change) observed in a large range of developed - and developing - countries to the current account reversals, defined as large and persistent decreases in the current account deficits.
disease originally refers to the (potentially) problematic characteristics of countries that are (newly) well-endowed in natural resources, with the recent examples of Norway (e.g. Heide and Holmøy, 2005; compare Bjornland, 1998), Australia (e.g., Corden, 2012; O'Neill and Weller, 2014) and Canada (e.g., Beine \textit{et al.}, 2012), it has been extended to describe deindustrialization arising from a substantial rise in the exports of any sector other than manufacturing, including for instance finance (e.g., Luxembourg, United Kingdom) and tourism (e.g., Greece).

### 3.3.3 Trade in Tasks and Structural Change

In addition to the sectoral specialization induced by trade, intensified international competition, combined with a comparatively higher level of labor costs in advanced countries, generate structural change by forcing global firms to devote efforts into implementing technological and organizational innovations (Peneder \textit{et al.}, 2003; Franke and Kalmbach, 2005). From a theoretical perspective, firms can raise their degree of competitiveness on international markets by modifying or rationalizing their processes production, notably through a change in their ratio of labor to capital, and by reorganizing themselves at a world scale. Regarding this point, due partly to declining trade costs (lower transport costs, etc.) and great progress in information and communication technology, firms have increasingly adopted offshoring strategies over the last few decades. Offshoring, usually defined as the relocation abroad of a company business process to unaffiliated firms - foreign outsourcing - or to own affiliates - foreign integration - (Farinas, 2016), has contributed to increasing the international fragmentation of production and enhancing further the international division of labor, an empirical evidence recently documented by Timmer \textit{et al.} (2014). These trends have been accompanied by a significant growth of trade volumes, with offshoring being one of the fastest growing components of trade (e.g., Grossman and Helpman, 2005; Grossman and Rossi-Hansberg, 2008; Blinder, 2009; Feenstra, 2010; Helpman, 2011; Baldwin and Robert-Nicoud, 2014), leading both countries and firms to specialize.  

\footnote{For studies formalizing this idea, see e.g. Corden and Neary (1982), Corden (1984), Matsuyama (1992) and Harding and Venable (2016).}
in specific activities along Global Value Chains (GVCs).

Broadly speaking, offshoring provides firms with the opportunity to simultaneously reap the benefits associated with the comparative advantage of various countries. Offshoring can involve supporting processes, such as accounting and administrative tasks, but also operational processes, such as manufacturing, with (potentially) important implications for relative manufacturing employment in developed and developing countries. As Williams (1997, p. 18) notes: “In the new international division of labor (NIDS), the control and command functions are located in a network of global cities in the developed nations whilst physical production is increasingly dispersed into a host of developing countries where new and efficient technology can be allied to lower labour costs”. The international fragmentation of production processes in GVCs has actually been reflected in increasing vertical flows of intermediates among ever more interconnected countries, an evolution challenging the nature of globalization as well as the functioning of the world economy (OECD, 2013). Thus, Lind (2014) documents that the explosion of GVCs has been associated with a reduced average manufacturing backward linkage toward the aggregate economy in advanced countries, and a reduced degree of vertical linkages within the broad domestic manufacturing sector. In line with this observation, using trade-linked input-output statistics from WIOD and introducing a new global value chain measure of comparative advantage, Peneder and Streicher (2017) argue that the declining share of manufacturing value added in domestic final expenditures appears to be the main cause of deindustrialization, with some national disparities - especially in the case of the employment shares - being driven by differences in comparative advantage between countries.

While offshoring has often been considered with fear as a major threat to manufacturing jobs and commonly associated with massive unemployment in advanced countries, which likely explains why the topic has been particularly prominent in the political and scientific debate, a number of scholars have qualified this pessimistic view, showing that trade in tasks can generate large efficiency gains and foster productivity advance (e.g., Bhagwati et al., 2004; Amiti and Konings, 2007; Frans and Woerz, 2008; Kasahara and Rodrigue, 2010).
2008; Amiti and Wei, 2009; Goldberg et al., 2010; Stijepic and Wagner, 2012; Wirtz et al., 2015), especially when slow-productivity-growth activities are offshored (e.g., Stijepic and Wagner, 2008; Stijepic, 2011). For instance, using the dynamic growth framework designed by Ngai and Pissarides (2007) in which intermediate inputs are introduced into the sectoral production functions, Stijepic (2011) shows that offshoring can spur the growth rate of the economy, and moreover slow down the process of structural change in the long run as offshoring, through the productivity-enhancing effect, induces a reallocation of labor from consumption-goods production, the one that is subject to structural change, to investment-goods production. Robert-Nicoud (2008) also argues that offshoring - and hence the specialization by function rather than by sector - relaxes the pressure to move abroad, especially to low-wage countries, the entire manufacturing production chain, thus helping to retain the core activities in manufacturing, namely the “complex” tasks such as design, marketing, research and development, or post-production activities, in advanced countries.

### 3.3.4 The Relative Importance of Globalization and Trade

As noted before, the trade channel has not been extensively incorporated in recent multi-sector models of structural change. As a consequence, only a very few studies have attempted to quantify, within a general equilibrium framework, the relative importance of “internal” factors and “external” factors, i.e. those linked to globalization, in driving structural change. The most convincing attempt is the study by Swiecki (2017). Using data for 45 diverse countries over the period 1970-2005, Swiecki (2017) calibrates a model combining in a common framework four different forces behind structural change - (i) sector-biased technological progress, (ii) nonhomothetic tastes, (iii) international trade and (iv) changing wedges between factor costs across sectors - and performs a set of counterfactual simulations to assess the relative importance of each factor.

---

31 See Garcia-Santana et al. (2016) for a recent study analyzing the role of investment demand on structural change.

32 See Fontagné and D’Isanto (2014) for a study, based on a survey of 28,000 French firms, focusing on the main determinants of international sourcing choices.
importance of the four drivers. Among others, Swiecki (2017) finds that sector-biased technological progress is overall the most important mechanism, but that trade can be important for individual countries, even though its impact on the relocation of labor appears to be less systematic.

Focusing on the movement of deindustrialization in advanced countries since the last third of the 20th century, a number of econometric studies also find a significant effect of globalization, especially of rising global integration between rich and poor countries, on relative manufacturing employment, but largely suggest that the internal factors have been quantitatively more important in accounting for deindustrialization in advanced countries (e.g., Rowthorn and Wells, 1987; Alderson, 1997, 1999; Saeger, 1997; Gaston, 1998; Rowthorn and Ramaswamy, 1999; Brady and Denniston, 2006; Boulhol and Fontagné, 2006; Kollmeyer, 2009; Kang and Lee, 2011; Rowthorn and Coutts, 2004, 2013; Skuffic and Druzin, 2016; OECD 2017).

4 Conclusion

The study of structural change has been the object of rising interest over recent decades. This surge of interest has been motivated, at least partly, by the numerous economic concerns and issues raised by deindustrialization which has particularly affected the world’s most economically successful countries since the last third of the 20th century, but also a number of low and middle income nations, especially in Latin America and Africa, since the 1980s. This is likely the reason why the literature on structural change has largely sought to account for the empirical facts associated with the familiar sectoral trichotomy agriculture, manufacturing and services. Recent multi-sector growth theories have identified various mechanisms through which the process of reallocation of activity across sectors can take place in market economies. This survey focuses on four forces behind structural change: changes in aggregate (real) income, changes in relative (sectoral) prices, changes in input-output (sectoral) linkages, and changes in comparative advantages via trade. While the two first forces now seem to be relatively well understood on theoretical grounds, the role of outsourcing and globalization (and trade) has been largely
overlooked in the theoretical literature on structural change. This is surprising as empirical research finds strong support for a key role of these factors in contributing to structural change, in particular to deindustrialization, in both developed and developing countries over recent decades. We show that recent models taking explicitly into account the role of input-output (sectoral) linkages and trade are promising in terms of improving the fit to the actual data on structural change.

In addition to emphasizing the potential importance of incorporating explicitly intermediates and trade into the models of structural change, this survey devotes efforts to presenting the main studies that seek to quantify the relative importance of the forces behind structural change. It notably shows that the findings of this literature are far from being consensual, especially for advanced countries, thus calling for further work in that direction as this research question has important implications, for instance in terms of growth perspectives and public policies. In this regard, it is noteworthy that the recent literature on structural change has been exceptionally silent on policy issues. As noted by Herrendorf and Valentinyi (2015), the likely reason is that it largely utilizes multi-sector models in which structural change arises as an efficient equilibrium outcome by construction, so that industrial policy can solely be harmful. Relaxing the strong usual assumptions that lead to the efficiency of the equilibrium (perfect competition, perfect mobility of factors across sectors, exogenous technological progress at the sector level, etc.) may therefore be very interesting for industrial policy purposes. In particular, it is widely accepted that structural change is associated with a number of costs resulting, among others, from the imperfect mobility of resources (labor and capital) across sectors. In line with this view, some degree of public intervention is potentially desirable to accompany the unavoidable structural change process, especially in an ever more global world in which international competition places considerable pressure on sectors like manufacturing.

We hope that this survey essay could be useful to scholars interested in understanding structural change and contributing to the growing literature on that topic, notably by suggesting amendments to the standard models. A major challenge for future research will be to better take into account the
observed heterogeneity, notably in terms of demand patterns, technology parameters and the use of intermediates, among different sub-sectors within the broad sectors agriculture, manufacturing and services, but especially within the service sector. Indeed, the heterogeneity of services - documented by recent studies like Jorgenson and Timmer (2011), Imbs (2014) and Duarte and Restuccia (2016) - has been largely ignored in most previous research. In this perspective, the availability of increasingly disaggregated data on economic activity at the sector level should help to refine the existing results, as well as to promote the use of other sectoral taxonomies that have the potential to deliver original information and to reveal new patterns of structural change.
5 Appendix A: Country and Variable Coverage in the GGDC 10-Sector Database

<table>
<thead>
<tr>
<th>Region of the World</th>
<th>Country</th>
<th>Value Added (Current Prices)</th>
<th>Employment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-Saharan Africa</td>
<td>Botswana</td>
<td>1964-2010</td>
<td>1964-2010</td>
</tr>
<tr>
<td></td>
<td>Ethiopia</td>
<td>1961-2010</td>
<td>1961-2010</td>
</tr>
<tr>
<td></td>
<td>Ghana</td>
<td>1960-2010</td>
<td>1960-2010</td>
</tr>
<tr>
<td></td>
<td>Kenya</td>
<td>1960-2010</td>
<td>1960-2010</td>
</tr>
<tr>
<td></td>
<td>Malawi</td>
<td>1960-2010</td>
<td>1966-2010</td>
</tr>
<tr>
<td></td>
<td>Mauritius</td>
<td>1960-2010</td>
<td>1970-2010</td>
</tr>
<tr>
<td></td>
<td>Nigeria</td>
<td>1960-2010</td>
<td>1960-2011</td>
</tr>
<tr>
<td></td>
<td>Senegal</td>
<td>1960-2010</td>
<td>1970-2010</td>
</tr>
<tr>
<td></td>
<td>South Africa</td>
<td>1960-2010</td>
<td>1960-2010</td>
</tr>
<tr>
<td>North Africa</td>
<td>Egypt</td>
<td>1960-2013</td>
<td>1960-2010</td>
</tr>
<tr>
<td></td>
<td>Morocco</td>
<td>1970-2012</td>
<td>1960-2010</td>
</tr>
<tr>
<td>Asia</td>
<td>China</td>
<td>1952-2011</td>
<td>1952-2011</td>
</tr>
<tr>
<td></td>
<td>India</td>
<td>1950-2012</td>
<td>1960-2010</td>
</tr>
<tr>
<td></td>
<td>Indonesia</td>
<td>1966-2012</td>
<td>1961-2012</td>
</tr>
<tr>
<td></td>
<td>Japan</td>
<td>1953-2011</td>
<td>1953-2012</td>
</tr>
<tr>
<td></td>
<td>South Korea</td>
<td>1953-2011</td>
<td>1963-2011</td>
</tr>
<tr>
<td></td>
<td>Taiwan</td>
<td>1951-2012</td>
<td>1963-2012</td>
</tr>
<tr>
<td></td>
<td>Thailand</td>
<td>1951-2011</td>
<td>1960-2011</td>
</tr>
<tr>
<td>Latin America</td>
<td>Argentina</td>
<td>1950-2011</td>
<td>1950-2011</td>
</tr>
<tr>
<td></td>
<td>Bolivia</td>
<td>1958-2011</td>
<td>1950-2010</td>
</tr>
<tr>
<td></td>
<td>Brazil</td>
<td>1990-2011</td>
<td>1950-2011</td>
</tr>
<tr>
<td></td>
<td>Chile</td>
<td>1950-2011</td>
<td>1950-2012</td>
</tr>
<tr>
<td></td>
<td>Colombia</td>
<td>1950-2011</td>
<td>1950-2010</td>
</tr>
<tr>
<td></td>
<td>Mexico</td>
<td>1950-2011</td>
<td>1950-2012</td>
</tr>
<tr>
<td></td>
<td>Venezuela</td>
<td>1960-2012</td>
<td>1950-2011</td>
</tr>
<tr>
<td>North America</td>
<td>USA</td>
<td>1947-2010</td>
<td>1950-2010</td>
</tr>
<tr>
<td></td>
<td>Denmark</td>
<td>1970-2011</td>
<td>1948-2011</td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td>1960-2011</td>
<td>1948-2011</td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>1970-2011</td>
<td>1951-2011</td>
</tr>
</tbody>
</table>
Chapter 3

Globalization and Deindustrialization in Advanced Countries

Leif van Neuss

Abstract

The main goal of this paper is to provide an in-depth study of deindustrialization and to analyze the reasons why the world’s most economically successful countries have experienced a fall in relative manufacturing employment over recent decades. A strand of the empirical research on deindustrialization aims at quantifying the relative importance of the economic forces behind the reallocation of labor away from manufacturing, and in particular the relative importance of the “internal” factors and “external” factors, i.e. those linked to globalization and trade. The results of this literature are however very fragile, arguably because the commonly used indicators of manufacturing trade are not well defined to capture the contribution of global exchanges to deindustrialization. While this study does not necessarily contradict the widely accepted idea that internal factors are quantitatively more important in accounting for deindustrialization in the OECD taken as a whole, our econometric results, based on panel data for 15 OECD countries over the period from 1970 to 2006, however suggest that global exchanges have the potential to affect significantly and substantially a country’s sectoral patterns of employment, and that the overall contribution of globalization, especially of increasing North-South integration, to deindustrialization may be revised upwards in advanced countries when resorting to better-defined indicators of trade in manufactures.

Keywords: Structural Change, Deindustrialization, Globalization, Trade

JEL classification: O1, O3, F1, F41, F43
1 Introduction

Structural change or structural transformation, usually defined as the reallocation of economic activity across the three broad sectors agriculture, manufacturing and services, has been the object of growing interest in the scientific literature over the last few decades. The surge of interest in structural change has relied, at least partly, on the numerous issues and economic concerns raised by the movement of deindustrialization that has particularly affected the world’s most economically successful countries since the last third of the 20th century. Thus, deindustrialization has often been associated with social troubles and rising inequality (e.g., Bluestone and Harrison, 1982), growing labor market polarization (e.g., Barany and Siegel, 2016) and relative economic decline. Deindustrialization has particularly been blamed for generating massive unemployment (e.g., Kollmeyer, 2013) and inferior economic growth and productivity improvements (e.g., Baumol, 1967; Baumol et al., 1985; Kitson and Michie, 2014). As evidenced by Palma (2014), the advanced OECD countries began deindustrializing in the late 1960s. By way of illustration, the share of manufacturing in total workforce declined from 28.2% to 15.6% in the EU15 between 1970 and 2007, while it fell from 22.4% to 9.9% in the United States (EU KLEMS Database - O’Mahony and Timmer, 2009). In 2015, the manufacturing share of total employment in the EU15 and the US is estimated to amount respectively 12.5% (Eurostat) and 8.9% (Bureau of Labor Statistics).

The study of deindustrialization really took off in the 1980s as the scale and regional consequences of economic restructuring and job losses in manufacturing became much more apparent and tangible, notably in the United Kingdom (UK) and the United States (US), thus contradicting Lawrence’s (1983) early claim that deindustrialization was a “myth” (Strangleman and Rhodes, 2014). As part of the research agenda, literature has devoted a great deal of effort to identify the economic forces behind the process of reallocation of labor away from manufacturing. Though this research question calls for specific considerations, it is part of a larger debate on the drivers of structural change. Theoretical literature has thus revealed various channels
through which structural change and, hence, deindustrialization can occur. Accordingly, the factors potentially responsible for the fall in relative manufacturing workforce range from preferences (e.g., Kongsamut *et al.*, 2001; Foellmi and Zweimiller, 2008) to technology (e.g., Ngai and Pissarides, 2007; Acemoglu and Guerrieri, 2008; Alvarez-Cuadrado *et al.*, 2016), and also include elements about outsourcing (e.g., Berlingieri, 2014; Sposi, 2016) and globalization and international trade (e.g., Matsuyama, 2009; Uy *et al.*, 2013).

While the economic forces behind deindustrialization now seem to be relatively well understood on theoretical grounds, a number of empirical questions still remain largely open. For instance, the relative importance of the forces behind the process of deindustrialization in advanced countries is not well established. In particular, the role and relative importance of the “internal” factors and “external” factors, i.e. those linked to globalization and trade, is not consensual. Resolving this issue has important implications for how policies, notably trade policies, influence deindustrialization. It is also essential to ensure the appropriate policy response to deindustrialization, especially as structural change is associated with a number of costs resulting, among others, from the imperfect geographical and sectoral mobility of workers (OECD, 2017). As said by Kollmeyer (2009), the unclearness of empirical results is partly because most previous work on deindustrialization has either ignored one or more of the explanations of deindustrialization altogether, or has failed to test all of them simultaneously. In addition to this potential omitted variable bias, we argue that the common measure of globalization in previous econometric studies is inaccurate and can therefore lead to misleading results. Adapting the usual measure of globalization to better capture the contribution of trade to deindustrialization, this study, based on panel data for 15 OECD countries for the period 1970-2006, shows that global exchanges have the potential to affect significantly and substantially a country’s sectoral patterns of employment, and that the contribution of globalization, especially of rising North-South integration, to the observed deindustrialization in advanced countries may be revised upwards when using better-defined indicators of trade in manufactures.

The remainder of this paper is organized as follows. Section II defines the
phenomenon of deindustrialization and provides evidence of the decline of relative manufacturing workforce in advanced countries since 1970. Section III explores the factors identified as “causal” in the occurrence of deindustrialization by the structural change literature. Section IV briefly reviews previous empirical research on the economic forces behind deindustrialization. Section V discusses the empirical methodology used in this study to assess the role of globalization and trade in shaping the decline in relative manufacturing employment. Section VI describes the data. The main empirical results are presented in Section VII. Section VIII deals with the potential endogeneity associated with gross trade flows in our empirical exercise. Section IX gives concluding thoughts and remarks.

2  Deindustrialization: Definition and Empirical Evidence

2.1 What Is Deindustrialization?
Nowadays deindustrialization is commonly defined as the decline of the share of manufacturing in a country’s total economic activity. The most common measures of economic activity are employment and value added, two production-side measures, and consumption (Herrendorf et al., 2014). Even though the different measures of deindustrialization exhibit very interesting features, the employment-based measure of deindustrialization is, by far, the most investigated in the literature. This is probably due to the fact that manufacturing employment is the most visible measure of the size of manufacturing in a country, the one that tends to determine public perceptions of the issue. It is also arguably the most interesting question from a social perspective, especially when concerns about deindustrialization are based on the costs of adjustment across sectors (Nordhaus, 2008).

2.2 Evidence for Currently Rich Countries
Based on the EU KLEMS Database (O’Mahony and Timmer, 2009), which provides comparable data for a large set of advanced countries over the pe-
period from 1970 to 2007, it can be observed from figure 1 that the share of manufacture in total employment and total nominal value added has been declining in advanced countries since 1970. By way of illustration, the manufacturing share of total employment decreased from 28.2% to 15.6% in the EU15 between 1970 and 2007, while it decreased from 22.4% to 9.9% in the US. Over the same period, the manufacturing share of nominal value added went from 26.6% to 18.1% in the EU15, while it went from 23.5% to 13.1% in the US. As shown in Appendix A, the UK is the country characterized by the largest absolute fall in relative manufacturing employment for the period 1970-2007, with the share of manufacturing in total employment declining from 33.2% to 11.8% (-21.4%). Partly because of the substantial growth of financial services, Luxembourg is however the country that experienced the strongest drop in the share of manufacturing in nominal value added, which fell from 41.1% to 8.6% (-32.5%).

While figure 1 reveals a deindustrialization process in advanced countries over the last few decades, it remains to be seen whether the decline in the manufacturing share of total employment and nominal value added has reflected an absolute decline in manufacturing employment and nominal value added. As shown in table 1, the answer depends on the variable considered. Thus, table 1 evidences that value added grew at a positive average annual growth rate between 1970 and 2007 in all advanced countries. While the average annual growth rate of real value added in manufacturing is lower than the one in the whole economy for the EU15, Australia and the US, it is remarkably higher for Japan, as well as for some individual European countries like Austria, Belgium, Finland, Ireland, Italy or Sweden (see Appendix B), thus reflecting the crucial role of prices in the evolution of the manufacturing share of nominal value added over time. Manufacturing employment contrastingly fell in absolute terms in the large majority of advanced countries between 1970 and 2007. The EU15, where the number of manufacturing workers declined by around 1% per year on average, exhibits a relatively large decline, surpassing Australia, Japan and the US. As shown in appendix B, this figure however hides a high degree of heterogeneity among European countries. Thus, while the number of manufacturing workers fell by around 0.2% and
0.5% per year on average in Italy and Finland respectively, it fell by 1.8% and 2.5% per year on average in Belgium and the UK respectively. Accordingly, manufacturing employment was slashed by roughly half in Belgium between 1970 and 2007 (-48.4%) and more than halved in the UK (-60.7%). The heterogeneity among advanced countries actually results from differences in both
the timing and the intensity of deindustrialization, which depend on a range of factors, such as the level of economic development and the performance of national manufacturing on global markets.

Table 1: Average annual growth rate of real value added and employment, 1970-2007 (in %)

<table>
<thead>
<tr>
<th></th>
<th>Real value added</th>
<th>Employment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Economy</td>
<td>Manuf.</td>
</tr>
<tr>
<td>EU15</td>
<td>2.5</td>
<td>1.9</td>
</tr>
<tr>
<td>United States</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td>Australia</td>
<td>3.3</td>
<td>1.5</td>
</tr>
<tr>
<td>Japan</td>
<td>3.2</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Source: EU KLEMS Database (O’Mahony and Timmer, 2009), author’s calculations.

3 Existing Explanations for Deindustrialization

Literature on structural change has identified a number of channels through which the reallocation of economic activity and, hence, deindustrialization can potentially take place in market economies\(^1\). In particular, research highlights four primary sources of deindustrialization: (i) non-homothetic tastes, (ii) technology, (iii) outsourcing and (iv) international trade. As this empirical study seeks to investigate the economic forces behind the fall in the manufacturing share of national employment in advanced countries, we will first and foremost focus on the effects of the driving forces on relative manufacturing employment.

Economists have long recognized the association between the level of economic development, i.e. the economic maturity, and the sectoral structure of national employment. Based on an extrapolation of the Engel’s law, Clark (1940) even predicted deindustrialization before it became a reality. Exploiting cross-national data, Clark (1940) argues that the fraction of income spent

\(^1\)See Krüger (2008), Matsuyama (2008), Herrendorf et al. (2014) and van Neuss (2017) for recent surveys of the structural change literature.
on manufactures should rise during the first stages of economic development, then stabilize and eventually fall beyond a threshold of per capita income, a situation sometimes labelled as the Bell’s (1976) law. According to this view, deindustrialization is driven by changes in the sectoral structure of demand resulting from changes in real income. In more technical terms, structural change is driven by non-homothetic preferences generating non-linear Engel curves (Kongsamut et al., 2001). Under the assumption that productivity growth is the same across the broad sectors agriculture, manufacturing, and services, the sequence of changes in the composition of demand then theoretically induces a curvilinear, inverted U-shaped relationship between income per capita and relative manufacturing employment.

A second source of structural change identified by research is technological heterogeneity across sectors. In a pioneer work, Baumol (1967) early formulates the “cost disease hypothesis” according to which economic resources, especially labor, move from the “dynamic” sectors, i.e. those characterized by a relatively high rate of productivity growth, to the “stagnant” sectors, i.e. those characterized by a relatively low rate of productivity growth, provided that the demand in the stagnant sectors is inelastic. Structural change is then the result of relative (sectoral) price effects driven by cross-sector differences in productivity growth (Ngai and Pissarides, 2007). Using the EU Klems Database (O’Mahony and Timmer, 2009) that provides statistics on employment and both nominal and real value added for most advanced countries, thus allowing the computation of an apparent labor productivity index and a price index (deflator), table 2 shows information on the growth of apparent labor productivity and prices in manufacturing and services in advanced countries over the period from 1970 to 2007. The results are consistent with a process of reallocation of labor from the most technologically dynamic sector, namely manufacturing, towards services as a result of a fall in the relative price of manufactures.

Besides the studies analyzing the final demand channels through which deindustrialization can take place, a number of scholars stress the importance of taking into consideration intermediates, as firms producing final goods are in turn “consumer” of goods as intermediates (e.g. Pasinetti, 1981; Fixler and
Table 2: Average annual growth rate of labor productivity and prices, 1970-2007 (in %)

<table>
<thead>
<tr>
<th></th>
<th>Labor Productivity</th>
<th>Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Manuf. Services</td>
<td>Diff.</td>
</tr>
<tr>
<td>EU15</td>
<td>3.3</td>
<td>1.5</td>
</tr>
<tr>
<td>United States</td>
<td>3.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Australia</td>
<td>2.4</td>
<td>1.3</td>
</tr>
<tr>
<td>Japan</td>
<td>4.9</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Source: EU KLEMS Database (O’Mahony and Timmer, 2009), author’s calculations.

Siegel, 1999; Berlingieri, 2014). Thus, Berlingieri (2014) points out increased “outsourcing” of services, a phenomenon reflected in the large growth of services for which final demand plays only a relatively small role, like professional and business services, as a significant contributor to the increase in services employment and to the fall in manufacturing\(^2\). Over recent decades, firms, including manufacturing firms, have indeed increasingly attempted to specialize in their “core” competencies, outsourcing non-core activities to specialized service providers. Some part of deindustrialization may then be explained by production processes becoming more dependent on service intermediates, not the least within the manufacturing sector (e.g. Falk and Peng, 2013; Lind, 2014).

A fourth potential source of deindustrialization is globalization and trade. According to traditional trade theory, opening up the borders brings about some reallocation of activity across sectors within every country involved in international exchanges. This sectoral reallocation largely takes place as a result of comparative advantage(s) driven by, among others, technology and factor endowments. A country’s economic structure is therefore directly affected by the specialization patterns induced by trade.

\(^2\)Facilitated by the introduction and progressive diffusion of information and communication technologies in the whole economy, the outsourcing strategy has been spurred by rising globalization (e.g., Windrum and Tomlinson, 1999; Franke and Kalmbach, 2005), the increasing complexity of business operations (e.g., Holweg and Pil, 2012; Berlingieri, 2015), the rising cost of monitoring workers, and the fast rate of technological progress in ever more specialized service activities.
The deindustrialization process observed in advanced countries has actually coincided with an expansion of economic linkages between the North and the South and a rapid growth of manufacturing in many parts of what was originally the developing world, particularly in China and the rest of East Asia (Weiss and Tribe, 2016). As a consequence, global integration between rich and poor countries has commonly been blamed for the drop in relative manufacturing employment in advanced countries, what Krugman (1996) has called the “deindustrialization hypothesis”. For instance, Autor et al. (2013) conclude from the main specification of their Ricardian model that growing Chinese import competition may explain one-quarter of the aggregate drop in US manufacturing workforce between 1990 and 2007. The model of Autor et al. (2013) interestingly shows that international trade can influence the allocation of labor between manufacturing and services if trade in manufactures is imbalanced. Accordingly, the severe deterioration of the manufacturing trade balances recorded by countries like the UK and the USA since the 1980s, an evolution which has coincided with booming manufactured imports from the developing world, in particular from Asia, may have contributed significantly to the fall of relative manufacturing workforce in these countries. Intuitively, another way to break the symmetry between the job-content of manufactured imports and the job-content of manufactures exports is to consider a right version of the Heckscher-Ohlin or factor-specific model, the latter assuming heterogeneous sectoral production functions in terms of factor intensities. Such trade model indeed allows for a country’s comparative advantage(s) to be also driven by national disparities in relative factor endowments. The job-content of manufactured exports can then be different from the job-content of manufactured imports even if trade in manufactures is balanced. In line with this view, using data on OECD countries for the period 1978-1995 to estimate the labor content embodied in the variations in manufacturing output resulting from changing patterns of trade, some scholars like Wood (1994, 1995) and Kucera and Milberg (2003) find that the changes in manufactured trade has had a significant negative net effect on manufacturing workforce in advanced countries, with the labor-intensive manufacturing sectors having experienced comparatively large job losses due to growing North-South trade.
4 Review of the Empirical Literature

There seems to be an increasing consensus on the idea that deindustrialization in advanced countries is the result of a combination of the forces described in the previous section. But how important is each of these forces as a source of deindustrialization? In addition to being crucial for understanding the process of deindustrialization, the answer to this question has important implications, for instance in terms of public policies and growth perspectives (Herrendorf et al., 2013). Yet the relative importance of the driving forces, especially of the “internal” factors and “external” factors, i.e. those associated with globalization, is far from being consensual in the empirical literature. The goal of this section is to provide a brief critical overview of the empirical literature which has been devoted to evaluating the impact of globalization and trade on deindustrialization in advanced countries.

The first scientific studies on the role of globalization and trade in deindustrialization are essentially based on the fears that first emerged in the US during the late 1970s and early 1980s, when firms from countries such as Japan and Germany were considered as the main competitive threat for US manufacturers. As a result, most of the pioneer studies exploring the relationship between globalization and deindustrialization focus on the impact of intra OECD trade, and largely conclude that it is of only limited importance. Works typical of this approach are the ones by Lawrence (1983, 1987 and 1991). First assimilating the process of deindustrialization to a “myth” in the United States because US manufacturing employment increased in absolute terms until the late 1970s (see Lawrence, 1983), Lawrence (1987, 1991) then argues that US deindustrialization has occurred overwhelmingly as a natural outcome of economic growth, and more particularly as a result of faster relative productivity growth in manufacturing.

In the same vein, Dollar and Wolff (1993) consider that international trade has contributed only little to deindustrialization in the United States. More particularly, they argue that US manufacturing did not suffer from international trade as its share of OECD manufacturing output and exports only marginally declined between 1970 and 1987. However, looking at relative
shares of OECD manufacturing output and exports as indicators of deindustrialization may lead to misleading conclusions given the fact that deindustrialization has taken place across the OECD as a whole.

Rowthorn and Wells (1987) are pioneer in addressing the causes of deindustrialization within and econometric framework. Using panel data to investigate the experiences of 12 OECD countries between 1953 and 1978, they conduct a set of econometric regressions that include three explanatory variables: real per capita GDP, unemployment and the ratio of net manufactured exports to total GDP. They find that internationalization, through trade specialization, accounts for persistent cross-country differences in relative manufacturing employment, but that changes in manufacturing trade balances cannot account for deindustrialization within countries. The study is however particularly silent on how trade specialization is determined and additionally fails to take into account the potential for North-South global integration to contribute to the secular decline in relative manufacturing employment3.

The study by Wood (1994), relying on accounting techniques and input-output data, is one of the first empirical works to explicitly investigate the link between deindustrialization and global integration between rich and poor countries. In particular, Wood (1994) documents an inverse relationship between the gross import penetration of manufactured goods from the South and the manufacturing share of total employment. This finding - corroborated by a number of studies of the same kind (e.g. Woodall, 1994; Sachs and Shatz, 1994; Kucera and Milberg, 2003) - is interpreted by Wood (1994) as an evidence that North-South trade has contributed significantly to a large acceleration of deindustrialization in advanced countries. There is however no attempt to control for alternative explanations for the secular decline in relative manufacturing employment.


---

3In a more recent econometric study exploring the determinants of deindustrialization in European countries, Skuffic and Druzic (2016) also fail to take convincingly into account the role of North-South trade, as the authors do not distinguish manufacturing imports according to the country of origin.
years between 1963 and 1994 and including North-South trade variables into the regressions, they find that North-South trade has contributed very little to deindustrialization in advanced countries. If the study convincingly shows the role played by trade specialization in driving cross-country differences in the sectoral structure of national employment, it nonetheless provides little evidence on the relationship between globalization and deindustrialization within the countries. This is because the authors fail to incorporate the North-South trade variables in their fixed-effects specifications, restricting their use to their pooled specification, thus making it hard to evaluate the accuracy of their final conclusion that “North-South trade has had very little to do with deindustrialization” (Rowthorn and Ramaswamy, 1997, p. 7)\textsuperscript{4}.

Relying on a data set covering the years 1970, 1975, 1980, 1985 and 1990 for 23 OECD countries, Saeger (1997) explicitly addresses, within an econometric framework, the role of North-South integration in driving deindustrialization in the developed world. Using gross trade flows for the North and the South and disaggregating them into imports and exports, Saeger (1997) estimates a set of fixed-effects regressions and finds that North-South manufactured trade is significantly related to deindustrialization in the 1970s and 1980s. In particular, he finds that the imports in manufactures from the South are a statistically significant predictor of the manufacturing share of total employment in advanced countries. In the same vein, providing random effects estimates of relative manufacturing employment for 18 OECD countries for the period 1968-1992, Alderson (1999) concludes that increasing southern import penetration has been a significant force behind deindustrialization in the OECD.

Using data for 16 OECD countries for the period 1970-2002, Boulhol and Fontagné (2006) extend the estimations by Rowthorn and Ramaswami (1999) within a dynamic panel methodology including North-South trade variables. Among others, the authors find that trade with developing countries has contributed 20% on average of the drop in relative manufacturing employment

\footnote{Similarly, it is hard to trust the results delivered by the studies of Rowthorn and Coutts (2004, 2013) as the authors compute the contribution of globalization and trade to deindustrialization in advanced countries by using the estimated coefficients from their pooled specification.}
in the OECD. This result is fairly close to that obtained by Fontagné and Lorenzi (2005) with the same set of data based on a static panel methodology.

Exploring the experience of 18 OECD countries over the period 1970-2003, the econometric analysis by Kollmeyer (2009) delivers results suggesting that global trade exerts both direct and indirect effects on the sectoral employment patterns of advanced countries. Including simultaneously explanatory variables intended to capture the role of growing affluence of consumers, unbalanced productivity growth and economic globalization within a fixed-effects specification\(^5\), Kollmeyer (2009) indeed shows that international trade affects relative manufacturing employment by inducing some economic specialization (direct effects) and by fostering the growth of aggregate income and relative productivity in the manufacturing sector (indirect effects). While the direct contribution of North-South trade to deindustrialization is estimated to amount around 13% in the OECD on average, the total contribution (direct plus indirect effects) is estimated to be around 24%. As in most previous econometric studies, the contribution of North-North trade appears very small, thus suggesting that global trade mainly impact relative manufacturing workforce through rising North-South integration.

Generally speaking, the econometric approach finds some support for a significant contribution of global exchanges, in particular of rising North-South trade in manufactures, on relative manufacturing workforce in advanced countries. The size of the contribution however appears quite limited compared to the combined effects of the “internal” factors, such as rising affluence and unbalanced productivity growth, thus challenging the popular belief that the deindustrialization of OECD countries is primarily the result of outsourcing manufacturing business and jobs to low-wage economies, as well as the main findings of the accounting studies by scholar like e.g. Wood (1994, 1995), Sachs and Shatz (1994) and Kucera and Milberg (2003). Nonetheless, it is noteworthy that the econometric results vary to a greater or lesser degree, depending notably on the selected specification, time and country data coverage, as well as on the chosen estimation technique. In a number of cases,

---

\(^5\)This study is the first to test these three major explanations for deindustrialization simultaneously.
the econometric analysis even fails to detect any significant relationship between global exchanges, or more particularly the import penetration ratio from developing countries, and the drop in relative manufacturing workforce observed in advanced countries. For instance, Saeger (1997) does not find a significant effect of manufacturing imports from the South when gross trade flows with the South and the North are disaggregated into imports and exports. More surprisingly, looking at the effects of foreign direct investment inflows and outflows on the sectoral structure of total employment in a range of OECD countries as from the 1980s, Kang and Lee (2011) find a positive and significant impact of manufacturing imports from developing countries on relative manufacturing workforce, as well as a negative and significant impact of the trade balance in manufactures.

5 Empirical Methodology

5.1 Estimation strategy

In order to study the causes of deindustrialization in advanced countries and to provide a quantification of the relative importance of globalization and trade, we adopt an econometric specification close to the one of Kollmeyer (2009) and Kang and Lee (2011). In line with this strategy, we use a panel-data model in which the determinants of a country’s manufacturing share of total national employment include the variables intended to capture the four primary explanations for deindustrialization briefly discussed in section 3 \((X_{i,t})\) and a set of control variables \((Y_{i,t})\). The manufacturing share of national employment, \(EMPMAN\), is therefore of the following type:

\[
EMPMAN_{i,t} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{i,t}^k + \sum_{j=K+1}^{J} \beta_j Y_{i,t}^j + \epsilon_{i,t}
\]

(1)

While the pooled specification can be very interesting to investigate the role of different factors in shaping differences in the sectoral structure of employment across advanced countries (e.g., Rowthorn and Wells, 1987; Saeger, 1997; Rowthorn and Ramaswamy, 1997, 1999, Alderson, 1999; Rowthorn and
Coutts, 2004, 2013), the use of “error components” panel models controlling for persistent and time invariant country-specific effects is better suited to analyze the driving forces behind the process of deindustrialization in advanced countries. We adopt a fixed-effects specification, a common choice for macroeconomists (Judson and Owen, 1999), which is furthermore validated by the performance of a Haussman test that indeed fails to show that the individual effects are uncorrelated with the model’s regressors. As previous research, we include year-specific effects (time dummies) in our specification to account for unmeasured effects occurring across time but being constant across groups. The use of time dummies is validated by a formal test rejecting the null at conventional levels of significance that the coefficients for all years are jointly equal to zero. The final model, which can be estimated with a standard two-way fixed effects approach (LSDV), thus has the following form:

\[
EMPMAN_{i,t} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{i,t}^k + \sum_{j=K+1}^{J} \beta_j Y_{i,t}^j + \theta_t + \rho_i + \epsilon_{i,t}
\]  

(2)

A number of additional tests are performed to detect some potential specification problems arising from the use of panel data. These tests seemingly reveal the presence of spatial correlation (cross-sectional dependence), serial correlation, as well as group-wise heteroscedasticity. To deal with these complications, we resort to the Panel-Corrected Standard Error (PCSE) procedure originally proposed by Beck and Katz (1995) to estimate the standard errors. The PCSE procedure is actually an alternative to the widely-used Parks-Kmenta FGLS approach - first described by Parks (1967) - for fitting linear panel models when the disturbances are not assumed to be independent and identically distributed. Among others, Beck and Katz (1995) show that the full FGLS variance-covariance estimates are typically unacceptably optimistic in the presence of the type of data commonly used by many macroeconomists, that is to say 10-20 panels with 10-40 periods per panel.
5.2 Measuring the Contribution of Globalization

How should one measure the effects of globalization on relative manufacturing employment? This is a key question for any researcher interested in analyzing empirically the drivers of deindustrialization in advanced countries. Given the difficulty of identifying and then measuring the channels through which globalization may affect the sectoral structure of total employment, previous research has largely used gross trade flows in manufacturing to construct trade variables, like the trade balance in manufacturing or the import penetration ratio from developing countries, used as proxies to capture the contribution of global exchanges to deindustrialization. To facilitate international comparisons, trade variables are usually expressed as a share of total GDP. We argue that this way of doing is not appropriate to investigate the potential relationship between globalization and deindustrialization and may therefore lead to misleading results.

A simple example will help to illustrate this point. Let’s consider a simple representative economy with two goods, manufactures and services, in which the final demand for manufactures is subject to saturation, so that any expansion in manufacturing TFP only translates into higher demand for services and a reallocation of labor away from manufacturing and towards services. Under the assumption that the economy’s position in manufacturing trade remains unchanged, the entire process of reallocation of labor towards services should then theoretically be captured by “internal” factors. Nonetheless, the variables measuring manufacturing imports and exports will move over time when gross trade flows are expressed as a share of total GDP given that total GDP also includes the services production, thus implying some contribution of globalization to the decline in manufacturing employment. Conversely, the ratios of manufacturing imports and exports to total GDP may theoretically stay constant even if deep changes in manufacturing trade affect the sectoral structure of national employment.

In light of this simple example, a better way to define trade variables intended to capture the role of globalization in deindustrialization is to express gross trade flows in manufacturing as a percentage of manufacturing produc-
tion, instead of production in the whole economy. Broadly speaking, it seems quite natural to look at the share of net exports, or exports and imports separately, in manufacturing production to quantify the relative contribution of manufacturing trade to manufacturing production and, hence, to manufacturing employment. It is exactly the approach adopted by studies resorting to accounting data to assess the role of trade in driving the evolution of manufacturing employment over the last few decades in the OECD (e.g., Bonnaz et al., 1994; Sachs and Shatz, 1994; Wood, 1994 and 1995; Cortès et al., 1999; Kucera and Milberg, 2003; Demmou, 2010). To be said differently, as manufacturing activity is influenced by both internal and external factors, the relative contribution of manufacturing trade to manufacturing employment should logically be assessed by comparing gross trade flows in manufactures with manufacturing production.

Traces of this idea can be found in the recent work by Comin et al. (2015). Relying on a structural equation, which is derived from a theoretical model of structural change, to estimate the main parameters of a utility function (the elasticity of substitution between the three broad sectors agriculture, manufacturing and services, and the sectoral income elasticities), the authors regress the sectoral employment shares on relative (sectoral) prices and aggregate consumption (real income) and take into account, in line with the implications of the model, trade controls - defined as the share of net (sectoral) exports over total production in each sector - to account for the facts that some goods can be imported and exported, thus affecting the sectoral composition of employment. As noted by the authors, this particular specification of the trade controls follows from their theoretical model. To derive this result, Comin et al. (2015) note that:

\[ p^c_{i,t} C^c_{i,t} = p^c_{i,t} Y^c_{i,t} - N X^c_{i,t} \]  

Where \( p^c_{i,t} \) is the price of sectoral good \( i \), time \( t \) and country \( c \). \( C^c_{i,t} \) is the real consumption of good \( i \), time \( t \) and country \( c \). \( Y^c_{i,t} \) is the production of good \( i \), time \( t \) and country \( c \). \( N X^c_{i,t} \) denotes the nominal value of net exports (exports minus imports) in sector \( i \), time \( t \) and country \( c \). In accordance
with this equation, the total amount of labor needed to produce the amount consumed in sector \( i \) needs to be adjusted by \( \frac{NX_{i,t}^c}{p_{i,t}^c Y_{i,t}^c} \).

By contrast with previous research, we therefore decide to express trade flows in manufacturing as a percentage of manufacturing production. Manufacturing production can be measured based on the concept of value added or final output. While it may seem to some extent relevant, for consistency purposes, to take manufacturing trade flows as a share of manufacturing gross output given that trade flows are usually expressed in gross terms\(^6\), empirical studies have long normalized statistics on trade by using either sectoral or total GDP. In order to see the impact of changing the definition of the trade variables in our empirical study on the causes of deindustrialization, we will first express gross trade flows as a proportion of manufacturing GDP, thus looking at the effects of switching from total GDP to manufacturing GDP, then we will express trade flows as a share of manufacturing gross output, thus looking at the effects of switching from manufacturing GDP to manufacturing gross output.

While one may be easily tempted to believe that expressing gross trade flows in manufacturing as a percentage of total GDP or manufacturing production - either manufacturing GDP or manufacturing gross output - is not a matter for great concern, the case of Norway, for instance, perfectly illustrates the importance of distinguishing between the two alternatives and opting for the one that seems to be the most theoretically justified. Indeed, while the ratio of gross manufacturing imports to total GDP fell from 0.21 to 0.15 between 1970 and 2007 in Norway, the ratio of gross manufacturing imports to manufacturing GDP (manufacturing gross output) grew from 1.16 (0.37) to 1.74 (0.48) over the same period.

\(^6\)While recent efforts have been made to better take into consideration the growing countries’ specialization in global industrial value chains and derive statistics on trade in “value added” terms, most historical time series of economic indicators on trade are expressed in gross terms.
We use panel data for 15 OECD countries\(^7\) for the period between 1970 and 2006. The selection of countries is dictated by the availability of data during the whole period of investigation. Our sample therefore contains 555 separate observations. Our equation is of the following type:

\[
EMPMAN_{i,t} = \beta_0 + \beta_1(YCAP_{i,t}) + \beta_2(YCAP_{i,t})^2 + \beta_3(RELPROD_{i,t}) + \beta_4(OUTSOURCING_{i,t}) + \beta_5(EXPNO_{i,t}) + \beta_6(IMPNO_{i,t}) + \beta_7(EXPSO_{i,t}) + \beta_8(IMPSO_{i,t}) + \beta_9(UNEMP_{i,t}) + \beta_{10}(FIXCAP_{i,t}) + \beta_{11}t + \theta + \rho_i + \epsilon_{i,t}
\]  

The dependent variable measures the share of a country's workforce employed in the manufacturing sector. As shown in figure 2 below, this variable captures the extent to which deindustrialization has taken place in the selected advanced countries since 1970. Data on employment at industry-level are from the EU Klem's Database (ISIC Rev.3) for virtually all countries, with the exception of Canada and Norway. For these two countries, data are from the OECD's STAN (STructural ANalysis) Database (ISIC Rev.3).

The key explanatory variables are designed to capture the primary explanations for the process of deindustrialization that we briefly presented in Section 3. The first explanatory variable measures a country’s GDP per capita expressed in 1990 international dollars as reported by Maddison (2013). In order to allow for a non-monotonous relationship between real income (national affluence) and a country’s manufacturing share of employment and, hence, to test the hump-shaped relationship between national affluence and relative manufacturing employment observed in the data, we include a squared term for this variable into the different regressions.

The second explanatory variable aims at capturing the contribution of cross-sector heterogeneity in labor productivity growth (unbalanced product-

---

\(^7\)These 15 countries are Australia, Austria, Canada, Denmark, Finland, France, Germany, Italy, Japan, Norway, Portugal, Spain, Sweden, the UK and the US.
Figure 2: Percentage change in the share of manufacturing employment, 1970-2006

Source: EU KLEMS Database (O'Mahony and Timmer, 2009), STAN Database, author’s calculations

tivity growth) to the decline in relative manufacturing workforce. It measures the ratio of labor productivity in manufacturing over labor productivity in the service sector. Labor productivity in manufacturing and services is computed by use of data on sectoral employment and real value added from the OECD’s STAN Database (ISIC Rev. 3) for Canada and Norway and from the EU Klems Database (ISIC Rev. 3) for all the other countries. The indicator of employment is “total hours worked”, and not “total persons engaged”, which dramatically improves the accuracy of the productivity measure. Under the realistic assumption that the elasticity of substitution between manufacturing goods and services is relatively low, this second explanatory variable is expected to be negatively correlated with relative manufacturing employment.

The third explanatory variable is related to outsourcing. In the absence of direct and internationally comparable data on this phenomenon, outsourcing of services from manufacturing is measured based on the indicator suggested by Demmou (2010), by looking at the evolution of the ratio of intermediate consumption (IC) to gross output (GO) in the manufacturing sector. Sectoral
data on intermediates and gross output are available in the OECD’s STAN Database (ISIC Rev.3) for both Canada and Norway and in the EU Klems Database (ISIC Rev.3) for the other countries. This variable is expected to be negatively correlated with the manufacturing share of total employment.

The fourth set of explanatory variables is intended to capture the contribution of globalization to deindustrialization. While many empirical studies have used data on trade balance in manufacturing to explore the role of trade in deindustrialization, we choose to follow the approach adopted by scholars such as Saeger (1997) and Kollmeyer (2009) by disaggregating gross trade flows into imports and exports for the North and the South. This choice allows North-South global integration to be associated with deindustrialization even if North-South trade in manufactures is roughly balanced and/or stable over time, as there is strong evidence that traded manufactures between the North and the South are of different factor intensities, in particular of different labor intensity (e.g., Wood 1994, 1995; Kucera and Milberg, 2003). Consequently, the coefficient on imports from the South is expected to be significantly larger in absolute value (and of opposite sign) compared to the coefficient on exports to the South. Data on trade flows are from the OECD’s International Trade by Commodity Statistics (ITCS) Database which categorizes the world’s countries into six geographic regions, namely Africa, Asia, Central and South America, Europe, North America and Oceania. In line with the study by Kollmeyer (2009), the South is defined as Africa, Asia, Central and South America, and Oceania. The North covers Europe and North America. A number of adjustments are made to these broad regional categories by moving Mexico (North America) and Turkey (Europe) to the South, and by moving Australia and New Zealand (Oceania) and Israel, Japan, and South Korea (Asia) to the North. Trade in manufactures is defined as both imports and exports in SITC (Standard International Trade Classification) sections 5 to 8. In order to see the influence of changing the normalization of gross trade flows on the results, monetary values for manufacturing imports and exports are expressed as a share of total GDP, as it is the case in existing research, as well as a share of manufacturing GDP and manufacturing gross output.
Finally, a set of control variables is also included into the regressions. Following existing research, two main control variables are considered. The first one, i.e. the unemployment rate, aims at capturing what Rowthorn and Wells (1987) call the “failure effect” or “negative deindustrialization”. According to this argument, deindustrialization may also take place as a result of a large and persistent structural disequilibrium in the macroeconomy, which is manifested in the poor performance of manufacturing and the economy as a whole. As part of this vicious circle, the labor shed by negative deindustrialization is not absorbed by services, contributing to a rise in the unemployment rate and the share of services in employment. The second control variable measures fixed capital investment expressed as a share of GDP. As noted by Rowthorn and Ramaswamy (1999), the rationale for this variable is that fixed capital investment is relatively focused on manufacturing, that is, is manufacturing intensive. The descriptive statistics are summarised in table 3.

7 Empirical Results

The regression results are shown in table 4 below. Model 1 isolates the variables intended to capture the contribution of growing national affluence (YCAP) and unbalanced productivity growth (RELPROD) to deindustrialization. Model 2 also incorporates the explanatory variable related to outsourcing. Both models confirm some of the results of previous empirical research. In particular, they give support to the curvilinear, inverted U-shaped relationship between income per capita and relative manufacturing workforce, and confirm the negative relationship between the relative productivity of manufacturing compared to the service sector and relative manufacturing workforce. Model 2 fails to detect a significant impact of the ratio of intermediates over gross output in manufacturing on relative manufacturing workforce, thus suggesting that this variable is seemingly not well designed to capture the broad contribution of outsourcing to deindustrialization in OECD advanced countries. This variable is therefore excluded from the following models⁸.

It is noteworthy that its inclusion would not significantly change the results.
Table 3: Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMPMAN</td>
<td>Share of manufacturing in total employment</td>
<td>555</td>
<td>0.204</td>
<td>0.043</td>
<td>0.101</td>
<td>0.339</td>
</tr>
<tr>
<td>YCAP</td>
<td>Per capita GDP</td>
<td>555</td>
<td>16,313</td>
<td>4,596</td>
<td>5,472</td>
<td>31,357</td>
</tr>
<tr>
<td>RELPROD</td>
<td>Relative productivity of manufacturing vs services</td>
<td>555</td>
<td>1.410</td>
<td>0.426</td>
<td>0.808</td>
<td>3.514</td>
</tr>
<tr>
<td>OUTSOURCING</td>
<td>Ratio of intermediates to GO in manufacturing</td>
<td>555</td>
<td>0.676</td>
<td>0.033</td>
<td>0.605</td>
<td>0.781</td>
</tr>
<tr>
<td>EXP-NO-GDP</td>
<td>Exports to the North (% total GDP)</td>
<td>555</td>
<td>0.111</td>
<td>0.061</td>
<td>0.008</td>
<td>0.340</td>
</tr>
<tr>
<td>EXP-SO-GDP</td>
<td>Exports to the South (% total GDP)</td>
<td>555</td>
<td>0.026</td>
<td>0.012</td>
<td>0.006</td>
<td>0.068</td>
</tr>
<tr>
<td>IMP-NO-GDP</td>
<td>Imports from the North (% total GDP)</td>
<td>555</td>
<td>0.122</td>
<td>0.057</td>
<td>0.014</td>
<td>0.273</td>
</tr>
<tr>
<td>IMP-SO-GDP</td>
<td>Imports from the South (% total GDP)</td>
<td>555</td>
<td>0.018</td>
<td>0.009</td>
<td>0.002</td>
<td>0.055</td>
</tr>
<tr>
<td>EXP-NO-GDPMAN</td>
<td>Exports to the North (% manufacturing GDP)</td>
<td>555</td>
<td>0.628</td>
<td>0.366</td>
<td>0.052</td>
<td>1.696</td>
</tr>
<tr>
<td>EXP-SO-GDPMAN</td>
<td>Exports to the South (% manufacturing GDP)</td>
<td>555</td>
<td>0.144</td>
<td>0.063</td>
<td>0.038</td>
<td>0.416</td>
</tr>
<tr>
<td>IMP-NO-GDPMAN</td>
<td>Imports from the North (% manufacturing GDP)</td>
<td>555</td>
<td>0.726</td>
<td>0.412</td>
<td>0.053</td>
<td>1.714</td>
</tr>
<tr>
<td>IMP-SO-GDPMAN</td>
<td>Imports from the South (% manufacturing GDP)</td>
<td>555</td>
<td>0.109</td>
<td>0.077</td>
<td>0.013</td>
<td>0.496</td>
</tr>
<tr>
<td>EXP-NO-GOMAN</td>
<td>Exports to the North (% manufacturing GO)</td>
<td>555</td>
<td>0.201</td>
<td>0.119</td>
<td>0.016</td>
<td>0.553</td>
</tr>
<tr>
<td>EXP-SO-GOMAN</td>
<td>Exports to the South (% manufacturing GO)</td>
<td>555</td>
<td>0.046</td>
<td>0.020</td>
<td>0.011</td>
<td>0.117</td>
</tr>
<tr>
<td>IMP-NO-GOMAN</td>
<td>Imports from the North (% manufacturing GO)</td>
<td>555</td>
<td>0.230</td>
<td>0.128</td>
<td>0.017</td>
<td>0.554</td>
</tr>
<tr>
<td>IMP-SO-GOMAN</td>
<td>Imports from the South (% manufacturing GO)</td>
<td>555</td>
<td>0.035</td>
<td>0.024</td>
<td>0.004</td>
<td>0.145</td>
</tr>
<tr>
<td>UNEMP</td>
<td>Unemployment rate</td>
<td>555</td>
<td>0.064</td>
<td>0.038</td>
<td>0.006</td>
<td>0.242</td>
</tr>
<tr>
<td>FIXCAP</td>
<td>Domestic fixed investment as a share of GDP</td>
<td>555</td>
<td>0.241</td>
<td>0.036</td>
<td>0.175</td>
<td>0.372</td>
</tr>
</tbody>
</table>
Models 3 and 4 incorporate the variables intended to capture the broad effects of globalization on deindustrialization. These variables relate to manufacturing trade, both exports and imports, with the North and the South. While the trade variables are expressed as a percentage of total GDP in model 3, a normalization applied in previous research, they are taken as a share of manufacturing GDP in model 4 and as a share of manufacturing gross output in model 5. Unsurprisingly, the results appear different, even though the coefficients all exhibit the expected signs, meaning that the manufacturing exports to the North and the South are positively correlated with relative manufacturing employment, whereas the manufacturing imports from the North and the South are negatively correlated with relative manufacturing employment. A key difference comes from the coefficient linked to the variable measuring the imports from the South. Indeed, this coefficient is not significantly different from zero at the 0.05 level of significance in model 3. Moreover, in model 3, the p-value of the test aimed at determining whether this coefficient is significantly different from zero turns out to be very sensitive to the selected period of investigation and/or the selected set of countries. While this fragility may be partly explained with statistical elements, we argue that it largely results from an inaccurate definition of the trade variables. In model 4, when gross trade flows are expressed as a share of manufacturing GDP, the coefficients associated with the trade variables are all robustly significant at the 0.05 level of significance. Additionally, the coefficient of the variable measuring the imports from the South is estimated more precisely and is significantly higher in magnitude than the coefficients of the other trade variables, which is consistent with the theoretical expectations. One can therefore reject the restriction that solely imbalances in North-South trade are related to deindustrialization. This finding gives support to the studies by e.g. Wood (1994, 1995) and Kucera and Milberg (2003) which indeed find strong empirical evidence that traded manufactures between the North and the South are of

---

\(^9\)For instance, removing Sweden from the regression leads the p-value of this test to rise substantially, making the coefficient not significantly different from zero at the 0.1 level of significance. This fragility may potentially explain why some studies on deindustrialization fail to detect a significant impact of manufacturing imports from the South on relative manufacturing employment in advanced countries (e.g., Saeger, 1997; Kang and Lee, 2011).
differential labor intensity. It also confirms the importance of disaggregating trade flows, in particular trade flows with the South, into both exports and imports, an approach not followed by e.g. Skuffic and Druzin (2016) in their recent analysis of the determinants of deindustrialization in the EU. The results delivered by model 5, in which gross trade flows are expressed as a share of manufacturing gross output, are qualitatively similar to the results of model 4. The main difference comes from the magnitude of the coefficients linked to the trade variables. The coefficients are indeed higher in model 5, due in part to wider trade data spread when gross trade flows are taken as a share of manufacturing gross output compared to manufacturing GDP (see table 3).

Using directly the regression coefficients from models 4 and 5 along with the actual data, we now quantify the relative importance of manufacturing trade in explaining the observed deindustrialization in the selected OECD countries. The results are shown in table 5 below. Generally speaking, though the contribution of globalization to deindustrialization tends to be somewhat higher when using the coefficients from model 5, the choice of expressing gross trade flows in manufactures as a share of manufacturing GDP or as a share of manufacturing gross output does not fundamentally change the results. Based on our calculations, manufacturing trade affected negatively relative manufacturing employment, and so contributed positively to deindustrialization, in virtually all selected countries over the period between 1970 and 2006, except Finland. While the direct contribution of manufacturing trade to the decline in relative manufacturing employment is relatively small for countries such as Germany, Italy, Japan and Sweden, it is comparatively large for countries such as Australia, Canada, Portugal, Spain, the UK and the US. For information purposes, our empirical estimates exceed those produced by Kollmeyer (2009, p. 1666) for G-7, namely Canada, France, Germany, Italy, Japan, the UK and the US. By way of illustration, the results of Kollmeyer (2009) show that the direct contribution of trade to deindustrialization amounts to 18.9% and 16.7% for the UK and the US respectively over the period 1970-2003.
### Table 4: FE-PCSE estimates of relative manufacturing employment (log) for 15 OECD countries, 1970-2006

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>YCAP</td>
<td>2.853***</td>
<td>2.877***</td>
<td>3.219***</td>
<td>4.887***</td>
<td>4.639***</td>
</tr>
<tr>
<td>(YCAP)^2</td>
<td>-1.134***</td>
<td>-1.135***</td>
<td>-1.152***</td>
<td>-1.241***</td>
<td>-1.229***</td>
</tr>
<tr>
<td>RELPROD</td>
<td>-.114***</td>
<td>-.111***</td>
<td>-.117***</td>
<td>-.075***</td>
<td>-.063***</td>
</tr>
<tr>
<td>OUTSOURCING</td>
<td>.122 (1.14)</td>
<td>.901 (1.19)***</td>
<td>-.308 (1.21)**</td>
<td>.1088 (2.38)***</td>
<td>-.769 (2.40)*</td>
</tr>
<tr>
<td>EXP-NO-GDP</td>
<td>-.010 (0.015)</td>
<td>.386 (0.076)***</td>
<td>-.013 (0.075)***</td>
<td>.574 (0.163)***</td>
<td>-.1695 (0.200)***</td>
</tr>
<tr>
<td>IMP-NO-GDPMAN</td>
<td>-.007 (0.005)</td>
<td>-.007 (0.005)</td>
<td>-.008 (0.005)</td>
<td>-.014 (0.006)**</td>
<td>-.016 (0.006)**</td>
</tr>
<tr>
<td>EXP-SO-GDPMAN</td>
<td>.1010 (0.020)</td>
<td>.1008 (0.020)</td>
<td>.072 (0.024)***</td>
<td>.109 (0.028)***</td>
<td>.117 (0.027)***</td>
</tr>
<tr>
<td>UNEMP</td>
<td>-.007 (0.005)</td>
<td>-.007 (0.005)</td>
<td>-.008 (0.005)</td>
<td>-.014 (0.006)**</td>
<td>-.016 (0.006)**</td>
</tr>
<tr>
<td>FIXCAP</td>
<td>.110 (0.020)</td>
<td>.088 (0.020)</td>
<td>.072 (0.024)***</td>
<td>.109 (0.028)***</td>
<td>.117 (0.027)***</td>
</tr>
<tr>
<td>Observations</td>
<td>555</td>
<td>555</td>
<td>555</td>
<td>555</td>
<td>555</td>
</tr>
</tbody>
</table>

All regressions include country dummies, time dummies and a linear trend. Variables are described in Table 3. The variables YCAP, UBP, UNEMP and FIXCAP are turned into logarithms. Standard errors, between parentheses, are robust to heteroscedasticity and autocorrelation and computed from Beck and Katz (1995) - see Stata’s xtpcse command. ***, ** and * denote significance at the 1, 5 and 10 per cent levels, respectively.
Table 5: Direct contribution of manufacturing trade to deindustrialization, 1970-2006 (in % of total change)

<table>
<thead>
<tr>
<th>Country</th>
<th>Model 4</th>
<th></th>
<th></th>
<th>Model 5</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>North-North</td>
<td>North-South</td>
<td>Total</td>
<td>North-North</td>
<td>North-South</td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>Trade</td>
<td>Trade</td>
<td>Trade</td>
<td>Trade</td>
<td>Trade</td>
<td>Trade</td>
</tr>
<tr>
<td>Australia</td>
<td>12.7</td>
<td>34.6</td>
<td>47.3</td>
<td>10.1</td>
<td>36.3</td>
<td>46.3</td>
</tr>
<tr>
<td>Austria</td>
<td>-1.4</td>
<td>10.1</td>
<td>8.6</td>
<td>4.4</td>
<td>12.8</td>
<td>17.2</td>
</tr>
<tr>
<td>Canada</td>
<td>3.4</td>
<td>33.1</td>
<td>36.4</td>
<td>5.6</td>
<td>34.9</td>
<td>40.5</td>
</tr>
<tr>
<td>Denmark</td>
<td>-2.6</td>
<td>13.6</td>
<td>11.0</td>
<td>0.7</td>
<td>16.6</td>
<td>17.3</td>
</tr>
<tr>
<td>Finland</td>
<td>-27.2</td>
<td>16.0</td>
<td>-11.2</td>
<td>-23.6</td>
<td>19.8</td>
<td>-3.8</td>
</tr>
<tr>
<td>France</td>
<td>8.9</td>
<td>12.9</td>
<td>21.8</td>
<td>11.4</td>
<td>11.7</td>
<td>23.1</td>
</tr>
<tr>
<td>Germany</td>
<td>-8.7</td>
<td>12.6</td>
<td>4.0</td>
<td>-0.8</td>
<td>15.5</td>
<td>14.6</td>
</tr>
<tr>
<td>Italy</td>
<td>3.2</td>
<td>12.7</td>
<td>15.9</td>
<td>6.7</td>
<td>12.6</td>
<td>19.4</td>
</tr>
<tr>
<td>Japan</td>
<td>-3.2</td>
<td>14.3</td>
<td>11.1</td>
<td>-2.1</td>
<td>18.8</td>
<td>16.8</td>
</tr>
<tr>
<td>Norway</td>
<td>14.6</td>
<td>15.0</td>
<td>29.6</td>
<td>10.7</td>
<td>14.5</td>
<td>25.3</td>
</tr>
<tr>
<td>Portugal</td>
<td>28.6</td>
<td>7.0</td>
<td>35.6</td>
<td>34.4</td>
<td>6.6</td>
<td>41.0</td>
</tr>
<tr>
<td>Spain</td>
<td>17.7</td>
<td>29.3</td>
<td>47.1</td>
<td>19.4</td>
<td>27.3</td>
<td>46.7</td>
</tr>
<tr>
<td>Sweden</td>
<td>-2.5</td>
<td>3.0</td>
<td>0.5</td>
<td>-1.1</td>
<td>3.7</td>
<td>2.6</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>10.2</td>
<td>24.1</td>
<td>34.3</td>
<td>14.9</td>
<td>28.4</td>
<td>43.2</td>
</tr>
<tr>
<td>United States</td>
<td>5.6</td>
<td>25.7</td>
<td>31.3</td>
<td>7.0</td>
<td>31.2</td>
<td>38.2</td>
</tr>
</tbody>
</table>
While the estimations presented in table 5 do not contradict the widely accepted idea that the “internal” factors are quantitatively more important in accounting for deindustrialization in the OECD as a whole, they suggest that trade in manufactures may have played an important role in accelerating the decline of relative manufacturing employment in a number of OECD countries, especially as trade has also affected manufacturing workforce in an indirect manner by enhancing the growth of GDP per capita and relative productivity in the manufacturing sector. With respect to this point, Kollmeyer (2009) shows that the total contribution of trade, in particular of trade with developing countries, may be doubled when taking into consideration the indirect effects of trade on income and productivity. It is noteworthy that the total impact of globalization on relative manufacturing employment may also be revised upwards when considering explicitly the role of foreign direct investment (e.g., Alderson, 1997, 1999; Kang and Lee, 2011) and offshore outsourcing.

In reality, it is very difficult to compare our results with those of previous research as the confidence intervals are usually not reported in those studies. Thus, for instance, while the use of the parameters from models 4 and 5 truly leads the direct contribution of manufacturing trade to the decline in relative manufacturing workforce to be greater than that suggested by the results of Kollmeyer (2009) for the group of G-7 countries, we do not know whether our estimates are significantly different. In Appendix C, we compute all the 95% confidence intervals associated with the estimated contributions of North-North trade and North-South trade to deindustrialization in our empirical exercise. Finland appears to be the only country where manufacturing trade with the North contributed significantly to a growth of relative manufacturing workforce. While the import penetration ratio in manufactures from the North rose only slightly in Finland over the period 1970-2006, manufacturing exports to the North increased substantially. By contrast, North-North trade is significantly related to deindustrialization in countries...
such as Australia, Norway, the UK and the US. Quite interestingly, this list brings out countries whose manufacturing activity is believed to have suffered from the “Dutch disease”\textsuperscript{11}. With respect to North-South trade, our estimations suggest that growing global integration between rich and poor countries has significantly affected the process of deindustrialization in virtually all selected OECD countries over recent decades, with the exception of Sweden where the import penetration ratio in manufactures from the South was still relatively low in the 2000s. This situation contrasts with the very rapid growth of manufacturing imports from the South in countries such as the UK and the US, an evolution that has fed numerous fears concerning the impact of North-South trade on manufacturing activity in advanced countries.

8 Endogeneity of Trade Flows

Like existing econometric research on the drivers of deindustrialization, this paper uses gross trade flows in manufactures to measure the effects of globalization and increasing North-South integration on manufacturing workforce. There are at least two main criticisms which can be addressed against such a strategy. First, there is little theoretical rationale for measuring the impact of globalization and trade by use of a quantity variable. Second, manufacturing trade flows are likely to be not exogenous, as they are determined simultaneously with both manufacturing employment and production.

From an empirical point of view, the endogeneity of manufacturing trade flows poses a problem for isolating the effects of globalization and trade. So far, this econometric study therefore suffers from the same limitations than most previous research. In particular, it cannot establish a causal relationship between globalization and deindustrialization. The partial correlations obtained from the econometric regressions in table 4 however provide a strong

\textsuperscript{11}The Dutch disease is a version of the “resource curse” theory that establishes a causal relationship between booming exports in natural resources and/or primary goods, leading to an appreciating real exchange rate, and deindustrialization. Though the term “Dutch disease” originally refers to the supposedly negative effects on manufacturing of the discovery of natural resources, like mineral resources, gas and oil, it has been extended to describe deindustrialization arising from a substantial growth in the exports of any sector other than manufacturing, like e.g. finance (Luxembourg, the UK) and tourism (Greece).
evidence on the existence of a close link between globalization and deindustrialization, and can arguably be used to make some “reasonable” decomposition of the contraction of relative manufacturing employment in advanced countries over recent decades.

Although this issue has been largely overlooked in the empirical literature, a few empirical studies on deindustrialization have explicitly attempted to deal with the potential endogeneity of trade flows by using the instrumental variables method (e.g. Bouhlool and Fontagné, 2006; Demmou, 2010; Kang and Lee, 2011; Autor et al., 2013). Following Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998), most of these studies resort to a GMM framework, which allows to use the past values of the endogenous trade variables to construct instrumental variables. For instance, Boulhol and Fontagné (2006) and Demmou (2010) apply the GMM methodology developed by Arellano and Bond (1991) and use lags of the trade variables as instruments. The resulting estimator is called “GMM difference”. As a robustness test, we now propose to check the validity of the results presented in table 4 by adopting the same strategy.

The GMM approach begins with the first-differenced version of equation (4), which leads to the disappearance of the country-specific effects. However, it is well known that the independent variable with some weakly exogenous properties retains some endogeneity problems even after the first-differencing process. We therefore need to differentiate the equation to remove the fixed effects, then estimate with instrumental variables by using as the instruments the values which are lagged by two or more periods (Roodman, 2009). In order to limit the number of instruments, which increases rapidly with the number of periods, we use the second to fourth lags of the trade variables. We run our main model for different definitions of the manufacturing trade variables. While the trade variables are expressed as a share of total GDP in model 6, they are taken as a share of manufacturing GDP in model 7 and as a share of manufacturing gross output in model 8 (table 6). The coefficients associated with the trade variables all exhibit the expected signs, meaning that the manufacturing exports to the North and the South are positively correlated with relative manufacturing workforce, whereas the manufacturing
imports from the North and the South are negatively correlated with relative manufacturing workforce. In model 6, the coefficient of the imports from the South is significantly different from 0 at the 0.05 level of significance (p-value = 0.047), but does not appear to be significantly different from the coefficients of the other trade variables. In models 7 and 8, the coefficient of the imports from the South is significantly different from 0 at the 0.01 level of significance and, moreover, is significantly higher in absolute value than the coefficients of the other trade variables, a result in line with theoretical expectations.

GMM estimation that is solely based on the methodology developed by Arellano and Bond (1991) can be remarkably inefficient. Indeed, a well-known problem with the original Arellano-Bond estimator is that lagged levels often prove to be poor instruments for first differences. Arellano and Bover (1995) and Blundell and Bond (1998) show how a number of additional moment conditions can be brought to bear to increase efficiency when adding original equations in levels to the system. In these equations, predetermined and/or endogenous variables in levels are instrumented with the suitable lags of their own first differences (Roodman, 2009). This estimator is called “system GMM”. We use this estimator in models 9 to 11 (table 7). While the results are similar to those derived with the “difference GMM” estimator, the estimations appear more precise.

Broadly speaking, the GMM methodology that we applied in this section to deal with the potential endogeneity of trade flows confirm the econometric results presented in table 4. The definition of the trade variables does have a large influence on the econometric estimates and, hence, on the computed contribution of globalization to deindustrialization. In particular, the choice of utilizing manufacturing gross trade flows as a percentage of total production or manufacturing production - either manufacturing value added or manufacturing gross output - really matters. With respect to this point, taking manufacturing gross trade flows as a share of manufacturing production, instead of total production, seemingly leads to more precise estimates and, furthermore, corroborates the well-documented empirical evidence that manufacturing imports from the South are of different labor intensity than manufacturing imports from the North and manufacturing exports.

158
Table 6: Difference GMM estimates of relative manufacturing employment (log) for 15 OECD countries, 1970-2006

<table>
<thead>
<tr>
<th></th>
<th>Model 6</th>
<th>Model 7</th>
<th>Model 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>YCAP</td>
<td>5.024</td>
<td>7.862</td>
<td>7.124</td>
</tr>
<tr>
<td></td>
<td>(1.243)***</td>
<td>(.784)***</td>
<td>(.708)***</td>
</tr>
<tr>
<td>(YCAP)^2</td>
<td>-2.45</td>
<td>-3.391</td>
<td>-3.354</td>
</tr>
<tr>
<td></td>
<td>(.067)***</td>
<td>(.042)***</td>
<td>(.037)***</td>
</tr>
<tr>
<td>RELPROD</td>
<td>-0.061</td>
<td>-0.136</td>
<td>-0.127</td>
</tr>
<tr>
<td></td>
<td>(.039)</td>
<td>(.025)***</td>
<td>(.023)***</td>
</tr>
<tr>
<td>EXP-NO-GDP</td>
<td>1.374</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.333)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMP-NO-GDP</td>
<td>-1.075</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.458)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXP-SO-GDP</td>
<td>3.151</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.789)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMP-SO-GDP</td>
<td>-2.064</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.037)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXP-NO-GDPMAN</td>
<td></td>
<td>.288</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.037)***</td>
<td></td>
</tr>
<tr>
<td>IMP-NO-GDPMAN</td>
<td></td>
<td>-.307</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.041)***</td>
<td></td>
</tr>
<tr>
<td>EXP-SO-GDPMAN</td>
<td></td>
<td>.328</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.094)***</td>
<td></td>
</tr>
<tr>
<td>IMP-SO-GDPMAN</td>
<td></td>
<td>-.651</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.088)***</td>
<td></td>
</tr>
<tr>
<td>EXP-NO-GOMAN</td>
<td></td>
<td></td>
<td>.947</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.111)***</td>
</tr>
<tr>
<td>IMP-NO-GOMAN</td>
<td></td>
<td>-1.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.127)***</td>
<td></td>
</tr>
<tr>
<td>EXP-SO-GOMAN</td>
<td></td>
<td>1.143</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.275)***</td>
<td></td>
</tr>
<tr>
<td>IMP-SO-GOMAN</td>
<td></td>
<td>-2.390</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.264)***</td>
<td></td>
</tr>
<tr>
<td>UNEMP</td>
<td>.004</td>
<td>-.010</td>
<td>-.030</td>
</tr>
<tr>
<td></td>
<td>(.019)</td>
<td>(.011)</td>
<td>(.011) *</td>
</tr>
<tr>
<td>FIXCAP</td>
<td>.106</td>
<td>.179</td>
<td>.190</td>
</tr>
<tr>
<td></td>
<td>(.074)</td>
<td>(.042)***</td>
<td>(.039)***</td>
</tr>
<tr>
<td>AR(1)p</td>
<td>.135</td>
<td>.647</td>
<td>.536</td>
</tr>
<tr>
<td></td>
<td>(.074)</td>
<td>(.042)***</td>
<td>(.039)***</td>
</tr>
<tr>
<td>AR(2)p</td>
<td>.358</td>
<td>.302</td>
<td>.211</td>
</tr>
<tr>
<td>Sargan-Hansen p</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Observations</td>
<td>555</td>
<td>555</td>
<td>555</td>
</tr>
</tbody>
</table>

All regressions include time dummies and a linear trend. Variables are described in Table 3. The variables YCAP, UBP, UNEMP and FIXCAP are turned into logarithms. Asymptotic standard errors, between parentheses, are robust to heteroscedasticity and autocorrelation - see Stata’s xtdpd command. ***, ** and * denote significance at the 1, 5 and 10 per cent levels, respectively. Serial correlation statistics are P-values for Arellano-Bond tests for first- and second-order correlation. For the Sargan-Hansen test, the number reported is the P-value.
Table 7: System GMM estimates of relative manufacturing employment (log) for 15 OECD countries, 1970-2006

<table>
<thead>
<tr>
<th></th>
<th>Model 9</th>
<th>Model 10</th>
<th>Model 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>YCAP</td>
<td>6.958</td>
<td>7.429</td>
<td>6.855</td>
</tr>
<tr>
<td></td>
<td>(0.302)***</td>
<td>(0.182)***</td>
<td>(0.190)***</td>
</tr>
<tr>
<td>(YCAP)^2</td>
<td>-0.379</td>
<td>-0.401</td>
<td>-0.360</td>
</tr>
<tr>
<td></td>
<td>(0.016)***</td>
<td>(0.009)***</td>
<td>(0.010)***</td>
</tr>
<tr>
<td>RELPROD</td>
<td>-0.082</td>
<td>-0.155</td>
<td>-0.147</td>
</tr>
<tr>
<td></td>
<td>(0.010)***</td>
<td>(0.006)***</td>
<td>(0.007)***</td>
</tr>
<tr>
<td>EXP-NO-GDP</td>
<td>1.537</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMP-NO-GDP</td>
<td>-1.385</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.071)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXP-NO-GDPMAN</td>
<td></td>
<td>.241</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.006)***</td>
<td></td>
</tr>
<tr>
<td>IMP-NO-GDPMAN</td>
<td>-0.318</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXP-NO-GOMAN</td>
<td>.225</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMP-NO-GODP</td>
<td>-1.784</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.254)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXP-NO-GDPMAN</td>
<td></td>
<td>.917</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.022)***</td>
<td></td>
</tr>
<tr>
<td>IMP-NO-GODP</td>
<td>-1.137</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.080)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RELPROD</td>
<td>-0.026</td>
<td>-0.042</td>
<td>-0.044</td>
</tr>
<tr>
<td></td>
<td>(0.005)***</td>
<td>(0.003)***</td>
<td>(0.003)***</td>
</tr>
<tr>
<td>EXP-NO-GDPMAN</td>
<td>.071</td>
<td>.039</td>
<td>.058</td>
</tr>
<tr>
<td></td>
<td>(0.018)***</td>
<td>(0.011)***</td>
<td>(0.012)***</td>
</tr>
<tr>
<td>IMP-NO-GDPMAN</td>
<td>.387</td>
<td>.354</td>
<td>.249</td>
</tr>
<tr>
<td></td>
<td>.387</td>
<td>.354</td>
<td>.249</td>
</tr>
<tr>
<td>Observations</td>
<td>555</td>
<td>555</td>
<td>555</td>
</tr>
</tbody>
</table>

All regressions include time dummies and a linear trend. Variables are described in Table 3. The variables YCAP, UBP, UNEMP and FIXCAP are turned into logarithms. Asymptotic standard errors, between parentheses, are robust to heteroscedasticity and autocorrelation - see Stata's xtdp command. ***, ** and * denote significance at the 1, 5 and 10 per cent levels, respectively. Serial correlation statistics are P-values for Arellano-Bond tests for first- and second-order correlation. For the Sargan-Hansen test, the number reported is the P-value.
9 Conclusion

Deindustrialization is one of the best-established stylized facts associated with economic development in advanced economies. Yet the causes and consequences of deindustrialization are still far from being fully understood. This contributes to fueling an abundant political discussion and scientific research, especially as there is a longstanding tradition in economics which argues that the manufacturing sector has a critical role in economic performance and growth (e.g., Szirmai, 2012; Weiss and Jalilian, 2016). This paper joins the empirical literature focusing on the determinants of deindustrialization. A strand of this literature aims at quantifying the relative importance of the economic forces driving the decline in relative manufacturing employment. Understanding the role played by each of these forces is indeed essential to ensure the appropriate policy response (OECD, 2017).

Relying on a regression-based decomposition of the evolution of relative manufacturing workforce in a set of advanced OECD countries over recent decades, this study explores more particularly the relationship between globalization and deindustrialization. While it does not necessarily contradict the widely accepted idea that internal factors are quantitatively more important in accounting for deindustrialization in the OECD taken as a whole, it nonetheless suggests that global exchanges have the potential to affect significantly and substantially a country’s sectoral patterns of employment, and that the contribution of trade, especially of trade with developing countries, to the observed deindustrialization in advanced countries may be revised upwards when using better-defined indicators of manufacturing trade. Regarding this point, the growing development of data on trade flows expressed in “value added” terms, not in gross terms, is likely to improve dramatically the empirical results in the future.

The role of international competition in structural change, especially in recent deindustrialization, has a number of important implications in terms of public policies in advanced countries. In particular, policy effort should be made to avoid the risk of loss of industrial substance in a competitive world economy, notably by helping firms to achieve successful participation
in increasingly globalized value chains. This implies defining and implement-
ing an effective industrial (supply-side) policy which is conducive to overall
economic development. In line with this view, Crafts (1996) early argued
in favor of a supply-side policy that would not place too much focus on the
deindustrialization of workforce per se, viewed as a “distraction”, but would
rather address the substantial questions over the human capital formation
and technological capabilities, as well as over investment attractiveness, de-

dined as a country’s ability to attract foreign direct investment (FDI) and to
benefit from knowledge available from abroad.

As trade integration inevitably induces a reallocation of resources both
within and between sectors, mainly as a result of comparative advantage(s),
policy effort should also be made to mitigate the costs and redistributive ef-
fects associated with international specialization, especially in regions which
are suffering from severe economic distortions due to tough import competi-
tion. The literature is full of examples of regions where trade liberalization
is associated with a large erosion of the manufacturing base and relative
economic decline, especially as the low geographical and sectoral mobility
of workers often prevent local economies from adjusting optimally to shocks
(OECD, 2017). With respect to this point, it is worth noting that labour
market policy crucially matters as a determinant of the speed of structural
change and the ease with which labor moves from one sector to another. Ac-
cordingly, Thompson et al. (2012) contend that the potential gains from trade
and increased specialization are directly linked to the degree of flexibility of
the economy, in particular of the labor market.\footnote{Some related research shows that the effects of international trade on unemployment are
highly dependent on labor market institutions, i.e. on differences in labor market frictions
across industries and countries (e.g. Stijepic and Wagner, 2008; Helpman and Itskhoki,
2010; Fugazza et al., 2014). See e.g. Belenkiy and Riker (2015) for a recent survey of the
literature linking international trade to unemployment.}
10 Appendix

10.1 Appendix A: Manufacturing Share of Nominal Value Added and Employment in Advanced Countries

Table 8: Share of Manufacturing in Nominal Value Added (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>24.6</td>
<td>19.8</td>
<td>14.5</td>
<td>12.7</td>
<td>10.5</td>
<td>-14.0</td>
</tr>
<tr>
<td>Austria</td>
<td>27.9</td>
<td>23.7</td>
<td>21.4</td>
<td>20.3</td>
<td>20.0</td>
<td>-7.9</td>
</tr>
<tr>
<td>Belgium</td>
<td>29.4</td>
<td>23.1</td>
<td>22.8</td>
<td>19.3</td>
<td>16.4</td>
<td>-13.0</td>
</tr>
<tr>
<td>Denmark</td>
<td>20.5</td>
<td>18.9</td>
<td>17.4</td>
<td>16.2</td>
<td>15.0</td>
<td>-5.5</td>
</tr>
<tr>
<td>Finland</td>
<td>26.2</td>
<td>27.5</td>
<td>22.6</td>
<td>26.2</td>
<td>23.6</td>
<td>-2.6</td>
</tr>
<tr>
<td>France</td>
<td>24.4</td>
<td>21.7</td>
<td>18.3</td>
<td>16.0</td>
<td>12.3</td>
<td>-12.1</td>
</tr>
<tr>
<td>Germany</td>
<td>34.9</td>
<td>29.7</td>
<td>28.1</td>
<td>22.9</td>
<td>23.6</td>
<td>-11.3</td>
</tr>
<tr>
<td>Greece</td>
<td>23.1</td>
<td>22.2</td>
<td>18.1</td>
<td>12.0</td>
<td>11.1</td>
<td>-11.9</td>
</tr>
<tr>
<td>Ireland</td>
<td>22.1</td>
<td>21.6</td>
<td>27.7</td>
<td>33.6</td>
<td>22.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Italy</td>
<td>27.6</td>
<td>28.9</td>
<td>23.3</td>
<td>21.0</td>
<td>19.0</td>
<td>-8.5</td>
</tr>
<tr>
<td>Japan</td>
<td>33.5</td>
<td>27.2</td>
<td>26.1</td>
<td>21.3</td>
<td>20.6</td>
<td>-12.9</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>41.1</td>
<td>25.4</td>
<td>20.7</td>
<td>11.3</td>
<td>8.6</td>
<td>-32.5</td>
</tr>
<tr>
<td>Netherlands</td>
<td>25.3</td>
<td>18.1</td>
<td>18.6</td>
<td>15.6</td>
<td>13.7</td>
<td>-11.6</td>
</tr>
<tr>
<td>Portugal</td>
<td>16.8</td>
<td>20.6</td>
<td>19.2</td>
<td>17.1</td>
<td>14.4</td>
<td>-2.3</td>
</tr>
<tr>
<td>Spain</td>
<td>30.5</td>
<td>28.1</td>
<td>23.6</td>
<td>18.6</td>
<td>14.9</td>
<td>-15.6</td>
</tr>
<tr>
<td>Sweden</td>
<td>23.2</td>
<td>21.7</td>
<td>20.3</td>
<td>22.0</td>
<td>20.0</td>
<td>-3.2</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>34.2</td>
<td>36.8</td>
<td>22.8</td>
<td>17.3</td>
<td>12.4</td>
<td>-21.7</td>
</tr>
<tr>
<td>United States</td>
<td>23.5</td>
<td>20.8</td>
<td>17.5</td>
<td>15.2</td>
<td>13.2</td>
<td>-10.4</td>
</tr>
<tr>
<td>EU15</td>
<td>26.6</td>
<td>23.4</td>
<td>22.6</td>
<td>19.9</td>
<td>18.2</td>
<td>-8.4</td>
</tr>
</tbody>
</table>

Source: EU KLEMS Database (O’Mahony and Timmer, 2009), author’s calculations.
Table 9: Share of Manufacturing in Employment (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>27.0</td>
<td>21.0</td>
<td>15.7</td>
<td>13.5</td>
<td>10.9</td>
<td>-16.1</td>
</tr>
<tr>
<td>Austria</td>
<td>26.3</td>
<td>25.0</td>
<td>21.5</td>
<td>16.8</td>
<td>15.5</td>
<td>-10.8</td>
</tr>
<tr>
<td>Belgium</td>
<td>31.9</td>
<td>25.5</td>
<td>22.0</td>
<td>17.3</td>
<td>14.5</td>
<td>-17.4</td>
</tr>
<tr>
<td>Denmark</td>
<td>24.8</td>
<td>19.7</td>
<td>18.9</td>
<td>17.0</td>
<td>13.8</td>
<td>-10.9</td>
</tr>
<tr>
<td>Finland</td>
<td>21.6</td>
<td>23.1</td>
<td>18.9</td>
<td>18.7</td>
<td>16.7</td>
<td>-5.0</td>
</tr>
<tr>
<td>France</td>
<td>23.4</td>
<td>22.1</td>
<td>18.5</td>
<td>15.3</td>
<td>12.9</td>
<td>-10.6</td>
</tr>
<tr>
<td>Germany</td>
<td>33.9</td>
<td>29.6</td>
<td>26.7</td>
<td>20.9</td>
<td>19.3</td>
<td>-14.6</td>
</tr>
<tr>
<td>Greece</td>
<td>20.4</td>
<td>23.0</td>
<td>22.1</td>
<td>16.2</td>
<td>13.9</td>
<td>-6.6</td>
</tr>
<tr>
<td>Ireland</td>
<td>20.5</td>
<td>21.6</td>
<td>19.7</td>
<td>18.3</td>
<td>13.7</td>
<td>-6.8</td>
</tr>
<tr>
<td>Italy</td>
<td>27.4</td>
<td>27.9</td>
<td>23.6</td>
<td>20.8</td>
<td>19.2</td>
<td>-8.2</td>
</tr>
<tr>
<td>Japan</td>
<td>20.3</td>
<td>23.5</td>
<td>23.9</td>
<td>20.1</td>
<td>18.9</td>
<td>-7.5</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>24.0</td>
<td>24.0</td>
<td>20.7</td>
<td>13.2</td>
<td>11.2</td>
<td>-12.8</td>
</tr>
<tr>
<td>Netherlands</td>
<td>24.8</td>
<td>20.4</td>
<td>17.8</td>
<td>14.0</td>
<td>12.0</td>
<td>-12.7</td>
</tr>
<tr>
<td>Portugal</td>
<td>22.1</td>
<td>23.4</td>
<td>23.2</td>
<td>20.2</td>
<td>17.6</td>
<td>-4.6</td>
</tr>
<tr>
<td>Spain</td>
<td>21.9</td>
<td>22.7</td>
<td>19.6</td>
<td>18.2</td>
<td>14.8</td>
<td>-7.0</td>
</tr>
<tr>
<td>Sweden</td>
<td>27.1</td>
<td>23.5</td>
<td>20.4</td>
<td>19.6</td>
<td>17.3</td>
<td>-9.8</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>33.2</td>
<td>27.5</td>
<td>20.3</td>
<td>16.7</td>
<td>11.8</td>
<td>-21.4</td>
</tr>
<tr>
<td>United States</td>
<td>22.4</td>
<td>19.1</td>
<td>14.9</td>
<td>12.7</td>
<td>9.9</td>
<td>-12.5</td>
</tr>
<tr>
<td>EU15</td>
<td>26.3</td>
<td>25.8</td>
<td>22.0</td>
<td>18.3</td>
<td>15.6</td>
<td>-12.7</td>
</tr>
</tbody>
</table>

Source: EU KLEMS Database (O’Mahony and Timmer, 2009), author’s calculations.
### 10.2 Appendix B: Growth of Value Added and Employment in Advanced Countries

Table 10: Average annual growth rate of real value added and employment, 1970-2007 (%)

<table>
<thead>
<tr>
<th></th>
<th>Real value added Economy</th>
<th>Manufacturing</th>
<th>Employment Economy</th>
<th>Manufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>3.3</td>
<td>1.5</td>
<td>1.7</td>
<td>-0.9</td>
</tr>
<tr>
<td>Austria</td>
<td>2.7</td>
<td>3.1</td>
<td>0.3</td>
<td>-0.8</td>
</tr>
<tr>
<td>Belgium</td>
<td>2.4</td>
<td>2.7</td>
<td>0.4</td>
<td>-1.8</td>
</tr>
<tr>
<td>Denmark</td>
<td>1.9</td>
<td>1.0</td>
<td>0.5</td>
<td>-1.3</td>
</tr>
<tr>
<td>Finland</td>
<td>2.8</td>
<td>4.4</td>
<td>0.3</td>
<td>-0.5</td>
</tr>
<tr>
<td>France</td>
<td>2.5</td>
<td>1.9</td>
<td>0.6</td>
<td>-1.3</td>
</tr>
<tr>
<td>Germany</td>
<td>2.2</td>
<td>1.4</td>
<td>0.5</td>
<td>-1.2</td>
</tr>
<tr>
<td>Greece</td>
<td>3.0</td>
<td>1.4</td>
<td>0.8</td>
<td>-0.2</td>
</tr>
<tr>
<td>Ireland</td>
<td>4.9</td>
<td>7.1</td>
<td>1.9</td>
<td>0.6</td>
</tr>
<tr>
<td>Italy</td>
<td>2.3</td>
<td>2.5</td>
<td>0.6</td>
<td>-0.2</td>
</tr>
<tr>
<td>Japan</td>
<td>3.2</td>
<td>3.9</td>
<td>0.5</td>
<td>-0.6</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>5.0</td>
<td>1.8</td>
<td>2.4</td>
<td>-0.6</td>
</tr>
<tr>
<td>Netherlands</td>
<td>2.7</td>
<td>2.4</td>
<td>1.2</td>
<td>-0.9</td>
</tr>
<tr>
<td>Portugal</td>
<td>3.2</td>
<td>2.4</td>
<td>0.4</td>
<td>-0.3</td>
</tr>
<tr>
<td>Spain</td>
<td>3.1</td>
<td>2.8</td>
<td>1.4</td>
<td>0.3</td>
</tr>
<tr>
<td>Sweden</td>
<td>2.3</td>
<td>3.2</td>
<td>0.3</td>
<td>-1.2</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2.1</td>
<td>0.6</td>
<td>0.4</td>
<td>-2.5</td>
</tr>
<tr>
<td>United States</td>
<td>2.9</td>
<td>2.8</td>
<td>1.6</td>
<td>-0.7</td>
</tr>
<tr>
<td>EU15</td>
<td>2.5</td>
<td>1.9</td>
<td>0.6</td>
<td>-1.0</td>
</tr>
</tbody>
</table>

Source: EU KLEMS Database (O’Mahony and Timmer, 2009), author’s calculations.
### 10.3 Appendix C: Direct Contribution of Manufacturing Trade to Deindustrialization in Advanced Countries

Table 11: Estimated contribution of manufacturing trade to deindustrialization based on the coefficients from model 4, 1970-2006 (in % of total change)

<table>
<thead>
<tr>
<th>Country</th>
<th>North-North Trade</th>
<th>North-South Trade</th>
<th>Total Trade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>12.7 [7.3; 17.9]</td>
<td>34.6 [21.6; 47.0]</td>
<td>47.3 [28.9; 64.9]</td>
</tr>
<tr>
<td>Austria</td>
<td>-1.4 [-26.9; 22.8]</td>
<td>10.1 [1.0; 19.0]</td>
<td>8.6 [-26.0; 41.7]</td>
</tr>
<tr>
<td>Canada</td>
<td>3.4 [-12.7; 18.9]</td>
<td>33.1 [23.0; 42.7]</td>
<td>36.4 [10.3; 61.6]</td>
</tr>
<tr>
<td>Denmark</td>
<td>-2.6 [-21.0; 15.1]</td>
<td>13.6 [3.3; 23.5]</td>
<td>11.0 [-17.6; 38.7]</td>
</tr>
<tr>
<td>Finland</td>
<td>-27.2 [-48.0; -6.9]</td>
<td>16.0 [2.6; 29.6]</td>
<td>41.6 [-45.4; 22.7]</td>
</tr>
<tr>
<td>France</td>
<td>8.9 [-12.0; 29.0]</td>
<td>12.9 [3.5; 22.0]</td>
<td>21.8 [-8.5; 51.0]</td>
</tr>
<tr>
<td>Germany</td>
<td>-8.7 [-35.2; 16.4]</td>
<td>12.6 [0.6; 24.3]</td>
<td>4.0 [-34.6; 40.7]</td>
</tr>
<tr>
<td>Italy</td>
<td>3.2 [-15.9; 21.7]</td>
<td>12.7 [2.4; 22.9]</td>
<td>15.9 [-13.5; 44.6]</td>
</tr>
<tr>
<td>Japan</td>
<td>-3.2 [-7.7; 1.4]</td>
<td>14.3 [1.5; 27.1]</td>
<td>11.1 [-6.2; 28.5]</td>
</tr>
<tr>
<td>Norway</td>
<td>14.6 [9.4; 19.7]</td>
<td>15.0 [9.1; 20.8]</td>
<td>29.6 [18.5; 40.5]</td>
</tr>
<tr>
<td>Portugal</td>
<td>28.6 [18.5; 73.6]</td>
<td>7.0 [2.5; 11.5]</td>
<td>35.6 [-16.2; 85.2]</td>
</tr>
<tr>
<td>Spain</td>
<td>17.7 [-8.7; 43.1]</td>
<td>29.3 [15.6; 42.7]</td>
<td>47.1 [6.9; 85.8]</td>
</tr>
<tr>
<td>Sweden</td>
<td>-2.5 [-22.9; 17.2]</td>
<td>3.0 [-3.2; 9.1]</td>
<td>0.5 [-26.1; 26.3]</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>10.2 [2.2; 18.3]</td>
<td>24.1 [15.6; 32.4]</td>
<td>34.3 [17.8; 50.7]</td>
</tr>
<tr>
<td>United States</td>
<td>5.6 [1.1; 10.1]</td>
<td>25.7 [14.5; 36.5]</td>
<td>31.3 [15.6; 46.6]</td>
</tr>
</tbody>
</table>

166
Table 12: Estimated contribution of manufacturing trade to deindustrialization based on the coefficients from model 5, 1970-2006 (in % of total change)

<table>
<thead>
<tr>
<th>Country</th>
<th>North-North Trade</th>
<th>North-South Trade</th>
<th>Total Trade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>10.1</td>
<td>36.3</td>
<td>46.3</td>
</tr>
<tr>
<td></td>
<td>[5.9; 14.2]</td>
<td>[24.8; 47.2]</td>
<td>[30.7; 61.4]</td>
</tr>
<tr>
<td>Austria</td>
<td>4.4</td>
<td>12.8</td>
<td>17.2</td>
</tr>
<tr>
<td></td>
<td>[-20.4; 28.0]</td>
<td>[3.7; 21.7]</td>
<td>[-16.7; 49.7]</td>
</tr>
<tr>
<td>Canada</td>
<td>5.6</td>
<td>34.9</td>
<td>40.5</td>
</tr>
<tr>
<td></td>
<td>[-7.4; 18.2]</td>
<td>[25.9; 43.6]</td>
<td>[18.5; 61.9]</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.7</td>
<td>16.6</td>
<td>17.3</td>
</tr>
<tr>
<td></td>
<td>[-17.1; 17.8]</td>
<td>[6.4; 26.5]</td>
<td>[-10.7; 44.3]</td>
</tr>
<tr>
<td>Finland</td>
<td>-23.6</td>
<td>19.8</td>
<td>-3.8</td>
</tr>
<tr>
<td></td>
<td>[-37.8; -9.7]</td>
<td>[4.5; 35.1]</td>
<td>[-33.3; 25.4]</td>
</tr>
<tr>
<td>France</td>
<td>11.4</td>
<td>11.7</td>
<td>23.1</td>
</tr>
<tr>
<td></td>
<td>[-3.6; 25.9]</td>
<td>[5.5; 17.9]</td>
<td>[1.9; 43.8]</td>
</tr>
<tr>
<td>Germany</td>
<td>-0.8</td>
<td>15.5</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>[-25.3; 22.4]</td>
<td>[4.5; 26.1]</td>
<td>[-20.8; 48.5]</td>
</tr>
<tr>
<td>Italy</td>
<td>6.7</td>
<td>12.6</td>
<td>19.4</td>
</tr>
<tr>
<td></td>
<td>[-6.1; 19.3]</td>
<td>[5.5; 19.7]</td>
<td>[-0.6; 39.0]</td>
</tr>
<tr>
<td>Japan</td>
<td>-2.1</td>
<td>18.8</td>
<td>16.8</td>
</tr>
<tr>
<td></td>
<td>[-6.9; 2.8]</td>
<td>[2.9; 34.3]</td>
<td>[-3.9; 37.1]</td>
</tr>
<tr>
<td>Norway</td>
<td>10.7</td>
<td>14.5</td>
<td>25.3</td>
</tr>
<tr>
<td></td>
<td>[8.1; 13.3]</td>
<td>[9.9; 19.1]</td>
<td>[18.0; 32.4]</td>
</tr>
<tr>
<td>Portugal</td>
<td>34.4</td>
<td>6.6</td>
<td>41.0</td>
</tr>
<tr>
<td></td>
<td>[4.7; 64.1]</td>
<td>[3.1; 10.1]</td>
<td>[7.8; 74.2]</td>
</tr>
<tr>
<td>Spain</td>
<td>19.4</td>
<td>27.3</td>
<td>46.7</td>
</tr>
<tr>
<td></td>
<td>[6.1; 32.7]</td>
<td>[16.6; 37.8]</td>
<td>[22.7; 70.5]</td>
</tr>
<tr>
<td>Sweden</td>
<td>-1.1</td>
<td>3.7</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>[-17.1; 14.4]</td>
<td>[-1.1; 8.6]</td>
<td>[-18.3; 23.0]</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>14.9</td>
<td>28.4</td>
<td>43.2</td>
</tr>
<tr>
<td></td>
<td>[5.0; 24.8]</td>
<td>[21.8; 34.9]</td>
<td>[26.8; 59.7]</td>
</tr>
<tr>
<td>United States</td>
<td>7.0</td>
<td>31.2</td>
<td>38.2</td>
</tr>
<tr>
<td></td>
<td>[2.3; 11.7]</td>
<td>[19.6; 42.3]</td>
<td>[21.9; 54.0]</td>
</tr>
</tbody>
</table>
Chapter 4

A New Shift-Share Method

Lionel Artige and Leif van Neuss

Abstract

Shift-share analysis is a decomposition technique widely used in regional studies to quantify an industry-mix effect and a competitive effect on the growth of regional employment (or any other relevant variable) relative to the national average. This technique has always been subject to criticism for its lack of theoretical basis. This paper presents a critical assessment of the methods suggested by Dunn (1960) and Esteban-Marquillas (1972) and proposes a new shift-share method, which separates out the two effects unambiguously. By way of illustration, we provide an application to manufacturing employment in the Belgian provinces between 1995 and 2007.

Keywords: Economic Structure, Regional Economics, Shift-Share, Belgian Manufacturing Employment

JEL classification: R10, R11, R12, R58
1 Introduction

Shift-share analysis is a decomposition technique widely used in regional studies to identify sectoral effects - the one resulting from the sectors' weights in the economy and the other from the sectors' growth rates - leading to inequality in employment growth across regions (Murray, 2010). Although the method was developed in the early 1940s, it is generally attributed to Dunn (1960) in the literature. The objective of shift-share analysis is to compare the sectoral distributions of employment growth between two geographical areas (usually a region versus the nation as a whole) in order to answer three questions: i) Does the regional economic structure yield more growth than the national one? ii) Is the regional sectoral growth higher on average than the national one? iii) From the results to i) and ii), which one from the structure or the sectoral efficiency contributes more to the observed differential in aggregate employment growth between the region and the nation?

What shift-share analysis can offer is to propose ordinal variables to answer i) and ii) and a decomposition technique to answer iii). As the shift-share technique is an accounting identity, any formula satisfying this identity is mathematically correct. Therefore, whereas many decompositions are mathematically possible, only one should answer questions i), ii) and iii) unambiguously. While various shift-share decompositions have been proposed in the literature, none is fully convincing yet. The first important decomposition method was proposed by Dunn (1960), who defines a growth effect from the economic structure of regional employment - which the literature calls an “industry-mix effect” after Esteban-Marquillas (1972) - and finds a residual, which is meant to measure what Esteban-Marquillas (1972) calls a “competitive effect” or, for others in the literature, a “regional effect.”

1The method has been applied to many other indicators such as income, population and productivity. This paper, like many previous studies using this technique, focuses on employment as this data is easily available at regional level.
2The origins of shift-share decomposition are not too clear as the literature variously attributes its authorship. Ray (1990) cites Jones (1940) as the first publication using shift-share analysis.
3Although we prefer “sectoral efficiency effect” to designate the effect of the sectors’ growth rates, we will use, in this paper, the term “competitive effect” as in the literature.
Rosenfeld (1959) soon criticized this residual arguing that the competitive effect in Dunn’s method was not properly defined, as it included some of the industry-mix effect. As a response to this criticism, Esteban-Marquillas (1972) modified the shift-share technique by adding a third component to construct another competitive effect. This third component, called the “allocation effect”, is the residual required by the accounting identity.

Both methods raised a lot of criticism (Houston, 1967; Richardson, 1978). Within this, we can single out the lack of theoretical foundations (see, e.g., Bartels et al., 1982) and Cunningham’s (1969) observation that both Dunn and Esteban-Marquillas’ decompositions yielded two solutions with different values for the industry-mix and competitive effects. These deficiencies sparked off many shift-share reformulations so as to deepen the analysis of regional effects of growth (Arceles, 1984); to include interregional and international trade flows in the analysis (Dinc and Haynes, 2005; Markusen et al., 1991; Sihag and McDonough, 1989); and to take short-term fluctuations into account within the study periods (Barff and Knight, 1988). None of these corrections, however, fundamentally departs from the methods of Dunn (1960) and Esteban-Marquillas (1972) as all of these extensions, in fact, remained based on either.

In the present paper, we argue that the decomposition methods proposed by the shift-share literature do not solve the methodological problems identified by Rosenfeld (1959) and Cunningham (1969). In particular, we consider that the definition of the competitive effect is not only flawed in Dunn (1960) but also in Esteban-Marquillas (1972). As a result, both methods fail to separate out a structural effect and a competitive effect relative to the national average. This may lead to incorrect numerical results in empirical studies and inaccurate policy advice. The contribution of this paper is to provide: 1) a comprehensive study of the methods of Dunn (1960) and Esteban-Marquillas (1972); 2) a test to assess the validity of any shift-share technique and 3) a new technique, which solves the definitional and technical shortcomings of the traditional shift-share methods.

Dunn’s shift-share method was first published in French Dunn (1939) with Rosenfeld’s reply appearing in the same publication.
The paper is organized as follows. Section 2 discusses the usefulness of shift-share analysis. The methods of Dunn (1960) and Esteban-Marquillas (1972) are presented and examined in Section 3 and Section 4 respectively. Section 5 develops a test for shift-share methods. Section 6 presents a new shift-share decomposition. Section 7 provides an application of this new technique to employment variations in the manufacturing sector of the Belgian provinces between 1995 and 2007 and compares the results with those of the other two methods. Finally, section 8 presents our conclusions.

2 On the Merits of Shift-Share Analysis

The growth rate of aggregate employment at the regional (or national) level can be disaggregated into a sum of sectoral growth rates weighted by the shares of sectors in regional (or national) employment. The aggregate employment growth performance thus depends on the economic structure (the weights) and the growth rate of each sector. If we observe that the growth rate of regional employment is lower than the national one, it can be interesting to investigate the extent to which the difference is attributable to the effect of the weights, on the one hand, and to the effect of the mean of the sectoral growth rates, on the other. This investigation requires to separate the effect of the economic structure from that of the average growth performance of all sectors.

If there were observable ordinal variables to measure these two effects, regressing the employment growth differential observed annually between the region and the nation on the differentials of these two variables would be good enough to realize this investigation\(^5\). Such a variable can easily be created in order to measure the effect of the growth performance of sectors: it suffices to take the sum of the sectoral growth rates weighted by a uniform distribution of sectors, which eliminates any effect of the economic structure\(^6\). Once this

\(^5\)As an alternative to shift-share accounting, Weeden (1974), Buck and Atkins (1976), Berzeg (1978), and Patterson (1991) have developed econometric analyses of structural effects on regional growth based on binary variables.

\(^6\)Let us emphasize that whenever the distribution of sectors is non-uniform the growth performance of sectors is not purged of any effect of the economic structure.
variable is available for the region and the nation, the differential between the two territories can be put in as an explanatory variable.

Yet, there is no obvious way of constructing an ordinal variable to measure the economic structure because an economic structure defined as the distribution of sectors is not a variable with an intrinsic ordering. Therefore, taking the difference between the regional and national distributions of sectors would not make any sense and taking the difference between the regional and national shares of sectors weighted by a uniform distribution of sectoral growth rates would necessarily equal zero. The construction of such a variable thus requires a non-uniform distribution of sectoral growth rates, which means that the economic structure cannot be isolated from the sectoral growth rates. The question is then: what non-uniform distribution? As mathematically there is an infinity of non-uniform distributions of sectoral growth rates, it is impossible to state whether, in total, the regional or the national economic structure yields more employment growth.

The contribution of shift-share analysis lies in yielding ordinal variables to measure an effect of the economic structure and an effect of the sectoral growth rates on the observed differential in aggregate employment growth between two territories. Separating these two effects clearly amounts to an accounting exercise and shift-share analysis aims at providing a technique to do so.

3 Dunn’s (1960) Shift-Share Method

3.1 The Decomposition Method

Shift-share analysis organizes data along three dimensions: geography, sectors of activity and time. The shift-share method proposed by Dunn (1960) consists in comparing regional employment growth observed in the data with a hypothetical employment growth that the region would have experienced, were its growth rate equal to the national one. The objective of the method is to decompose the difference between these two employment variations into two components: a structural effect (industry-mix effect) and a competitive
effect\(^7\). Formerly, for a region \(j\), we have\(^8\)

\[
\sum_{i=1}^{I} (n_{i,t+1}^j - n_{i,t}^j) = \sum_{i=1}^{I} n_{i,t}^j(r_{i,t+1} - r_{t+1}) + \sum_{i=1}^{I} n_{i,t}^j(g_{i,t+1}^j - r_{i,t+1})
\]

(1)

where \(n_{i,t+1}^j\) is employment in sector \(i = 1, \ldots, I\) of region \(j\) at time \(t + 1\), \(g_{i,t+1}^j\) is the employment growth rate between time \(t\) and \(t + 1\) in sector \(i\) of region \(j\), and \(r_{i,t+1}\) and \(r_{t+1}\) are the national employment growth rates between time \(t\) and \(t + 1\) in, respectively, sector \(i\) and the total economy. The left-hand side of Equation (1) is the difference in observed and hypothetical regional employment growth between time \(t\) and \(t + 1\). On the right-hand side, the first component (industry-mix effect), \(\sum_{i=1}^{I} n_{i,t}^j(r_{i,t+1} - r_{t+1})\), quantifies the effect of the economic structure of region \(j\) on the growth employment differential between the region and the nation from time \(t\) to \(t + 1\). If employment in all sectors at the national level were to grow at the rate of national employment or if the regional and national economic structures were identical, this component would equal zero and the economic structure of region \(j\) would not matter for employment growth. The second component (competitive effect), \(\sum_{i=1}^{I} n_{i,t}^j(g_{i,t+1}^j - r_{i,t+1})\), quantifies the effect of the relative sectoral growth performance of region \(j\) on the growth employment differential between the region and the nation from time \(t\) to \(t + 1\). In order for this component to equal zero, the employment growth rates in each sector would need to be the same at the regional and national levels.

If we want to express the difference between observed and hypothetical regional employment growth in terms of percentage change, we divide Equation (1) by \(\sum_{i=1}^{I} n_{i,t}^j\) and obtain

\[
g_{t+1}^j - r_{t+1} = \frac{\sum_{i=1}^{I} n_{i,t}^j(r_{i,t+1} - r_{t+1})}{\sum_{i=1}^{I} n_{i,t}^j} + \frac{\sum_{i=1}^{I} n_{i,t}^j(g_{i,t+1}^j - r_{i,t+1})}{\sum_{i=1}^{I} n_{i,t}^j}
\]

(2)

\(^7\)Dunn (1960) refers to the industry-mix and competitive effects respectively as the differential and proportionality effects.

\(^8\)The development of this decomposition is given in Appendix A.
where \( g^j_{t+1} = \frac{\sum_{i=1}^{I} (n^i_{j,t+1} - n^i_{j,t})}{\sum_{i=1}^{I} n^i_{j,t}} \) is the employment growth rate of region \( j \) between time \( t \) and \( t + 1 \). The left-hand side of Equation (2) is the difference between the observed regional and national growth rates, and the two components are now expressed in terms of percentage change. A better way to understand Dunn’s (1960) decomposition is to rewrite the industry-mix effect in Equation (2) as:

\[
g^j_{t+1} - r_{t+1} = \left( \frac{\sum_{i=1}^{I} n^i_{j,t} r^i_{t+1}}{\sum_{i=1}^{I} n^i_{j,t}} - \frac{\sum_{i=1}^{I} c^i_t r^i_{t+1}}{\sum_{i=1}^{I} c^i_t} \right) + \frac{\sum_{i=1}^{I} n^i_{j,t} (g^j_{i,t+1} - r_{i,t+1})}{\sum_{i=1}^{I} n^i_{j,t}}
\]

(3)

where \( c^i_t \) is the national employment in sector \( i \) at time \( t \). To rewrite the industry-mix effect, we used the fact that \( \frac{\sum_{i=1}^{I} n^i_{j,t} r^i_{t+1}}{\sum_{i=1}^{I} n^i_{j,t}} = r_{t+1} = \frac{\sum_{i=1}^{I} c^i_t r^i_{t+1}}{\sum_{i=1}^{I} c^i_t} \).

Finally, we can rewrite Equation (3) in terms of shares in total employment in region \( j \) and at the national level:

\[
g^j_{t+1} - r_{t+1} = \sum_{i=1}^{I} (\omega^i_{j,t} - \theta^i_{t}) r_{i,t+1} + \sum_{i=1}^{I} \omega^i_{j,t} (g^j_{i,t+1} - r_{i,t+1})
\]

(4)

where \( \omega^i_{j,t} = \frac{n^i_{j,t}}{\sum_{i=1}^{I} n^i_{j,t}} \) is the share of sector \( i \) in region \( j \) in total employment in region \( j \) and \( \theta^i_{t} = \frac{c^i_t}{\sum_{i=1}^{I} c^i_t} \) is the share of sector \( i \) at the national level in total national employment.

### 3.2 A Critical Assessment

In Equation (4), the industry-mix effect, \( \sum_{i=1}^{I} (\omega^i_{j,t} - \theta^i_{t}) r_{i,t+1} \), is obtained by associating the national sectoral growth rates to the regional and the national economic structures while the competitive effect, \( \sum_{i=1}^{I} \omega^i_{j,t} (g^j_{i,t+1} - r_{i,t+1}) \), is obtained by associating the regional economic structure to the regional and national sectoral growth rates. In other words, this method makes a choice on the territorial basis of the growth rates to compute the industry-mix effect, and on the basis of the economic structure to compute the competitive effect. Dunn (1960) chose the national growth rates to calculate the industry-mix effect and the regional economic structure to calculate the competitive effect.
We argue that this choice is arbitrary. He could equally have chosen the regional growth rates for the industry-mix effect and the national economic structure for the competitive effect. This decomposition leads to the same difference between the regional and national employment growth rate but yields different values for the two effects if the region and the country have different economic structures and growth rates:

\[
g_{t+1}^j - r_{t+1} = \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t})g_{i,t+1}^j + \sum_{i=1}^{I} \theta_{i,t}(g_{i,t+1}^j - r_{i,t+1}) \quad (5)
\]

In the absence of any explicit criterion, there is no a priori reason to prefer one decomposition to the other\(^9\). Therefore, the shift-share method based on Dunn (1960) cannot deliver a unique value for each of the two effects.

Moreover, Rosenfeld (1959) emphasized that the competitive effect in Equation (4) was inconsistent: if two regions have identical sectoral growth rates but different economic structures they will have different competitive effects, which means that the economic structure affects the value of the competitive effect. From Equations (4) and (5) it clearly appears that neither decomposition suppresses all influence of the economic structure on the competitive effect.

4 Esteban-Marquillas’s (1972) Shift-Share Method

4.1 The Decomposition Method

When Dunn presented his shift-share method in 1959, Rosenfeld (1959) immediately identified an inconsistency in his definition of the competitive effect as already mentioned in the previous paragraph. He showed that if two regions have identical growth rates by sector but different economic structures, they will have different competitive effects relative to the national average because the competitive effect depends on the regional economic structure. As a result, the competitive effect is not purged of any industry-mix effect. Esteban-

---

\(^9\)In Appendix B we show how to move from Equation (4) to Equation (5).

\(^{10}\)Cunningham (1969) came to the same conclusion.
Marquillas (1972) proposed a solution that has since become the standard shift-share method. His solution computes the competitive effect as the difference between the sectoral regional and national growth rates weighted by the national economic structure. This implies the addition of a third component, called the “allocation” component, to Equation (1). Formerly, we have

\[
\sum_{i=1}^{I} (n_{i,t+1}^{j} - n_{i,t}^{j}) - \sum_{i=1}^{I} n_{i,t}^{j} r_{t+1} = \sum_{i=1}^{I} n_{i,t}^{j} (r_{i,t+1} - r_{t+1}) \\
+ \sum_{i=1}^{I} m_{i,t} (g_{i,t+1}^{j} - r_{i,t+1}) \\
+ \sum_{i=1}^{I} (n_{i,t}^{j} - m_{i,t}) (g_{i,t+1}^{j} - r_{i,t+1})
\]

(6)

where \(\sum_{i=1}^{I} m_{i,t} (g_{i,t+1}^{j} - r_{i,t+1})\) is the newly-defined competitive effect and \(m_{i,t} = \sum_{i=1}^{I} n_{i,t}^{j} \left( \frac{\sum_{j=1}^{J} n_{i,t}^{j}}{\sum_{i=1}^{I} \sum_{j=1}^{J} n_{i,t}^{j}} \right)\) is the “homothetic employment”, i.e., the hypothetical employment that region \(j\) would have, were its economic structure identical to the national one. The last term in Equation (6) is the allocation effect, i.e., the product of the difference between the observed and hypothetical economic structure of region \(j\) and the difference between the regional and national employment growth rate in sector \(i\). In applied papers, the economic interpretation of the allocation effect is evasive and often omitted. In their original works, both Esteban-Marquillas (1972) and Cunningham (1969) interpret a positive allocation effect as the contribution of regional specialization in sectors in which the regional growth rates are relatively the highest, and a negative allocation effect as a lack of regional specialization in the fastest growing sectors. In addition, Cunningham (1969) hints that the allocation effect can be indicative of a convergence (negative allocation effect) or a divergence (positive allocation effect) of the regional and the national economic structures.

In order to better understand the solution proposed by Esteban-Marquillas...
(1972), let us divide Equation (6) by \( \sum_{i=1}^{I} n_{i,t}^j \) in order to rewrite it in terms of percentage change

\[
g_{i+1}^j - r_{t+1} = \frac{\sum_{i=1}^{I} n_{i,t}^j (r_{i,t+1} - r_{t+1})}{\sum_{i=1}^{I} n_{i,t}^j} + \frac{\sum_{i=1}^{I} m_{i,t} (g_{i,t+1}^j - r_{i,t+1})}{\sum_{i=1}^{I} n_{i,t}^j} + \frac{\sum_{i=1}^{I} (n_{i,t}^j - m_{i,t}) (g_{i,t+1}^j - r_{i,t+1})}{\sum_{i=1}^{I} n_{i,t}^j}
\]

and, then, in terms of employment shares:

\[
g_{i+1}^j - r_{t+1} = \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) r_{i,t+1} + \sum_{i=1}^{I} \theta_{i,t} (g_{i,t+1}^j - r_{i,t+1}) + \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) (g_{i,t+1}^j - r_{i,t+1})
\]

### 4.2 A Critical Assessment

We argue that Esteban-Marquillas’s (1972) decomposition method brings no improvement to Dunn’s (1960) method for the following three reasons:

1. This method does not solve the main problem posed by the absence of unique values for the industry-mix and competitive effects in Dunn’s (1960) method. By comparing Equation (4) with Equation (8), we can observe that the solution proposed by Esteban-Marquillas (1972) uses the same territorial basis to compute the two effects: the national growth rates to compute the industry-mix effect and the national economic structure to compute the competitive effect. Not only is this choice arbitrary but it also requires a residual term (allocation effect) to satisfy the equality. It would have been possible to use the regional territorial basis to compute both effects, which requires the same allocation effect.\(^{11}\)

\(^{11}\)As Esteban-Marquillas (1972) looked for a competitive effect that would not vary for
\[ g_{t+1}^j - r_{t+1} = \sum_{i=1}^I (\omega^j_{i,t} - \theta_{i,t}) g_{i,t+1}^j + \sum_{i=1}^I \omega^j_{i,t} (g_{i,t+1}^j - r_{i,t+1}) - \sum_{i=1}^I (\omega^j_{i,t} - \theta_{i,t}) (g_{i,t+1}^j - r_{i,t+1}) \]  

Equation (9) changes the territorial basis of the sectoral growth rates in Equation (5) to compute an industry-mix effect with the same territorial basis as in the competitive effect, and adds a residual term to satisfy the equality. The values of the two effects are different between Equations (8) and (9) while the residual terms - the allocation effects - are of opposite signs. The allocation effect allows the modification of the competitive effect in Equation (8) and of the industry-mix effect in Equation (9). No more than Dunn’s solution can Esteban-Marquillas’ deliver a unique value for the industry-mix and competitive effects.  

2. This method is unnecessary to solve the inconsistent example identified by Rosenfeld (1959). In fact, this inconsistency can be solved without adding a third component by Equation (5). Let us recall that Equation (5) is:

\[ g_{t+1}^j - r_{t+1} = \sum_{i=1}^I (\omega^j_{i,t} - \theta_{i,t}) g_{i,t+1}^j + \sum_{i=1}^I \theta_{i,t} (g_{i,t+1}^j - r_{i,t+1}) \]

where \( \sum_{i=1}^I \theta_{i,t} (g_{i,t+1}^j - r_{i,t+1}) \) is the same competitive effect as the one constructed by Esteban-Marquillas (1972) in Equation (8). To the best of our knowledge, nobody has yet thought of this solution to Rosenfeld’s inconsistent example. Nevertheless, contrary to what is commonly believed, the solution proposed by Esteban-Marquillas (1972) does not solve the inconsistency identified by Rosenfeld as the former does not solve two regions with the same sectoral growth rates, Equation (8) is the most appropriate one. From a theoretical point of view, though, it is no longer justified to use Equation (8) rather than Equation (9).  

\[ ^{12} \text{Once again, Cunningham (1969) came to the same conclusion.} \]
succeed in removing any influence of the economic structure on the computation of the competitive effect.

3. This method adds a problematic residual term (allocation effect). First, it is unnecessary (see previous point). Second, when its value is different from zero, the value of the competitive effect is necessarily different from that of Dunn’s method based on Equation (4), and the value of the industry-mix effect is different from that of Dunn’s method based on Equation (5). How to justify this unless one proves that Dunn’s method is wrong? Third, the economic interpretation of this allocation effect given by Esteban-Marquillas (1972) and Cunningham (1969) refers to an effect of the economic structure, which should in fact be captured by the industry-mix effect in the first place.

We can conclude that Esteban-Marquillas’s (1972) method does not bring any improvement to Dunn’s (1960) method. In the next two sections, we propose a simple test to identify a relevant shift-share method and then propose a new technique. We show that whereas both Esteban-Marquillas (1972) and Dunn’s (1960) methods fail this test, our technique comes out successfully.

5 A Shift-Share Test

Shift-share analysis aims at answering the following question: does a region’s economic structure impact its growth performance positively or negatively? If it does negatively, the effect of the structure may be offset by the average growth performance of all sectors. Therefore, it would be interesting to discover whether the economic structure of a region is a relative strength (or weakness) in terms of growth and whether this strength (or weakness) is reinforced or offset by relatively higher (or lower) sectoral efficiency. A shift-share technique should be able to separate out these two effects unambiguously.

5.1 Two Difficulties

The first difficulty to tackle in shift-share analysis is the absence of order for an economic structure. What we call an economic structure is a distribution
of sectors with no intrinsic ordering. A region may specialize in some sectors more than others. We will say that a region or a nation specializes in a given sector if its employment share is larger than the uniform share. There is no a priori good or bad specialization. The regional specialization may be different from the national one but, without considering an associated variable, it is impossible to conclude that the regional specialization is better than the national one. Moreover, the specialization may vary between the region and the nation while their distribution of employment across sectors may be identical. For instance, employment shares can be 20% in sector A and 80% in sector B for the region while being the opposite for the nation. Although the distributions of employment shares are identical, the regional and national specializations are different.

When the distribution of sectors is associated with their corresponding sectoral growth rates, it is possible to conclude that a specialization will yield more or less employment growth. Yet, another difficulty arises if we want to disentangle the effect of specialization from the effect of growth rates. This is precisely the objective of shift-share analysis. We will now show that neither of the methods proposed by Dunn (1960) and Esteban-Marquillas (1972) solves this difficulty.

5.2 A Simple Shift-Share Test

Table 1 presents two numerical examples. In Example 1, the region and the nation are identical in all respects: the economic structures and the growth rates are the same for each sector. Obviously, there is no difference between regional and national employment growth, and any shift-share technique should find zero industry-mix and competitive effects. Table 2 shows that the shift-share techniques of Dunn (1960) and Esteban-Marquillas (1972) find the expected results. In Example 2, we keep the same pairs of data (employment shares and growth rates) as in Example 1 but assign them to different sectors between the two geographical units. The distribution of shares and growth rates is exactly the same as previously but the specializations are now different: the region specializes in Sector B while the nation does in Sector A. The difference between the regional and national growth rates between $t$ and
$t + 1$ should be zero. Moreover, any shift-share technique should conclude that the industry-mix and the competitive effects are null. Each geographical unit specializes in the highest performing sector and none has a systematic advantage in employment performance in all sectors. We can observe that the shift-share techniques of both Dunn (1960) and Esteban-Marquillas (1972) fail this test (Table 2). In fact, by choosing the national growth rates to compute the industry-mix effect, both techniques implicitly consider that the national specialization is better than the regional one. In the case, this conclusion turns out to be wrong, as regional and national employment growth rates are identical.

Table 1: Two numerical examples

<table>
<thead>
<tr>
<th>Region</th>
<th>Share of total employment at $t$</th>
<th>Employment growth rate between $t$ and $t + 1$</th>
<th>Nation</th>
<th>Share of total employment at $t$</th>
<th>Employment growth rate between $t$ and $t + 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Example 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sector A</td>
<td>80%</td>
<td>5%</td>
<td>Sector A</td>
<td>80%</td>
<td>5%</td>
</tr>
<tr>
<td>Sector B</td>
<td>20%</td>
<td>4%</td>
<td>Sector B</td>
<td>20%</td>
<td>4%</td>
</tr>
<tr>
<td><strong>Example 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sector A</td>
<td>20%</td>
<td>4%</td>
<td>Sector A</td>
<td>80%</td>
<td>5%</td>
</tr>
<tr>
<td>Sector B</td>
<td>80%</td>
<td>5%</td>
<td>Sector B</td>
<td>20%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Table 2: Dunn and Esteban-Marquillas’s shift-share methods tested

<table>
<thead>
<tr>
<th>Example</th>
<th>Growth rate differential (%)</th>
<th>Industry-mix effect (%)</th>
<th>Competitive effect (%)</th>
<th>Allocation effect (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Example 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dunn’s decomposition</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>EM’s decomposition</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Example 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dunn’s decomposition</td>
<td>0.0</td>
<td>-0.6</td>
<td>0.6</td>
<td>-</td>
</tr>
<tr>
<td>EM’s decomposition</td>
<td>0.0</td>
<td>-0.6</td>
<td>-0.6</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Notes: Growth rate differential is the difference between the regional and the national aggregate employment growth rates. EM, Esteban-Marquillas’ shift-share method.
A valid shift-share technique should result in a unique decomposition of the growth differential between two geographical units into an industry-mix effect and a competitive effect. In addition, this technique should solve Rosenfeld’s inconsistency and pass the test of Example 2. The new technique we now propose provides the solution to separate out unambiguously an effect of the economic structure and a competitive effect.

Our technique starts with the construction of the competitive effect. Rosenfeld (1959) rightly pointed out that the competitive effect should not be influenced by the economic structure if one wanted to separate out an effect from the economic structure and an effect from the sectoral growth rates. Both Dunn (1960) and Esteban-Marquillas’s (1972) methods fail in building such a competitive effect. As mentioned in the second section, the only way to purge the competitive effect from any influence of the economic structure is to associate a uniform distribution of sectors to the sectoral growth rates. Therefore, we define the competitive effect as

$$\sum_{i=1}^{I} \frac{1}{I} \left( g_{i,t+1}^{j} - r_{i,t+1} \right)$$

(10)

where $I$ is the number of sectors and $\frac{1}{I}$ is the employment share of each sector. Equation (10) is the difference between the arithmetic means of the regional and national sectoral growth rates. If Equation (10) is positive, the arithmetic mean of the sectoral growth rates is higher in the region than in the nation. In that case, the sectors, on average, yield more growth in the region than in the nation. Then, we calculate the effect of the economic structure (or the industry-mix effect) as the residual, i.e., as the difference between the differential in the aggregate employment growth rates ($g_{t+1}^{j} - r_{t+1}$), on the one hand, and Equation (10), on the other hand, which yields:

$$\sum_{i=1}^{I} \left( \omega_{i,t}^{j} - \frac{1}{I} \right) g_{i,t+1}^{j} - \sum_{i=1}^{I} \left( \theta_{i,t} - \frac{1}{I} \right) r_{i,t+1}$$

(11)
where $\sum_{i=1}^{I} \left( \omega_{j,i,t} - \frac{1}{I} \right)$ and $\sum_{i=1}^{I} \left( \theta_{i,t} - \frac{1}{I} \right)$ are the regional and national specializations respectively. As mentioned in Section 2, the economic structure is a distribution of sectors and the difference between two distributions is meaningless. As such, we cannot say which specialization is better than the other. Yet the specializations associated with their corresponding sectoral growth rates, as it comes out in Equation (11), are ordinal variables measuring employment growth due to specialization. Equation (11) allows us to determine which of the two specializations yields more employment growth. If Equation (11) is positive, then the regional economic structure yields more employment growth than the national one. Our new shift-share decomposition equation thus is the sum of Equations (10) and (11):

$$g_{j,t+1} - r_{t+1} = \left[ \sum_{i=1}^{I} \left( \omega_{j,i,t} - \frac{1}{I} \right) g_{j,i,t+1} - \sum_{i=1}^{I} \left( \theta_{i,t} - \frac{1}{I} \right) r_{i,t+1} \right]$$

$$+ \sum_{i=1}^{I} \frac{1}{I} \left( g_{i,t+1} - r_{i,t+1} \right)$$

Equation (12) accounts for the observed difference between the regional and national aggregate employment growth rates and separates out the industry-mix and the competitive effects unambiguously. Finally, Table 3 shows that our method passes the shift-share test as it yields the expected results for the industry-mix and competitive effects in Example 2.

<table>
<thead>
<tr>
<th>Table 3: Our shift-share method tested</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Example 1</th>
<th>Example 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our decomposition</td>
<td>0.0</td>
</tr>
</tbody>
</table>

184
Let us insist that our method is a major departure from the approach commonly used in the shift-share literature. Dunn (1960) and Esteban-Marquillas (1972) compare an actual regional growth rate with two hypothetical regional growth rates that would result if the sectoral growth rates or the economic structure were identical to those of the nation. Therefore, the actual and hypothetical growth effects of the economic structure and the sectoral growth rates are mixed up in their decomposition formula. Our method only uses actual data, computes actual growth effects in both geographical units and compare them. It enables us to decompose the aggregate growth rate of any geographical unit into two terms, which capture the two effects we are interested in: the growth effect of the economic structure and the growth effect of the sectoral growth performances. The regional and national decompositions are the following:

\[
g_{j,t+1} = \sum_{i=1}^{I} \left( \omega_{j,i,t} - \frac{1}{I} \right) g_{i,t+1} + \sum_{i=1}^{I} \frac{1}{I} g_{i,t+1}
\]

\[
r_{t+1} = \sum_{i=1}^{I} \left( \theta_{i,t} - \frac{1}{I} \right) r_{i,t+1} + \sum_{i=1}^{I} \frac{1}{I} r_{i,t+1}
\]

The growth effect of the economic structure is measured respectively \( \sum_{i=1}^{I} \left( \omega_{j,i,t} - \frac{1}{I} \right) g_{i,t+1} \) in the region and \( \sum_{i=1}^{I} \left( \theta_{i,t} - \frac{1}{I} \right) r_{i,t+1} \) in the nation. It is positive if the territory specializes, on average, in fast-growing sectors, i.e. in sectors which experience a relative high growth rate. The growth effect of the sectoral growth performances is measured respectively \( \sum_{i=1}^{I} \frac{1}{I} g_{i,t+1} \) and \( \sum_{i=1}^{I} \frac{1}{I} r_{i,t+1} \) in the region and in the nation. Equations (13) and (14) are independent from each other and can be used to create ordinal variables for the economic structures in the region and in the nation. With our new shift-share decomposition, we can compare the two growth effects of a geographical unit with those of any other geographical unit without defining a reference territory. For instance, if one wants to assess whether the specialization of the region is favorable or harmful to its

\[\text{In growth regressions, it would be possible to use this ordinal variable measuring the growth effect of the economic structure as an explanatory variable.}\]
growth performance, in comparison with the nation, it is enough to compare the growth effect of the regional economic structure \( \left( \sum_{i=1}^{I} \left( \omega_{i,t} - \frac{1}{I} \right) g_{i,t+1} \right) \) with the growth effect of the national one \( \left( \sum_{i=1}^{I} \left( \theta_{i,t} - \frac{1}{I} \right) r_{i,t+1} \right) \). If one wants to assess the total regional growth performance in terms of industrial specialization and sectoral efficiency relative to the nation, one has to take the difference between Equations (13) and (14), which yields Equation (12).

7 An Application to Employment in the Belgian Manufacturing Sector between 1995 and 2007

By way of illustration, we propose to carry out a shift-share analysis of employment variations in the manufacturing sector in the Belgian provinces and the Brussels region between 1995 and 2007, and to compare the results of our technique with those of the traditional shift-share methods. Data on 14 sub-sectors of the manufacturing sector was retrieved from the Belgian Central Bank’s database for the 10 Belgian provinces and for Brussels, as listed in the first column of Table 4. At the national level, data shows that manufacturing employment decreased by 13.4% over that period. The second column of the table displays the employment growth rate differential of each province and Brussels relative to the national growth rate. We then computed the industry-mix and the competitive effects using Dunn’s (1960) technique (third and fourth columns), the same two effects plus the allocation effect using Esteban-Marquillas’s (1972) technique (fifth to seventh columns) and the industry-mix and the competitive effects using our new technique (last two columns). This exercise clearly shows that Dunn (1960) and Esteban-Marquillas’s (1972) techniques can lead to very misleading measures of the competitive and the industry-mix effects. For instance, in the province of Liège, where employment fell 3.6% under the national average, the industry-mix effect and the competitive effect amount to respectively 4.0% and -7.6% with Dunn’s (1960) method, as against 4.0% and 0.1% (while the allocation effect reaches -7.7%) with Esteban-Marquillas’s (1972) method, and -0.8% and -2.8% with our own method. In terms of policy prescriptions, the con-
clusions based on our decomposition technique stood in clear contrast with those of Dunn (1960) and Esteban-Marquillas (1972): the economic structure of the Province of Liège does not provide it with a relative structural advantage in terms of employment growth. In light of these comparative evidence, we recommend the use of our technique in shift-share studies.

8 Conclusion

The shift-share method is an accounting technique which aims at determining whether the aggregate growth performance of a region relative to the national average is the result of its economic structure or/and the growth rates of its sectors. Hence, the accounting formula should be able to separate out the two components unambiguously. This paper attempts to show that the traditional shift-share methods proposed by Dunn (1960) and Esteban-Marquillas (1972) fail to do so due to a flawed definition of the competitive effect.

Instead of these, the shift-share decomposition technique we recommend here is based on a competitive effect defined as the sum of the sectoral growth rates weighted by a uniform distribution of sectors. This is the only way to eliminate any effect of the second component, the economic structure, which is computed as the residual. Thus the separation between the two components is unambiguous.

Since all accounting shift-share methods are mathematically correct, we designed a simple test to assess the conceptual accuracy of shift-share methods and rule out inaccurate ones. The test confirms the flaws that we identified in Dunn (1960) and Esteban-Marquillas’s (1972) methods and validates the relevance of our own.

Finally, our empirical application on employment in the Belgian manufacturing sector between 1995 and 2007 shows that the three methods can yield very different results for the industry-mix and competitive effects. Though shift-share analysis does not highlight the causes of regional growth, it is very useful in identifying and quantifying these possible sources of regional growth performance. Therefore, the conceptual accuracy of the accounting technique is compelling in order to deliver the right assessment in regional studies.
Table 4: A shift-share analysis of the industrial sector in the Belgian provinces and the region of Brussels between 1995 and 2007

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
</tr>
<tr>
<td>Antwerpen</td>
<td>2.0</td>
<td>2.5</td>
<td>-0.5</td>
<td>2.5</td>
<td>-1.7</td>
<td>1.2</td>
<td>5.8</td>
<td>.38</td>
</tr>
<tr>
<td>Limburg</td>
<td>1.3</td>
<td>0.0</td>
<td>1.3</td>
<td>0.0</td>
<td>5.0</td>
<td>-3.7</td>
<td>-.4.3</td>
<td>5.6</td>
</tr>
<tr>
<td>Oost-Vlaanderen</td>
<td>-4.4</td>
<td>-4.0</td>
<td>-0.4</td>
<td>-4.0</td>
<td>0.9</td>
<td>-1.3</td>
<td>-4.8</td>
<td>0.4</td>
</tr>
<tr>
<td>Vlaams-Brabant</td>
<td>-4.6</td>
<td>1.6</td>
<td>-6.2</td>
<td>1.6</td>
<td>-9.5</td>
<td>3.4</td>
<td>4.2</td>
<td>.8.7</td>
</tr>
<tr>
<td>West-Vlaanderen</td>
<td>5.9</td>
<td>-4.7</td>
<td>10.6</td>
<td>-4.7</td>
<td>10.7</td>
<td>-0.1</td>
<td>-1.3</td>
<td>7.2</td>
</tr>
<tr>
<td>Brabant wallon</td>
<td>35.5</td>
<td>5.0</td>
<td>30.5</td>
<td>5.0</td>
<td>24.7</td>
<td>5.7</td>
<td>14.0</td>
<td>21.5</td>
</tr>
<tr>
<td>Hainaut</td>
<td>1.0</td>
<td>1.0</td>
<td>-2.0</td>
<td>1.0</td>
<td>1.3</td>
<td>-3.3</td>
<td>8.7</td>
<td>7.7</td>
</tr>
<tr>
<td>Liège</td>
<td>-3.6</td>
<td>4.0</td>
<td>-7.6</td>
<td>4.0</td>
<td>0.1</td>
<td>-7.7</td>
<td>-0.8</td>
<td>.28</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>16.5</td>
<td>5.3</td>
<td>11.2</td>
<td>5.3</td>
<td>7.8</td>
<td>3.5</td>
<td>8.5</td>
<td>8.0</td>
</tr>
<tr>
<td>Namur</td>
<td>17.2</td>
<td>4.4</td>
<td>12.8</td>
<td>4.4</td>
<td>10.2</td>
<td>2.7</td>
<td>8.9</td>
<td>8.4</td>
</tr>
<tr>
<td>Brussels</td>
<td>-18.9</td>
<td>0.3</td>
<td>-19.2</td>
<td>0.3</td>
<td>-17.2</td>
<td>-2.9</td>
<td>-2.6</td>
<td>-16.3</td>
</tr>
</tbody>
</table>

Growth rate differential is the difference between the provincial and the national total employment growth rates.

Source: Belgostat; calculations: authors.
9 Appendix

9.1 Appendix A: The Original Decomposition of Dunn (1960)

Let us define employment in sector $i$ at time $t$ in region $j$ by $n_{i,t}^j$ and in the nation by $m_{i,t}$. Equation (1) is obtained from the difference between the regional and national employment variations:

$$\frac{\sum_{i=1}^{I} n_{i,t+1}^j}{\sum_{i=1}^{I} n_{i,t}^j} - \frac{\sum_{i=1}^{I} m_{i,t+1}}{\sum_{i=1}^{I} m_{i,t}} = (1 + g_{t+1}^j) - (1 + r_{t+1})$$

$$= \frac{\sum_{i=1}^{I} n_{i,t}^j (1 + g_{t+1}^j)}{\sum_{i=1}^{I} n_{i,t}^j} - \frac{\sum_{i=1}^{I} n_{i,t}^j (1 + r_{t+1})}{\sum_{i=1}^{I} n_{i,t}^j}$$

$$= \frac{\sum_{i=1}^{I} n_{i,t}^j r_{i,t+1}}{\sum_{i=1}^{I} n_{i,t}^j} - \frac{\sum_{i=1}^{I} n_{i,t}^j r_{i,t+1}}{\sum_{i=1}^{I} n_{i,t}^j}$$

$$= \frac{\sum_{i=1}^{I} n_{i,t}^j (r_{i,t+1} - r_{i,t+1})}{\sum_{i=1}^{I} n_{i,t}^j} + \frac{\sum_{i=1}^{I} n_{i,t}^j (g_{t+1}^j - r_{i,t+1})}{\sum_{i=1}^{I} n_{i,t}^j}$$

where we used the fact that $\sum_{i=1}^{I} n_{i,t}^j g_{t+1}^j = \sum_{i=1}^{I} n_{i,t}^j (g_{i,t+1} - r_{i,t+1})$. Therefore, $g_{t+1}^j - r_{t+1} = \frac{\sum_{i=1}^{I} n_{i,t}^j (r_{i,t+1} - r_{i,t+1})}{\sum_{i=1}^{I} n_{i,t}^j} + \frac{\sum_{i=1}^{I} n_{i,t}^j (g_{i,t+1} - r_{i,t+1})}{\sum_{i=1}^{I} n_{i,t}^j}$. By multiplying both sides by $\sum_{i=1}^{I} n_{i,t}^j$, and taking the fact that $\sum_{i=1}^{I} n_{i,t}^j = \sum_{i=1}^{I} (n_{i,t+1}^j - n_{i,t}^j)$, we obtain Equation (1).
9.2 Appendix B: Two Possible Decompositions Following Dunn (1960)

Equation (4) is the rewriting of the decomposition proposed by Dunn (1960):

\[ g_{t+1}^j - r_{t+1} = \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) r_{i,t+1} + \sum_{i=1}^{I} \omega_{i,t}^j (g_{i,t+1}^j - r_{i,t+1}) \]  \hspace{1cm} (15)

By adding and subtracting \( \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) g_{i,t+1}^j \) and \( \sum_{i=1}^{I} \theta_{i,t} (g_{i,t+1}^j - r_{i,t+1}) \) to Equation (15) we obtain

\[ g_{t+1}^j - r_{t+1} = \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) r_{i,t+1} + \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) g_{i,t+1}^j 
- \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) g_{i,t+1}^j + \sum_{i=1}^{I} \theta_{i,t} (g_{i,t+1}^j - r_{i,t+1}) 
- \sum_{i=1}^{I} \theta_{i,t} (g_{i,t+1}^j - r_{i,t+1}) + \sum_{i=1}^{I} \omega_{i,t}^j (g_{i,t+1}^j - r_{i,t+1}) \]

After rearranging the terms,

\[ g_{t+1}^j - r_{t+1} = \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) g_{i,t+1}^j + \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) (r_{i,t+1} - g_{i,t+1}^j) 
+ \sum_{i=1}^{I} \theta_{i,t} (g_{i,t+1}^j - r_{i,t+1}) + \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) (g_{i,t+1}^j - r_{i,t+1}) \]

and, finally, since two terms cancel out, we obtain another decomposition:

\[ g_{t+1}^j - r_{t+1} = \sum_{i=1}^{I} (\omega_{i,t}^j - \theta_{i,t}) g_{i,t+1}^j + \sum_{i=1}^{I} \theta_{i,t} (g_{i,t+1}^j - r_{i,t+1}) \]  \hspace{1cm} (16)

which is Equation (5), a decomposition that yields different values for the industry-mix and the competitive effects compared to Equation (4), if the region and the country have different economic structures and growth rates.
Avenues for further research

In this doctoral thesis I have studied several aspects of structural change. Though scholars have long been interested in this major topic, with a real revival over recent decades due to the movement of deindustrialization that has particularly affected the world’s most economically successful countries, its economic analysis within a rigorous multi-sector growth framework is still relatively young. As a consequence, the open questions are many and varied. Let me highlight some in what follows.

While human capital has been identified as a major driver of economic growth in advanced countries, only a few attempts have been made to explore the interactions between structural change, a phenomenon narrowly linked to economic development, and the accumulation of human capital. Literature largely focuses on the role of physical capital accumulation, thus leaving the question about the role of human capital in affecting structural change largely unanswered. This calls for further work in that direction, especially as there is strong evidence that rises in GDP per capita are associated with a systematic shift in the sectoral composition of the economy towards sectors that are intensive in high-skill labor, an evolution sometimes labelled as “skill-biased structural change”.

In most of recent literature, structural change arises as an efficient equilibrium outcome by construction, so that public intervention can only be harmful. As a result, recent literature on structural change has been exceptionally silent on policy issues. Amending the standard models of structural change by relaxing the strong usual assumptions that lead to the efficiency of the equilibrium (perfect competition, perfect mobility of factors across sectors, exogenous technological progress at the sector level, etc.) may therefore be very helpful for industrial policy purposes. By way of illustration, it is well established that the mobility of labor is imperfect, both sectorally and geographically, thus suggesting that the process of structural change is costly. In line with this reality, it would be particularly interesting to analyze how
public policies should react to absorb the costs associated with structural change.

Another set of open questions is related to the effects of structural change, for instance in terms of aggregate productivity, growth and employment. While recent literature, both theoretical and empirical, has devoted a great deal of effort to investigate the causes of structural change and, more particularly, of the process of deindustrialization over the last few decades, it has surprisingly paid less attention to the macroeconomic consequences of structural change. Existing literature is moreover largely inconclusive, giving rise to conflictual stories about the role of structural change and deindustrialization in driving the recent economic performance of advanced countries. Yet understanding the impact of deindustrialization is essential to ensure the appropriate policy response.

Finally, one could also mention the need for studies analyzing structural change out of the familiar sectoral trichotomy agriculture, manufacturing, and services. There is indeed large heterogeneity, for instance in terms of demand patterns, technology and intermediates, within these three broad sectors, but especially within the services. This heterogeneity has been largely ignored in previous research. In this perspective, the availability of increasingly disaggregated data on economic activity at the sector level should help to refine the existing results, as well as to promote the use of other sectoral taxonomies that have the potential to deliver original information and to reveal new patterns of structural change.
Bibliography


Broadberry, S., Campbell, B., and van Leeuwen, B. (2013). When did Britain industrialise? The sectoral distribution of the labour force and labour productivity in Britain, 1381-1851. Explorations in economic history 50, pp. 16-27.


Craighead, W.D. and Hineline, D.R. (2015). Current Account Reversals and Structural Change in


so ciales, Fonds social européen, Commission européenne, Bruxelles.


208


211


218


222


