What Drives Innovation and Productivity?
A Case Study Using Data for German Firms

Mazen MANSOUR

A thesis submitted to

University of Gloucestershire

in accordance with the requirements of the degree of

Doctor of Business Administration

in the Faculty of Business

October 2017

To my small and big family …
To their great consistent encouragement and unconditional support, my deeply gratefulness.

Acknowledgements

I would like to express my gratitude to all those who helped me during the development of this thesis.

First of all, I would like to express my deep appreciation to my supervisors, Dr. Xiaoling Hu and Professor Nong Zhu, for their continuous help and patient guidance throughout the writing of my thesis.

I want to thank the examiners Dr. Jun Li and Dr. Babak Jafarizadeh for their invaluable suggestions and contributions to improve this work. I would also like to express my gratitude to the chair Dr. Robin Bown.

I would also like to express my thanks to Dr Philippa Ward for her strong support, to the academic stuff at the University of Gloucestershire who accompanied me on my academic journey, and to my supervising panel members for their comments and assistance.

Finally, I am very grateful to Dr. Sandra Gottschalk and the team at the Centre for European Economic Research (ZEW) for allowing me access to the Mannheim Innovation Panel (MIP) and for their support in clarifying all queries related to the dataset.

Abstract

This work attempts to explain the relationship between innovation expenditure, innovation outputs, and firm productivity. It investigates the key factors that drive these relationships using unbalanced German manufacturing panel data at firm level captured by the Mannheim Innovation Panel (MIP) between 2003 and 2013. A structural equation model is employed to test the data consisting of three stages proposed by the Crepon, Duguet, and Mairesse econometric model (CDM) framework.

The first stage is a Heckman model to control for selection bias and to explain the firm's decision if participating in innovation activities or not, and the level of expenditure on innovation in relation to its previous labour productivity. The second stage is the knowledge production function in which innovation expenditure generates economically valuable knowledge in the form of different types of innovation. The third stage is the production function, which describes the relationship between generating innovation and labour productivity. This work focuses on testing the CDM and the expansions on process innovation and organisational innovation in the production function using the Principal Component Analysis (PCA) approach.

The results imply that the firm's decision to involve itself in innovation activities is positively associated with its previous labour productivity. However, for those firms which participate in innovation activities, the previous labour productivity affects the level of expenditure on innovation negatively. The estimation results of the knowledge production function suggest that product innovation in the form of new to the firm of clearly improved products rises with innovation expenditure. The estimation results of the production function promote the role of presenting market novelties, process innovation targeted at the reduction of average costs, and organisational innovation as sources for labour productivity. A set of determinants that might affect innovation and productivity were investigated. The empirical results suggest that market novelties are driven by qualified personnel, however, this study was unable to find drivers for process and organisational innovations.

Keywords: Productivity, Innovation, CIS, MIP, CDM, German Manufacturing, Structural Equation Model, Selection bias, Principal Component Analysis, Panel Data

Declaration

I declare that the work in this thesis was carried out in accordance with the regulations of the University of Gloucestershire and is original except where indicated by specific reference in the text. No part of the thesis has been submitted as part of any other academic award. The thesis has not been presented to any other education institution in the United Kingdom or overseas. Any views expressed in the thesis are those of the author and in no way represent those of the University.

Mazen Mansour

07.07.2017

Table of contents

List of figures

List of tables

Chapter 1

Introduction

It is common knowledge that different countries have different standard of living, based on the real output per capita for each country. Notably, the average real output in the United States, Japan or Germany is about twenty times higher than those in Kenya or Bangladesh (Romer, 2012).

Since the difference in outputs strongly affects almost all aspects of human welfare, economic research has been trying to understand the main sources of economic growth in the long-term (Romer, 2012). Robert Lucas (2002, p.21) said: 'Once one starts to think about (economic growth), it is hard to think about anything else'.

The aim of this chapter is to demonstrate the importance of the subject under investigation and establish the link between the motivation behind the work and its contribution to knowledge. Section 1.1 presents the background and the motivation for conducting this study. Section 1.2 introduces the evolution of economic growth models used in the modern macroeconomics, which have been contributed to promoting the relationship between productivity and innovation. Section 1.3 highlights the meaning of the knowledge- based economy and the importance of knowledge. Section 1.4 provides a short overview of the historical trend of productivity in Germany, compared with other developed countries. Section 1.5 addressed the main characteristics of this research, starting from the research aim, and moving to present research questions, research objectives, and summarising the work's contribution to knowledge. Section 1.6 presents an overview of the structure of this work and an aggregate summary of its contents.

1.1 Motivation

When I moved to Germany in 2005, I lived in the North Rhein-Westphalia federal state to obtain my master's degree in electronic engineering. At that time, globalisation had started to represent the development of world economics and its consequences were clear for the whole region. Coal mines, heavy steel industry and diverse plants that made Germany one of the world's top three industrial countries during the post war era, were abandoned. After the fall of the Berlin Wall, rapid outsourcing and the moving of production and plants to the Far East began.

Spiegel (2004) published an article 'Globalization's Toll, Goodbye Made in Germany' announcing that fundamental changes to the German economy since the mid-1990s mean that Germany is just starting to feel the serious downside of globalisation. The trend is for German firms to shift their production to the competitive low wage regions in Eastern Europe and the Far East in both the manufacturing and service sectors. The backbone of the German economy are Small and Medium Enterprises (SME) or what are known as 'Mittelstand'. These firms are world leaders in their market sectors even though they are more broadly unknown (Sinn, 2004). According the Institute of the German Economy (DIW), at least 60 per cent of German SMEs have already established plants outside the EU-15.

On the other hand, Spiegel (2004) also presented a promising future in the fields of high-tech, health, energy, and biotechnology, in which the number of firms in Germany are growing. As an example, they took a success story from BASF, which registers an average of five new patents a day. Hamilton and Quinlan (2008) consider that Germany had benefited from globalisation by penetrating the markets of developing countries and the United States, capital flow, labour mobility and the movement of ideas and investments.

Furthermore, an important aspect to be considered here is that in general economies show a clear trend of 'continuously and gradually shifting from tangible manufacture- based structures to intangible knowledge and service based business models' (Squicciarini and Torsti, 2008, p.12). In this context, Audretsch and Lehmann (2015) pointed out that an economy can be 'either a high-wage knowledge economy or a low-wage manufacturing economy'. Germany, which was known in the nineties of the last century as 'the sick man of Europe' was able to develop a unique combination of both economic models, and overcome the worst economic crisis after the Great Recession to become a 'Wirtschaftswunder', or economic miracle, which deserves to be analysed.

The reasons that Germany remained one of the world's leading economies despite globalisation were a strong motivation for starting this study, in order to try to understand what factors make German firms productive and successful. Moreover, it was a clear sign that a new era has started, in which production gained a new definition and there was a move from the 'Classical Economy' to the 'Knowledge Economy', in which the product good is the 'idea' that can be sold, exchanged and increase growth.

1.2 Productivity and Economic Growth

1.2.1 Economic Growth

Due to its unquestionable importance, economic growth has been a subject of research for a very long time. Economists agreed that the target of economic growth theory is to answer the question 'how to keep the economy going?', but they disagreed in approach and about whether this growth should be in a long-term or short term sense (Drucker, 1986).

The attempt to answer how nations grow and what drives economic growth has been started with Smith (1776). Since then the theoretical development of economic growth has moved on, and in the last half of the 20th century identified technology and technical progress as a main driver for fostering long-run growth (Lucas, 2002). Different models describing the development of economic growth were presented to understand the possibilities for increasing growth.

The theory of economic growth in the early 20th century was shaped by two great economists: Schumpeter and Keynes (Drucker, 1986). The Schumpeterian theory of technical innovation developed in the early 1910s opened a new research path to explain economic growth and economic cycles. Schumpeter (1934) emphasised the role of innovation and creative destruction in reallocating resources from old to new entities, which made yesterday's capital equipments and investments obsolete.

The Schumpeterian economy of innovation disproved the meaning of profit as an added value 'Mehrwert' stolen from workers as presented by Marx (1867), because it rationalises profit made by innovators as the only source of income and employment. Economic development was the contentious evolution of capitalist society, where capitalism is a method of economic development that 'innovates' the economic structure through the process of creative destruction (Schumpeter, 1942). This creative destruction forces businesses to use profit to bear costs caused by maintaining the future and keep economy going. In this context, 'capitalism' can be considered a moral system (Drucker, 1986).

On the contrary, Keynesian economics considered innovation as 'outside catastrophic, like earth quick or war', which is not a part of economics but which does influence it. Another important deviation between the Schumpeterian economy and the Keynesian economy is that they have a different understanding about what the 'real economy' of goods and services is versus the 'symbol economy' of money and credit. For Keynes, the symbol economy is real and goods and services are shadows of it (Drucker, 1986). Furthermore, labour productivity, worker educational level, quality, technology and technological change were not major issues in the Keynesian economy (Galbraith, 1994).

In this context, I will briefly introduce the modern economic growth models: the neoclassical growth model and the endogenous growth model:

The Neoclassical model of economic growth was proposed by Solow (1956), who was awarded the Nobel Prize for economics in 1987. Solow noticed that not all of the output of the Cobb-Douglas production function (Cobb, 1928) may be explained using the traditional input of labour force and physical capital stock. The un-explained part of growth represents technical and technological change as an exogenous force driving growth, which is called 'Solow's Residual' or Total Factor Productivity (TFP). Technical change may come from Research and Development (R&D) activities that may generate new knowledge in the form of either new products or processes. It also may come from foreign knowledge, which is generated in other countries and used in domestic enterprises (OECD, 2001b). The model proposed by Solow (1956) is referred to as an **exogenous** model of growth because it assumes that technology is determined by forces outside the economy.

Furthermore, Solow (1987) argues the discrepancy between investment in information technology and automation, which is expected to boost labour productivity and the resulting productivity output is a 'productivity paradox', saying: 'You can see the computer age everywhere but in the productivity statistics'.

The next development in economic growth theory was the New Growth Theory, which is also known as **endogenous** growth theory. Romer (1990) noticed that the Neoclassical growth theory described by Solow (1956) with constant exogenous rates of technological change is incomplete to explain long-term economic growth. Furthermore, Romer (1992) complained that the Neoclassical theory ignored the impact of 'ideas' on economic growth.

According to Romer (1990), technological changes drive capital acquisition and is thus one of the most important factors in terms of workers' growth output. These changes are stimulated by market conditions and are the result of deliberate choices. The Neoclassical exogenous theory of growth did not seek to explain why technology improved over time. Therefore, the New Growth Theory assumes that technology grows inside the growth model because it is a product of economic activities (Cortright, 2001).

Romer (1993, p.345) states that 'no amount of saving and investment can generate sustained economic growth unless it is accompanied by the countless large and small discoveries that are required to create more value from a fixed set of natural resources'. Hence, the New Growth Theory emphasises that economic growth is the result of increasing returns due to an increase in knowledge rather than labour or capital (Cortright, 2001).

Romer (1990) proposed an improved model that considers externalities in the accumulation of knowledge depending on inputs such as labour, capital, and an economy's stock knowledge, which rise over time. Furthermore, the model assumes that knowledge and ideas can be shared, reused and accumulated without limit.

The second wave of new growth theory promoted innovation as a source of growth (Helpman, 2004). A long time ago, Schumpeter (1942) promoted innovation as a basic driver of growth and comprehensively described the role of innovation by distinguishing clearly between the real innovators and the imitators. Aghion and Howitt (1998) present a model based on the Schumpeterian approach to explain endogenous growth. In this model, growth is created by a random sequence of innovation resulting from research activities to improve quality. Since new inventions make the previous technology obsolete, the model of Aghion and Howitt (1998) is based on creative destruction and has a Schumpeterian nature. Thus, the firm that succeeds in innovating monopolises the good sector until being replaced by the next innovator. The major differences in Schumpeter's theory compared with standard theories of firm behaviour is that it firstly recognises heterogeneity amongst producers; and secondly it underpins the continuous evolution of a firm's population caused by entry, exit, expansion and contraction as a key element for the creation of new products, processes, and markets. (Cotis, 2004).

Another relevant contribution is the endogenous growth model put forward by Lucas (1988), in which human capital accumulation increases the productivity of both labour and physical capital and drives economic growth. In this model, people divide their time between work and training, but training improves their knowledge, and hence raises their future productivity and future wages.

Table 1.1 summarises the above discussed economic growth theories and the main characteristics of their models.

Table 1.1 Overview of modern economic growth models

Growth theory	Growth is driven by	Model property	Market structure
Neuclassical	Technological change ’Technology’	Technology (exogenous)	Competitive in price
New Growth	Knowledge ’Ideas’	Knowledge (exogenous)	Monopolistic competition
		Technology (endogenous)	competitive to keep monopoly in product characteristics

1.2.2 Productivity

Macroeconomic Level

Productivity is considered essential for economic growth and rising living standards because a worker produces more output for the same amount of work (Atkinson and Wial, 2008). The Cobb- Douglas production function (Cobb, 1928) presents a simplified framework for analysing growth at the aggregate level by modelling the process of converting inputs like labour, capital, machinery and materials into outputs like goods and services (Jorgenson, 1991). One of the most widely used productivity measures on a macroeconomic level is labour productivity as Gross domestic product (GDP) per worked hour of labour (OECD, 2001a). Increased productivity enables companies to generate increased output while maintaining the same input level, resulting higher GDP in the long term for closed economies. Therefore, Krugman (1997, p.11) stated that ’A country’s ability to improve its standard of living over time depends almost entirely on its ability to raise its output per worker’.

The role of productivity in the economics literature as a primary determinant of a country’s economic growth and life standards is very clear. Therefore, policy makers were interested to raise productivity (Mankiw, 2007) to sustain the run of economic growth in the long term. In turn, the aggregate productivity of an economy depends on the productivity of all firms in that economy (Bloom et al., 2013).

An additional aspect should be considered in the German case, in which labour productivity earns an additional recognition because of the shrinking demographic trend in Germany. Because labour force participation is decreasing driven by the dramatic decline in birth rates and the distinct ageing of society (Weidmann, 2014), labour productivity obtains an important meaning so as to cushion the negative impact of demographic change on economic growth.

Microeconomic Level

On the microeconomic level, a firm's productivity is a reflection of the way they use available inputs, like physical capital and human capital, which takes the operational environment like macro-economic fundamentals, market competition and firm dynamics into consideration (Raes et al., 2004). The Cobb- Douglas production function creates the link between the entire economy and microeconomics (Jorgenson, 1991). Here, labour productivity is the most common measure of productivity (Atkinson and Wial, 2008). It is the value that each worker creates per unit of his or her input, which has two determinants: Firstly, human capital, which represents the accumulation of workrer's knowledge and skills. Secondly, technological change as combination of invention and innovation (Cotis, 2004).

Because labour productivity is considered a proxy of firm performance (Belderbos et al., 2004), factors that positively affect labour productivity have a positive impact on the success of the enterprise and drive its performance. The rapid acceleration of globalisation brought big changes and new challenges like facing new competitors and diversification of demand (Nguyen and Martin, 2010). These challenges stimulate productivity to reduce per unit cost, and sustain a firm's profitability and competitiveness (Fallahi et al., 2010). Furthermore, the rapid development of information technology associated with globalisation has provided firms with new opportunities to organise their production processes by outsourcing or offshoring and entering new markets. These new operational conditions create new challenges to keep firms competitive and assert themselves in the market. Hence, it is extremely important to achieve and maintain a high level of productivity to survive.

1.3 The Economy of Knowledge

The term 'knowledge-based economy' recognises the role played by knowledge and technology in driving productivity and economic growth in the next era (OECD, 1996, p.9), based on history, institutions and geography (Cortright, 2001). Knowledge is a non-rival good, which can be embodied in technology and accessed by everybody, but economically valuable knowledge can be used only by its owner to obtain benefits. Van Ark et al. (2009) and similarly Crass et al. (2015) addressed three pillars on which intangible assets of the knowledge economy are based on:

- Economic competency: investments in human capital such as employee training, investments in organisational structure, and brand equity as achievement of marketing research. This is the largest intangible investment in the US, the UK and France, and the second largest in Germany.
- Innovation property: investments in scientific research and development, investments in mineral exploration which leads to increase future sales, investments in copyright protected works and license costs, development of new finance products, and investments in architectural and engineering designs. This is the second largest intangible investment in the US, the UK and France, but the largest investment in Germany.
- Computerised information: investments in computer software and database development. It is the smallest investment in all mentioned countries.

The New Growth Economy emphasises that sustaining growth can be done by investing in activities that create new knowledge such as research and development (R&D), improving the education system and openness to trade. It supports economists in understanding the rapid transition from a resource-based economy to one based on knowledge, and the new shaping of growth on macro and micro levels (Cortright, 2001). Similarly, Hall (2011a) argues that the structural changes in advanced economies has shifted from manufacturing towards services, so that pioneer economics concepts like 'technical change' and 'R&D activities' partially describe the sources of productivity growth and arouse the interest in innovation as a source of growth.

In the 'economics of ideas', ideas are economic goods that can be produced, shared, reused or accumulated without limits, and knowledge has become a strategic assets for successful enterprises (Squicciarini and Torsti, 2008). In the economics of ideas, economic development can be achieved through two strategies: using ideas and producing ideas. Ideas in a limited physical world are classified as big ideas or discoveries, and millions of small ideas that ensure persistence of economic growth. Furthermore, deviating from physical objects, ideas are not scarce and discovering new ideas is not affected by diminishing returns (Romer, 1992).

The second wave of the New Growth Theory promoted innovation as a main source of productivity growth (Van Ark et al., 2009). Therefore, over the last two decades economists tried to understand the relationship between innovation and productivity and went beyond innovation to search for the determinants that drive it.

Innovation is not only a new technology of production but also has intangible effects in the fields of organisation, management, global corporations, new marketing strategies and education (Rao et al., 2001; Van Ark et al., 2009). Furthermore, innovation must be supported by regulatory changes to ensure the condition of innovation creation, investment in education and training on an individual and public level, and to foster research.

The 'Innovation Economy' (Hall et al., 2009) positions knowledge, technology, entrepreneurship, and innovation in the centre of the economic growth model and calls for a 'smart public-private partnership' to achieve higher productivity. In innovation economy, innovation and invention are the source of technical change, create knowledge and spillover, and diffuse to entities that were not involved in the original creation (Hall et al., 2009). In addition to hardware and software, the human brain 'wetware' is a key factor in the innovation economy and therefore increases the importance of non-manufacturing knowledge-intensive service sectors in that economy.

Helpman (2004) insists on the relationship between productivity and knowledge, in which the productivity of a firm depends on its private stock of knowledge and on the economy's aggregate public stock of knowledge. The private stock of knowledge contributes to the public stock, which raises every one's productivity.

1.4 The Historical Trend of Productivity in Germany

The historical trend of labour productivity growth in Germany has different phases. However, Germany's economic success and the international recognition of it makes it easy to overlook the face that productivity growth in the German economy is slowing down (Schneider, 2013).

After the Second World War, between 1945 and 1950, Europe was deeply involved in reconstruction efforts which caused rapid growth and high labour productivity. Technology imitation, increased innovation and the creation of new institutions enabled productivity in Europe to catch up with the United States after the Depression of 1930. At that time Europe had a relatively well educated population and a set of institutions that generated the needed human capital to drive this growth revolution by absorbing new technologies developed by other nations (Van Ark et al., 2009). West Germany was rapidly catching up with other European countries like the United Kingdom or Sweden; productivity growth increased to 5.7 percent per year which was more than twice as faster as in the United States (Van Ark et al., 2009). Because of the international assistance received through the Marshal plan and the stable economic political system, and through cooperation between the financial sector, industry and worker's movements, the German economy became strong and modern.

Between 1973 and 1995, labour productivity in Germany was about the same as in the United States and the gap continued to narrow; however, Germany's strong educational system and highly skilled labour kept the country in a good position in the medium high-tech manufacturing sector (Van Ark et al., 2009). German reunification in 1990 was seen as a reason for further productivity slowdown in Germany. The OECD (2001b) shows that Germany's Multi Factor Productivity (MFP) growth between 1980 and 1998 was lower than the The Organization for Economic Cooperation and Development (OECD) average. This was not primarily related to the Eastern part, where productivity increased dramatically during the first five years after reunification because of the transformation process, but in the Western part related to the decline in employment due to the rationalisation of inefficient enterprises (Van Ark et al., 2009).

Between 1995 and 2007 the divergence between productivity growth in Europe and the United States becomes much wider. Labour productivity growth in the United States accelerated, driven by the industrial capital deepening from investment in Information and Communication Technology (ICT), the increase in innovation and the reduction of semiconductor prices. In parallel, labour productivity growth in Europe slowed down because the implementation of ICT and modernisation of production processes was much slower (Van Ark et al., 2009). In the years 2008 and 2009, the German economy suffered from the global recession but productivity growth continued its low trend (Schneider, 2013).

1.5 Research Characteristics

In this section, I will present the aim of this work, state the research questions and the research objectives, and summarise the gap in knowledge.

1.5.1 The Research Aim

The main goal of the study is to examine the relationship between innovation and productivity in German manufacturing firms, investigate determinants that drive innovation and productivity, and investigate how firms can improve their performance based on the research findings.

1.5.2 Research Questions

To be able to realise the research aim laid out above, this research is expected to answer the following research questions:

1. What is the relationship between innovation and productivity?
2. What are the key determinants affecting innovation and productivity?

1.5.3 Research Objectives

Based on the research questions, undertaking the study involved the following objectives:

1. Investigate the link between innovation and productivity in the German manufacturing sector.
2. Explore the key variables that drive innovation and productivity in Germany.

1.5.4 Research Methodology and Methods

The research follows a quantitative deductive approach and uses secondary unbalanced data for German manufacturing firms collected between 2003 and 2013 from MIP. Using balanced data with less samples would carry the risk of bias caused by selecting only frequently-responding firms. However, to improve data balance, the samples which are present in the data for fewer than four years have been dropped from the analysed dataset.

This research accounts for different econometric issues such as heterogeneity, multicollinearity, endogeneity, simultaneity, and selection bias. It also addresses data issues and explains how to deal with them, such as censoring in the relevant variables. Furthermore, it develops a conceptual framework based on the literature review to implement the understanding gained in an econometric model using the CDM approach. Additionally, it presents the estimation strategy and provides the rationale for the single-equation estimation approach. Finally, the PCA approach is employed to reduce the number of explanatory variables and to account for multicollinearity.

1.5.5 The Gap in Knowledge

The growing body of empirical literature on the relationship between innovation and productivity and the global contributions to this topic confirm its importance. At the same time, they make identifying gaps in the existing knowledge challenging. Nevertheless, four gaps have been identified, which this research aims to close:

The first gap is in the analysed dataset. The most recent works on the relationship between innovation and productivity are based on the CDM framework of Crepon et al. (1998) and tests on German data done by Janz et al. (2004) and Peters (2007) using data from 1998-2000 and 2000-2003 respectively. Furthermore, the latest research of Peters et al. (2013) and Roberts and Vuong (2013) uses panel data for German manufacturing firms up to 2009. This leads to conclude that the German firm panel data between 2010-2013 has not yet been the subject of investigation within the context of the CDM approach.

The second gap is investigating the reciprocal link between innovation and productivity, in which a firm's previous productivity affects its propensity to become involved in activities which may lead to innovation. This was proposed by Raymond et al. (2013) and Baum et al. (2015). The research on this relationship is scarce for German firm panel data. Peters et al. (2013) and Roberts and Vuong (2013) model a mutual dynamic dependency between productivity and R&D activities, which at its core incorporates the firm's decision to engage in R&D activities as a dynamic programming problem using panel data for German manufacturing firms up to 2009, applied to different settings. However, they considered the impact of productivity on innovation decisions but not the innovation expenditure.

The third gap is the failure to include organisational innovation as proposed by Polder et al. (2010), in addition to process innovation as proposed by Parisi et al. (2006) and Peters (2007), and product innovation as proposed in the CDM framework of Crepon et al. (1998).

This combination has not been tested for German firm panel data between 2003-2013.

The fourth gap is that due to the high correlation among various types of innovation, it was impossible to consider them together as inputs in the production function. Therefore, Hall et al. (2012) estimates a firm's predicted probability of innovation and uses it as an input in the production function to proxy innovation. Peters et al. (2013) also takes binary indicators to proxy innovation outcomes. However, this level of abstraction does not allow refined understanding of which type of innovation is most relevant for productivity. Therefore, investigating the impact of different types of innovation, especially product innovation in the form of market novelties and organisational innovation has not been done before with German data.

1.6 Outline of the Work

Chapter 1 explains the motivation for the study and the importance of the investigated topic. It provides a summary of the relevant theories of modern economic growth and the link between economic growth and productivity on the macro and micro level, the rationale of the future trend towards the knowledge based economy. It also reviews the historical trend of productivity in Germany. Most importantly, the chapter introduces the research characteristics such as the research aim, research questions, research objectives, and highlights the gap in knowledge, thus the author's contribution to the knowledge.

Chapter 2 contains a critical review of the previous empirical literature from economics, business and management on innovation, productivity, examines the relationship between them, and investigates factors which may influence this relationship. Moreover, the chapter lays out the theoretical basis of this work by structuring the conceptual framework of the estimation model and the research hypotheses.

Chapter 3 explains the methodology followed in the work, its philosophical position and the rationale for a quantitative research approach. It also presents the research design and the research process.

Chapter 4 discusses the usage of secondary data and panel data, describes the main characteristics of the dataset used, addresses data issues classified according to their source. Furthermore, the chapter lays out the steps taken to prepare data for analysis, and presents the descriptive statistics.

Chapter 5 justifies the econometric methods applied in this research, addresses a set of relevant econometric issues, and propose solutions. The chapter describes the structural model proposed to estimate the relationship between innovation and productivity considering relevant determinants and the dynamic nature of the relationship. Furthermore, it describes the estimation strategy and the methods applied to test the model.

Chapter 6 contains the assessment of the proposed model and provides the empirical results and outcomes of the regression analyses along with some relevant inferential statistics, and the robustness check.

Chapter 7 contains the summary conclusion, which includes the main findings of the research, the major contribution to the knowledge, the implication for the results, the limitations of the work, and suggestions for future studies.

Chapter 2

Previous Empirical Studies

2.1 Introduction

This chapter evaluates the current state of knowledge embodied in the existing empirical literature on the relationship between innovation and productivity pertinent to the research questions. It also addresses major models and theoretical frameworks related to the topic in a detailed manner.

Section 2.2 presents the basic definitions of firm performance and productivity, and the link between them.

Section 2.3 deals with the definition of innovation, how to measure it, describes the innovation process and attributes, analyses the obstacles facing innovation, and identifies the regimes relevant to innovation. Furthermore, this section describes the knowledge production function as a milestone in modelling the innovation process, identifies types of innovations, explains the association between innovation types, and describes the role of knowledge accumulation and ICT in developing innovation.

Section 2.4 provides an overview of previous theoretical and empirical studies that modelled the relationship between productivity and innovation, especially the CDM model and the causality between innovation and productivity.

Section 2.5 discusses determinants that may affect innovation and productivity, such as firm size, the structure of the market, the geographic area of operation, measures carried out to protect innovation, receiving public subsidies, the impact of spillover, partners for cooperation and collaboration within innovation activities, the impact of human capital, physical capital, and financing innovation projects, firm ownership and membership of firms.

Section 2.6 summarises the analysis of previous literature, proposes the conceptual framework used for developing the structural model, and addresses the research hypotheses that will be tested in this work.

2.2 Firm Performance and Labour Productivity

2.2.1 Definition

As discussed in the previous chapter, productivity is key to micro-economic and macro-economic wealth (Peters, 2007). Underlining the importance of productivity on improving living standards, Krugman (1997, p.11) says: 'productivity isn't everything but in the long run it is almost everything'.

Productivity by definition is 'the quantity of goods and services that can be produced from each hour of a worker's time' (Mankiw, 2007, p.12) 'using the minimum necessary level of input to produce a certain level of output in sense of efficiency, using its technological knowledge, its organisation, its size and the operation environment' (Hall, 2011a, p.176).

Productivity as a ratio between output and inputs can grow in two ways: Firstly, by raising the value of produced goods and services. Secondly, by producing goods and services in a more efficient manner. As an example from the discussed topic, the impact of innovation on productivity is that raising productivity is not done by working longer or harder, which is not sustainable in long-term, but by generating product innovation which raises productivity by producing higher value-added products, and via organisational and process innovation that raise productivity by improving the efficiency of the production process (Atkinson and Wial, 2008).

Hall (2011b) states that productivity has to be defined in context. The most common measures of productivity are labour productivity and MFP (the OECD term) or TFP. Labour productivity is the value added per employee or per hour worked. However, MFP is a term which describes a measure that adjusts labour productivity for differences in capital and other inputs e.g. energy, purchased input, and materials. The relationship between labour productivity and MFP is that an increase in output per unit of labour input can be achieved by installing more capital per unit of labour input, which is called the capital-to-labour ratio, or by improving the efficiency with which inputs combine to produce output (Shanks and Zheng, 2006).

Araujo and Costa (2012) pointed out that sometimes productivity is used interchangeably with efficiency; however productivity assumes the existence of technical efficiency. De Loecker and Goldberg (2013) emphasised that the term 'productivity' is different from the term ’firm performance’ or ’profitability’ because the latter depends not only on physical efficiency but also on prices.

As discussed above, labour productivity was the most common proxy for firm performance. However, in addition to this, economic literature also used other proxies to express firm performance: Koellinger (2008) used changes in revenue, employment development, and profitability. Lööf and Heshmati (2002c) and Jefferson et al. (2002) used profitability instead of productivity.

Belderbos et al. (2004) utilised growth values including the increase sales of new products per employee and the rise of labour productivity and Moreno and Huergo (2010) used TFP growth. Klomp and van Leeuwen (2001) used both total sales growth and employment growth. Similarly, Lööf and Heshmati (2002b) proposed using growth versions such as growth in sales, growth in labour productivity, growth in employment, or growth in profit per employee.

Syverson (2011) addressed the main factors that act as levers of productivity: managerial practice and talent, a high quality of labour and capital, R&D, choices regarding firm’s structure, Information Technology (IT) usage, product innovation, and learning-by-doing.

2.2.2 The Production Function

The production function is based on the framework of the Cobb-Douglas production function (Cobb, 1928), which is a widely used function to describe the relationship between inputs, labour and physical capital, and the output that can be produced by those inputs:

$$Y = TFP.L^{\alpha}.C^{\beta} \tag{2.1}$$

Where output Y is the total productivity, TFP, L is labour input, C capital input, α and β are output elasticities of capital and labour. TFP is a measure of productivity when multiple inputs and outputs are used (Araujo and Costa, 2012).

Every measure of productivity is the ratio between output and input(s). As discussed above, the most frequently used measure of productivity is labour productivity, which indicates how efficiently labour is used in production. Syverson (2011) addressed two ways to measure productivity: as index number (Solow residual), or residual of production function estimation. Johansson and Lööf (2009) found that sale per employee is a good proxy for labour productivity. Lööf and Heshmati (2002b) proposed using value added per employee, the share of sales per employee. Peters (2005) used labour productivity by turnover per employee and proposed using both proxies of level and growth rate. However, using growth rate requires panel data to calculate difference over time.

Syverson (2011) states that productivity is a persistent measure because businesses that have high productivity in a given year tend to do so the next year as well and are more likely to survive, but businesses which have low productivity tend to stay low unless they shut down.

2.3 Innovation, Knowledge and Technology

2.3.1 Definition

The term innovation comes from the Latin 'innovatus', which means 'to renew or change' (Arun, 2013). There are many contributing definitions for innovation, which describe it from different perspectives: the Oslo Manual (OECD, 2005, p.46) defines innovation as 'the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organisational method in business practices, workplace organization or external relations'. According to Milbergs (2007, p.2) innovation is 'a process through which the nation creates and transforms new knowledge and technologies into useful products, services and processes for national and global markets leading to both value creation for stakeholders and higher standards of living'. For Morris (2008) innovation is the process by which ideas are generated and transformed in a useful form before being either released onto the market for sale, or implemented to boost a firm's operational efficiency.

The Advisory Committee on Measuring Innovation in the 21st Century Economy's definition of innovation is 'the design, invention, development and/or implementation of new or altered products, services, processes, systems, organisational structures, or business models for the purpose of creating new value for customers in a way that improves the financial returns for the firms' (Schramm, 2008). Soskice and Hall (2001) argues that innovation is 'one of the most crucial dimensions of economic success'. Sengupta (2011) describes innovation as a concept that includes both technology and knowledge capital, in which technology is based on accumulating physical capital while knowledge capital is the accumulation of human capital. Drucker (1999) defined innovation as any action which enables resources to generate wealth, and described two forms of technological innovation. Whereas technical innovation identifies new applications and gives them a new economic value, social innovation creates new administration and management tools in both the economy and society to obtain economic and social value.

Nevertheless, defining innovation and its indicators is a challenging topic for several reasons: Firstly, innovation is not an isolated event but it involves sources of information, organisation, processes, market development, and knowledge generation within the firm or multinational enterprises, or hiring people with knowledge. Secondly, the time scale is an issue in understanding innovation because some innovations need more than a decade before the product enters the market (Gault, 2013).

The Oslo Manual (OECD, 2005, p.150) links innovation to the market through implementation: 'A common feature of an innovation is that it must have been implemented' by introducing it to the market; also new processes, marketing methods, organisational methods are implemented when they be implemented in the firms' operation. A firm is considered as inventor if it introduced at least one innovation during the survey time-frame (OECD, 2005). One of the main challenges in the Oslo Manual is dealing with intangible investment as part of innovation (Gault, 2013). However, Knell and Nas (2006) believe that the methodology of the Oslo Manual must be improved over time to solve problems that appear in innovation surveys and can be avoided only by changes to the survey design.

The theory of economic development described in Schumpeter (1934) is still one of most important contributions, which started a new phase of economic development by highlighting the role of innovation in the business cycle. Firms 'who have carried out a successful innovation can be found in the class of capitalists' (Croitoru, 2012). Schumpeter has been named 'the prophet of innovation' by McCraw (2009).

The OECD (2010b) highlighted the need to understand why and how innovation happens in a firm and what processes are behind it. Hence, it is important to go into more depth to try to understand the process of innovation.

2.3.2 The Innovation Process

Stoneman (1995) summarised the 'Schumpeterian trilogy' using different phases:

1. The invention process phase and the generation of new ideas.
2. The innovation process phase and the development of the new idea into marketable products.
3. The diffusion phase and spreading of the new product across the potential market.

Schumpeter (1934) emphasised the difference between invention and innovation and stated that innovation is the real driver of economic growth because 'as long as they are not carried into practice, inventions are economically irrelevant'.

Morris (2008) visualised the innovation process as a funnel: The big end receives ideas but few ideas can be finished to come out the narrow end to the market. The innovation funnel consists of nine stages: (-1) Strategic thinking, (0) Portfolio management and metric, (1) Research and development, (2) Ideation, (3) Insight, (4) Targeting, (5) Innovation development, (6) Market development, and (7) Sales. As illustrated in figure 2.1, the orange arrow signals the feedback from output back to the input, in addition to the interactions between people in different stages.

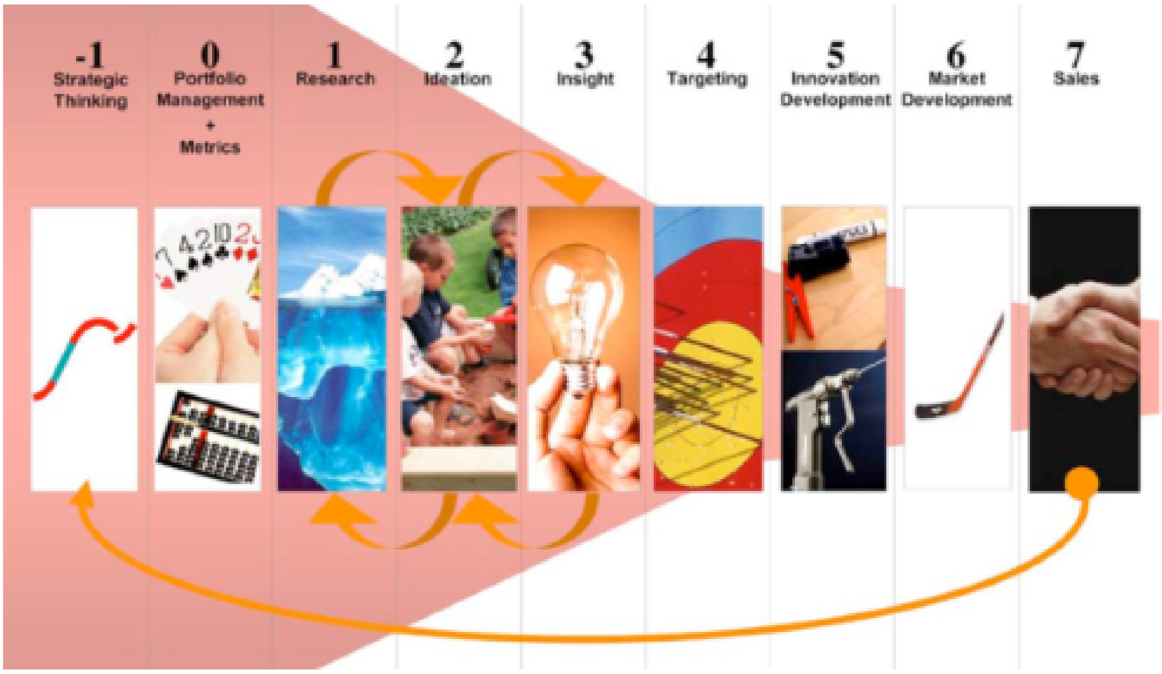


Fig. 2.1 Nine-stage Funnel innovation process (Morris, 2008)

Stone et al. (2008) describe the innovation process in three main stages: (1) Research stage, includes learning and discovery. (2) Development stage, including implementation and demonstrating technical feasibility. (3) Commercialization stage: includes promoting product diffusion and economic returns. Innovation activities involves a combination of inputs in creation of the outputs; the input to innovation could be characterised as resources and assets and the output of former activities becomes the inputs for later processes (Rose et al., 2009). The process and its inputs and outputs are illustrated in figure 2.2.

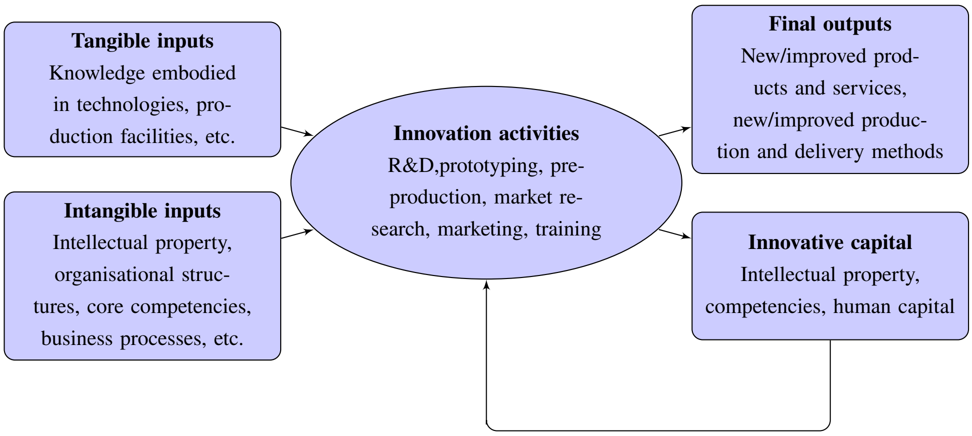


Fig. 2.2 Innovation process (Stone et al., 2008)

Araujo and Costa (2012) viewed innovation as a 'productive process' that has outputs and inputs using the TFP index. The average of TFP value in Portugal is close to that in Germany and higher than that of Sweden, even though Germany and Sweden are classified as innovation leaders and Portugal as a moderate innovator according to the report of the Innovation Union Scoreboard 2010. This result is based on the fact that fewer resources are collocated to innovation in Portugal than in Sweden or Germany, producing fewer innovation outputs.

Innovation activities: are all the scientific, organisational technological, financial and commercial stages leading to innovation being implemented (Stone et al., 2008; OECD, 2005). The aim of innovation activities is to 'create economic value through commercialization', in which customers obtain benefits and innovators obtains returns (Stone et al., 2008). The main innovation activity is Research and Development (R&D).

Frascati Manual (OECD, 2002, p.30) defines R&D as 'comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications'. Investment in R&D has two major effects on innovation: The first effect is the development of innovations and the second is the learning effect resulting from the progressive development of the knowledge stock, which is also called 'absorptive capacity' (Cohen and Levinthal, 1989).

Innovation inputs: can be tangibles or intangibles. Tangible inputs could be physical embodiment, knowledge embedded in technologies and production facilities. Intangible inputs, called knowledge or intellectual assets in business literature, could be intellectual property, competencies and knowledge, R&D activities, organisational structures, and business processes.

Innovation outputs: are the new or improved products and services, and the new or improved production and delivery methods (Stone et al., 2008). Innovation outputs can be tangible or intangible. Tangible outputs are new or improved products and new or improved processes and services. Intangible outputs are the knowledge and competencies gained, intellectual property, and the more experienced human capital and skills that will engage in future innovations.

Helpman (2004) emphasises the macro component of the innovation process in addition to the micro component to ensure the success of the innovation process. Both components interfere and interconnect in such a way that it is difficult to separate them.

2.3.3 Innovation Attributes

Stone et al. (2008) summarise a set of important attributes of innovation:

1. A combined set of inputs is involved in the creation of innovation outputs. The nature of those inputs depends on the required outputs.

2. Inputs can be tangible, and have a physical embodiment and cost, i.e. technology infrastructure, production materials and machinery. Inputs can also be intangible i.e. patents, databases, knowledge, skills of the workforce, which are known in economic literature as 'knowledge assets' and in the business literature as 'intellectual assets'.

3. Knowledge is the key input for innovation. Although more R&D generates more patents, it does not necessarily mean more innovation, because the best ideas come from changes in industry and market structure.

4. The inputs for innovation are assets. Intangible assets are difficult to measure but they are important for innovation. Investments in tangible capital and intangible capital drive innovation, which leads to both tangible and intangible outputs as shown in figure 2.2.

5. The target of innovation activities is to create economic value through commercialisation, with the customer obtaining the benefits and the innovator the returns, in order to develop the innovation process.

6. The innovation process is too complex to be reduced to measurable elements. It is not a simple linear combination of component factors and may happen at any point on its trend.

7. The innovation process is risky so that inputs often fail to produce the desired returns.

8. The outputs in innovation are unpredictable especially before the process of innovation is complete.

9. Knowledge is a key innovation output. Tangible and intangible outputs reflect the knowledge of the resources, technologies and markets of the firm.

Stone et al. (2008) argue that investing in both tangible and intangible capital drives innovation. Although these inputs are recognised as critical, measuring them is a challenging task. Rose et al. (2009) state that reducing the innovation process to measurable elements is not representative because it is complex and not linear. The literature confirms that measuring the innovation process is a challenging task because of the 'innovation uncertainty principle'. The way of measuring innovation may negatively affect the innovation process itself (Morris, 2008). Andrew et al. (2009) argue that 'two elements are routinely under-measured: the first is how fast the company's innovation processes work, and the second is neglecting the measurement of firms' innovation portfolio'.

As mentioned above, innovation is a risky process but engaging in this risk could bring pay-offs of higher productivity levels and rising sale rates. According to Fernandes and Paunov (2012), the chance of surviving the risk of innovation is dependent on different factors: Firstly, the diversity of the product that this firm produced. Firms producing a single product have a greater innovation risk than those which develop several products. If one of them fails, they may rely on the rest to resolve any possible mistakes. Secondly, the market risk represented by factors such market structure, market challenges, the number of competitors, and firms' sales strategy, which affect innovation risk. Thirdly, the technical risk represented by the degree of complexity of the product and product novelty on the market.

Fernandes and Paunov (2012) conclude that pay-offs resulting from risky innovation are not always higher than those from cautious innovation based on credit constraints. However, they see policy makers as having the responsibility to reduce the risk of becoming involved in innovation activities by assets dealing with failed innovations.

2.3.4 Obstacles to Innovation

There are some factors that hamper innovation activities or influence the firms' decision to innovate. D'Este et al. (2012) distinguish between two kinds of obstacles that firms face in carrying out innovation activities: The first type is 'revealed', which describes the firms' awareness about the difficulty of innovation and the learning process resulting from the firms' involvement in innovation activities. The second type is 'deterring', which describes obstacles that prevent the commitment of a firm to enter the innovation context, mainly financing innovation activities and market risk.

Community Innovation Survey (CIS) captures a set of constraints on innovation of financial factors, knowledge factors, market factors, or regulation factors. E.g. the survey of ZEW (2011) captures a set of factors that may hamper innovation such as:

- an excessive perceived economic risk,
- the substantial cost of innovation projects is too high,
- the lack of internal funding sources,
- the lack of suitable external funding sources,
- internal resistance to innovation projects,
- internal organisational problems within the enterprise,
- the lack of suitable specialised staff,
- the lack of technological information,
- the lack of information about the market,
- the lack of consumer acceptance concerning innovations,
- legislature, legal regulations, norms,
- long-winded administrative and authorisation procedures, and
- market control by other established enterprises.

D'Este et al. (2012) found that firms engaged intensively in innovation activities tend to assess obstacles as highly important compared to those which are not involved in innovation activities. However, the result of assessing obstacles is different from one obstacles to another, e.g. firms that have experience in innovation were clearly able to assess obstacles related to knowledge and cost.

2.3.5 Measuring Innovation

Innovation is intangible and has long-term effects. Therefore, measuring it is a very challenging task (Hall, 2011a). In general, innovation outputs are much noisier than innovation inputs because they are subjective (Mairesse and Mohnen, 2010). Innovation has historically been viewed as a residual measure, based on what remains once other growth factors have been taken into consideration (Rose et al., 2009).

A management survey carried out for US firms shows that while the firms recognise the significance of measuring innovation and the need to track it as a business operation they are not sure what to measure and what the scope of measurement is. According to the survey customer satisfaction, overall firm profitability, and the incremental revenue deriving from innovation are the most widely tracked components. Less firms track other components such as time to marketing, idea generation, R&D efficiency, time to volume, portfolio health and performance of life cycle. Furthermore, firms consider themselves more effective at measuring innovation output than tracking innovation inputs or evaluating the quality of innovation process (Andrew et al., 2009).'Bad metrics can lead to bad diagnosis that results in a poor policy with unintended consequences' (Milbergs, 2007).

Based on his description of the innovation process, Morris (2008) proposed a concept for measuring innovation using a set of two types of metrics: qualitative 'soft' metrics that provoke people into thinking about their work, and quantitative 'hard' metrics, which are based on statistical analyses.

Stone et al. (2008) explain how to measure intangible assets using proxies and techniques for indirect measurement. Intangible assets can be categorised into three groups depending on the degree of controllability and ownership:

- Assets which can be owned, controlled, and sold, such as databases or patents.
- Assets that can be owned and controlled, but not sold, such as organisational processes or R&D activities.
- Assets that may not be owned or controlled, such as labour skills and knowledge.

This categorisation clarifies the relationship between activities, inputs, outputs, and outputs that are considered as inputs into multiple activities within the innovation process. Accordingly, Stone et al. (2008) classified intangible assets used in innovation activities and to carry knowledge and skills by mechanism of their development into:

- Human capital, which is the skills and knowledge of individuals working in the firm.
- Organisational capital, which includes the firms' knowledge, intellectual property and the databases owned by the firm.
- Relational skills, which include the firms' knowledge as embodied in supplier, customers, and R&D collaborators.

Rose et al. (2009) proposed two main frameworks to measure innovation using three sources of information: business financial literature to evaluate how firms measure intangible innovative activities, interviews with senior leaders at national firms, and the international contribution to measuring innovation through CIS or collecting data systematically by adding questions to existing surveys.

The first framework is called **Measuring innovation activity** and focuses on measuring intangible capital which is generated and reinserted into the innovation process. This process happens at firm level but can be scaled to the national level. The intangible capital is categorised into three elements: Human capital proxies the knowledge and skills of individuals. Intellectual capital proxies technical inputs through the output of R&D and patents. Organisational capital represents ICT infrastructure, business models and processes which encourage the sharing of information among employees and foster innovation. This framework follows the existing literature but does not include all intangible assets such as brands since it does not feed into the innovation process. In this framework, it is difficult to capture incremental innovation because it may be missing from data collection schemes. The purpose of this framework is to identify data needs and its major advantage is data availability. Furthermore, it records government investment in R&D and ICT.

The second framework is called **Measuring innovation investments** and focuses on how to measure the intangible capital resulting from innovation activities and investments that encourage innovation to occur under specific assumptions instead of measuring innovation per se. Innovation measurement is not the only technological aspect, but it also includes business organization and marketing innovation. The framework has three pillars of investments: human capital in the form of educated and skilled workers, technical knowledge in the form of scientific and non-scientific R&D, patents, and ICT infrastructure that enables employees to organize and communicate information. This framework has the additional advantage of data availability since it includes new innovation sources. However, it considers the innovation process as a closed box. Therefore, it may be better to capture investments by government, non-profits and individuals and business. By measuring education, the framework can capture the inputs from human capital that lead to innovation.

One of the most used frameworks to measure innovation is the Oslo Manual (OECD, 2005), which provides a guideline for collecting and interpreting innovation indicators and composing survey questionnaires. Figure 2.3 describes the framework from the perspective of the firm and reflects how it guides the innovation surveys by target, which is one of the strengths of the Oslo Manual compared to other innovation measurement frameworks. The main topics of the Oslo Manual (OECD, 2005) are: the framework in which the company operates, as well as innovation within the firm, connections with other firms and/or with public research bodies and the role played by demand. Related innovation surveys were developed to collect information about types of innovation in firms, reasons for innovating or not, cooperation with public research sector, flow of data, and quantitative data on sales resulting from product innovation (OECD, 2010a). Furthermore, it shows the integration of firm-based innovation activities with innovation as a system. However, Donselaar et al. (2004) state that taking the innovation system as a whole into account is hampered by the lack of relevant data.

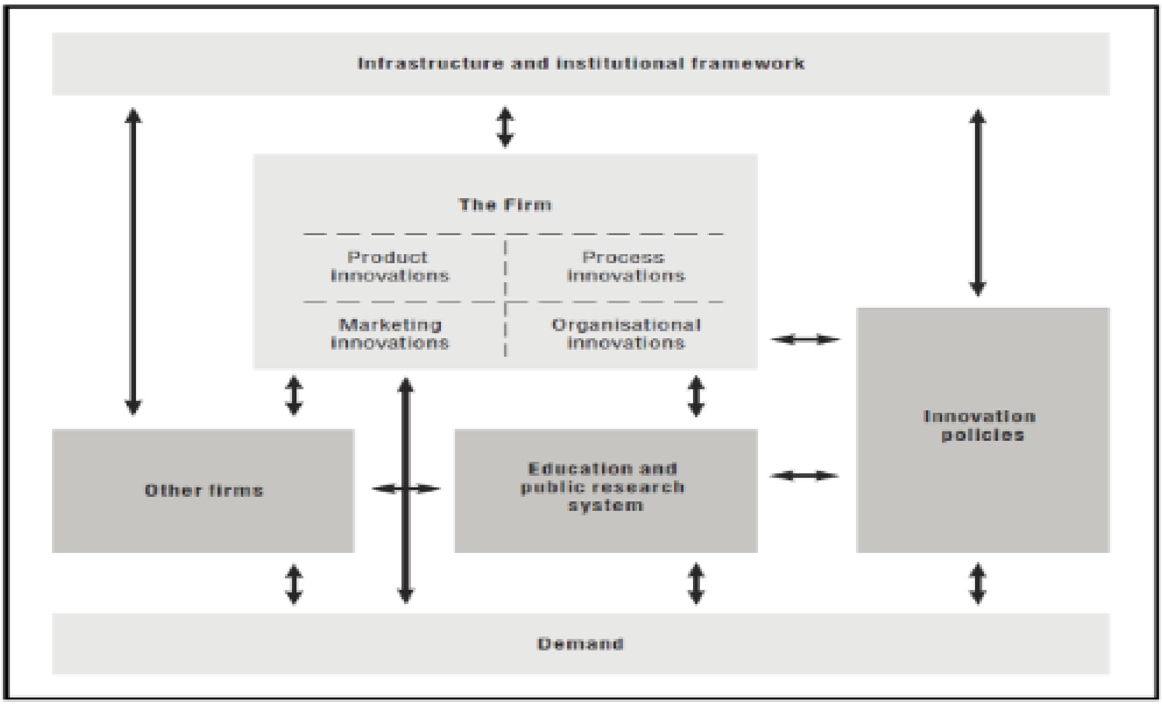


Fig. 2.3 Oslo Manual Innovation Measurement Framework (OECD, 2005)

Table 2.1 demonstrates an example of the evolution of innovation metrics by generation as presented by Milbergs (2007).

Table 2.1 Evolution of Innovation Metrics by Generation (Milbergs, 2007)

1^{st} **Generation**	2^{nd} **Generation**	3^{rd} **Generation**	4^{th} **Generation**
Input indicators	Output indicators	Innovation indicators	Process indicators
(1950s-60s)	(1970s-80s)	(1990s)	2000, emerging focus
R&D expenditures, Personnel, Capital, Tech intensity, ...	Patents, Publications, Products, Quality Change, ...	Innovation surveys, Indexing, Benchmarking, Innovation capacity, ...	Knowledge, Intangibles, Networks, Demand, Management techniques, Risk/Return, ...

2.3.6 Innovation Regimes

Schumpeter emphasised the twin forces of innovation regimes in market dynamics by which new firms and industries replace older ones: creative destruction and creative accumulation (Sengupta, 2011). Table 2.3 shows an overview of the main characteristics of each innovation regime.

Creative Destruction: in this regime, firms have relatively low entry barriers to the market, which has a high level of competition and a turbulent environment. New 'entrepreneurs' are the most significant generators of innovations. Schumpeterian creative destruction may affect the dynamic of the market in two ways: If the reaction of 'entrepreneurs' is slow, the selling price of the existing product remains over cost until the demand diminishes, which lets firms fail in the 'perennial gale of creative destruction'. Firms are motivated to innovate to augment efficiency and competence, and consequently force others to exit the market (Sengupta, 2014).

Creative Accumulation: in this regime, firms have relatively high entry barriers to the market, which is dominated by large established firms and is a stable environment. Entry barriers exist due to the importance of the accumulative nature of knowledge and the high cost of innovation (Filippetti et al., 2009). The old firms carry out innovation motivated by incentive for profit expectations and compete with new firms (Sengupta, 2014). This regime is characterised by creative accumulation of knowledge capital that requires risky investment in research and development but also very high profit expectations (Sengupta, 2014).

Table 2.3 Characteristics of innovation regimes (Filippetti et al., 2009)

Characteristics	Creative destruction	Creative accumulation
Innovation driven by	Small firms, new entrants	large firms, incumbent firms
Market structure	Low entry barriers, low levels of concentration	High entry barriers, appropriation
Key technologies	Radical innovations, relevance to applied knowledge	Incremental innovations, accumulated knowledge, formal R&D

2.3.7 The Knowledge Production Function

The knowledge production function describes the evolution of knowledge creation (Abdih and Joutz, 2005). For early philosophers, 'knowledge was pure objective impersonal, explicit and permanent' (Knell and Nas, 2006, p.4). However the current evolution considers knowledge as 'unobservable and controversial' because philosophers are still debating the nature of knowledge and have only basic understanding of factors that affect the learning process, the acquisition of knowledge, and creation of knowledge (Knell and Nas, 2006). Creation of knowledge is a complicated process that involves a marked amount of individuals and organisations and requires coordination and communication (OECD, 2010b, p.208).

Griliches (1979) concludes that the new knowledge generated depends on current and past R&D expenditure. The stock of knowledge generated by R&D activities enters the Cobb-Douglas production function as a separate input (Mairesse and Mohnen, 2004). The output of innovation activities is not only the observed innovation types discussed below, but also the future returns and the increased market value of the innovating firm (Knell and Nas, 2006). Therefore, in the knowledge production function, R&D is not related to measures of economic performance but to innovation indicators as input. Mohnen et al. (2006) described 'innovativity' as a residual from the knowledge production function, similar to productivity in the Cobb-Douglas production function.

On the one hand, the innovation process from the new idea to innovation outputs is a complicated process, and thus the knowledge production function should be estimated as a system of equations. On the other hand, by considering several links in the innovation processes and a simultaneous equation framework, if the variables tend to move together, their behaviour results in simultaneity bias (Johansson and Lööf, 2009).

Knell and Nas (2006) criticised the innovation surveys because they mix sources of knowledge, types of knowledge, and methods of knowledge production while neglecting to examine how the learning process effects how knowledge is generated.

2.3.8 Innovation Expenditure

Innovation expenditure captures the amount of resources provided by the firm to carry out innovation activities. Aschhoff (2013) defines it as the total sum of expenses caused by internal and external activities aimed at developing both product and process innovations, regardless of whether these innovations have been completed or introduced to the market. Rammer and Peters (2013) defines innovation expenditure as the amount of money spent on innovation activities.

Spending on innovation activities such as R&D has been widely used to deliver information about innovation targets and levels. However, micro data from innovation surveys shows that firm may introduce new products without necessarily undertaking R&D (OECD, 2010a). R&D activities are defined as ’creative work undertaken on a systematic basis in order to increase the stock of knowledge to devise new applications, such as new or significantly improved products or processes including software development’ (Aschhoff, 2013; OECD, 2002).

A breakdown of innovation expenditures that are needed to introduce a new product or to introduce a new process includes the following categories (Aschhoff, 2013):

1. Expenditure on R&D needed to expand existing knowledge which is carried out internally.

2. Expenditure on R&D needed to expand existing knowledge which is carried out externally; internal and external expenditure on R&D including investments in equipment and software.

3. Investments in tangible assets needed for innovation projects such as acquisition of machinery, equipments and software, which are considered part of capital expenditure.

4. Investments in intangible assets, such as acquisition of external knowledge, patents, trademarks and rights for intellectual property, which considered as a part of the capital expenditure.

5. Training and qualification activities for the workforce.

6. Marketing activities to introduce new or improved products to customers.
7. Other related innovation activities that could not be grouped under one of the categories mentioned above; this may include feasibility analysis, engineering, concept development, testing, tooling and software development etc.

This breakdown shows expenditure on R&D is a considerable part of overall expenditure on innovation, which is necessary but not sufficient to generate innovations. Lööf and Heshmati (2002a) modified the CDM model to use innovation expenditure rather than only R&D expenditure as an input for the knowledge production function. Gault (2013) had doubts about using R&D expenditure as a proxy for innovation because it excludes many SME that innovate rather than carry out R&D. Zemplinerova and Hromadkova (2012) argue that R&D expenditure is a relatively good determinant to measure innovation input rather than output because not all R&D investment leads to successful innovation. Lööf and Heshmati (2002c) and Janz et al. (2004) use innovation expenditure as innovation input instead of only R&D expenditures. However, Lööf and Heshmati (2002b) found that considering all innovation expenditure instead of only R&D as innovation input does not change the findings.

Firms do not separate capital expenditure for R&D purposes in their financial accounts; hence, it could be difficult to provide accurate data of their expenditure on innovation (Aschhoff, 2013). Across all industrial sectors in Germany, expenditure on R&D composes about 50% of total innovation expenditure (Rammer and Peters, 2013).

(Corrado et al., 2005) classifies intangibles into three groups: First, computer-based information like software and databases. Second, knowledge produced by R&D activities. Third, a firms' competencies such as brands, reputation, and specific human resources. Using this proposal for intangible classification, Aschhoff (2013) tried to provide quantitative estimates as to the total innovation expenditure on tangibles and intangibles using some indicators that can cover the estimation of intangibles.

Aschhoff (2013) noticed that the increase in the innovation expenditure of German firms in the last decade by an annual rate of 3.8 per cent partially reflects the increase in input costs. Salaries grew annually in manufacturing and service sectors by 2.1 and 2.4 per cent respectively; however, prices for goods grew only by 0.2 per cent. An additional factor which increased innovation expenditure was the increased use of external knowledge.

Johansson and Lööf (2009) proposed separating the ordinary labour force and knowledge labour to avoid the double counting of R&D workers wages as an R&D investment. Peters (2005) addresses the problem of double-counting R&D expenditures including human capital because German data shows that 40% of R&D personnel wages goes to non-graduates and, therefore, proposed to use independent R&D personnel.

Monjon and Waelbroeck (2003) argue that not all R&D expenses necessarily lead to an innovation, or a firm might keep its R&D results secret or non-commercialised. Moreover, innovation may occur in non-performing R&D firms as the development of new business models grounded on organisational innovation, or the development of new marketing strategies, or the implementation of new technology that increase firms' performance. Firm data from CIS shows that firms may introduce new products or processes without necessarily conducting R&D (OECD, 2010b).

Johansson and Lööf (2010) used manufacturing and service firms data from Sweden between 2002 and2004 to investigate the impact of R&D strategy, proxied by steadiness of innovation activities, on firm performance. They found that labour productivity is 13 per cent higher for firms that do R&D persistently and 9 per cent higher for firms that do R&D occasionally in comparison with firms do not do any R&D. Johansson and Lööf (2010) concludes that maintaining innovation activities may be achieved by persistent R&D rather than by the size of R&D expenditure.

R&D expenditure seems to be a stable and continuous indicator. Griliches (1998) found high correlation between R&D expenditure in two consecutive years. Similarly, Lööf and Heshmati (2002a) used R&D investments in cross-sectional data for a specific year as a proxy for permanent R&D because firms do not show strong variation in their R&D investments.

Finally, Aschhoff (2013) noticed that R&D expenditure remains constant for the majority of firms for two reasons, even in an economic downturn: firstly, the high entry costs and technical equipment, and secondly the significant part of R&D expenditure which is fixed in the form of salaries and material costs. Rammer and Peters (2013) found that the decision to innovate increases the probability of continuing innovating at a later time. This might be due to sunk costs using existed R&D facilities. Firm size, human capital, and financial resources are all attributes that may affect persistence of innovation. But there are also some unobserved attributes such as technological opportunities and managerial ability which contribute to persistence of innovation.

2.3.9 Innovation Types

Researcher's opinions differ as to which indicator can serve as a proxy for the 'economically valuable knowledge' resulting from innovation activities as output of the knowledge production function, but they agree about the need for a clear commercial value of innovation. Finding a suitable proxy is challenging because innovation outputs, due to their intangible nature, are noisy and subjective (Hall, 2011a; Mairesse and Mohnen, 2010). Two indicators were used as proxy for product innovation: how many patents were registered and the percentage of sales resulting from product innovation. For process and organisational innovation, firms were asked in the CIS survey whether or not they have gained advantages such as improved quality or cost reduction due to carrying out innovation.

Early in the last century, Schumpeter (1934) identified five types of innovation, which still inspire the innovation studies worldwide: new products, new production processes, new sources of supply, new forms of organisation, and new markets. The Oslo Manual (OECD, 2005) classifies four types of innovations that firms may elaborate: product innovation, process innovation, and organisational innovation, as well as marketing innovation. The term 'technological innovation' refers to products or process innovations; and the term 'non-technological innovation' refers to organisational and marketing innovations. Aschhoff (2013) considers non-technical innovations a 'necessary follow-up' to technological innovation activities. By nature and definition, non-technical innovations appear as supporting innovations that aims at increasing both the quantity and quality of technical innovations (Armbruster et al., 2008).

Product Innovation

Product innovation aims to satisfy customer needs by introducing a new good or a new service to the market, or by improving existing ones significantly (OECD, 2005, p.48). Successful product innovations are key competitive factors in the knowledge based economy (Aschhoff, 2013). However, product innovations are risky because they represent new products, which may be rejected by the market (Peters et al., 2013).

Product innovations may differ in their novelty. Dubner (2008) classifies them into four categories:

- **Radical innovation** demands entirely new knowledge and resources and requires significant advances in technology, rendering existing products non-competitive and obsolete. Therefore, this kind of innovation is considered as competence-destroying.

- **Incremental innovation** aims for cheaper, thinner, faster, product or service with more features than existing ones. It is based on existing knowledge and resources in a given company and involves smaller-scale technological changes to keep the existing product competitive in the market. Therefore, this kind of innovation is considered as competence-enhancing.
- **Architectural innovation** involves the restructuring of the elements that make up the product. Therefore, it is considered deeper than incremental innovation.
- **Disruptive innovation** coined by Christensen (1997) is seen as a real innovation and more interesting from a social point of view since it offers a different and interacting vision of the world. Therefore, disruptive innovation is seen as an input to the process of Schumpeterian creative destruction (Schneider, 2017). It presents a new product or service, which is simpler, more convenient, and/or cheaper than the existing products and services and impacts the way customers live at the lower end of the market (Mohr et al., 2010).

Disruptive innovation might start in two types of markets (Christensen et al., 2015): the first type is the low-end foot-holds. In this market disruptive innovation starts as a low-quality and low-cost product or service, which grows rapidly, solving the quality issues while retaining the same low cost. Then, the disruptive product or service threatens the incumbent market players. An example of this type is 'Amazon', which disrupts traditional bookstores by offering convenient online shopping (Grant et al., 2009). The second type of market is the new-market foot-holds. In this market, disruptive innovation creates a market where none existed, gaining non-customers as new customers. An example of this type is 'eBay', which enabled customers who had unwanted goods but were not able to access the traditional auction market, to sell their goods online, grew rapidly, and disrupted the auction market (Grant et al., 2009).

Radical innovations are considered an engine of technological evolution and economic growth because they are much more profitable than incremental innovations (Frietsch et al., 2010). Mohnen et al. (2006) argue that introducing radical innovation depends on different conditions such as spillovers, patent licensing, and conducting formal research. However, introducing incremental innovations depends on the adaptation of new equipments, products, and informal research. Lööf et al. (2001) states that radical innovation has less correlation with recent R&D investment.

Another type of product innovation mentioned by Aschhoff (2013) is product imitation, which is 'new or significantly improved products or services that are introduced by a firm to the market although these products or services were already offered by competitors at the same time'. This innovation aims to expand the product range of a firm in order to cover customer demands.

To proxy product innovations, the empirical literature used three approaches: The number of patents, the share of sales resulting from innovation and if available, both of these.

Number of Patents: Economists have long considered patents and patent statistics a compelling are for examination (Griliches, 1998). Therefore tracking the number of patents is widely used as a tangible measure, especially for innovations of a technical nature, taken over by a third party to avoid internal bias.

On the one hand, Pakes and Griliches (1984) used the number of patents as a measure of innovation output. Nagaoka et al. (2010) also provided an approach to using patents as an indicator of innovation. Hall (2013) sees a patent as a property right to knowledge asset and proposes the number of patents as a proxy of innovation output, weighted by the numbers of subsequent citations they receive. However, she stated that the simple assumption that patents are a stable measure of innovation outputs without understanding the conditions of this measure is not advisable. Furthermore, the technological value of a patent as a radical invention is higher compared with that of an incremental invention. The latter may consist of subsequent incremental inventions with a lower value. The commercial value of a patent is determined not only by its characteristics, but also by the firms' competitors and the market (Frietsch et al., 2010).

On the other hand, Dubner (2008) argues that many firms register trivial patents that yield no business value but increase their number of patents to achieve a virtual increase in their creativity. Dubner (2008) mentioned Motorola as an example, where 10 per cent of the patents drive 90 per cent of the value. Similarly, Zemplinerova and Hromadkova (2012) argue that the number of patents is not a representative measure of innovation output because not all patents have a considerable market value and most of them are inventions with minor market value. Hall (2011a) argues that although patent count reflects the success of invention and R&D activities, this measurement is very noisy because of the different usage and extent of the patents.

Additionally, Pavitt (1988) similarly addressed three sources of bias in patent counts: Firstly, different economic costs across countries, which depends on the market size. Secondly, differences among industrial sectors and technologies, in which the patent works as a protection measure against imitation. Thirdly, differences among firms like considering unimportant innovation, or presenting an innovation under a different name. To compensate for these bias sources, Dubner (2008) proposed an approach involving evaluating the percentage of returns from products existing for less than three to five years.

In terms of newly-implemented innovations, Duguet and Lelarge (2012) emphasise the difference between those which are new to the market, and those which are only new to the firm itself. From this point of view, Dubner (2008) claims that CIS surveys are not representative because they measure new products and services, but although these are new to one firm they may have already been introduced by another firm or another country, or may be 'an old wine presented to the market in a new bottle'. Therefore, the CIS survey may show firms in Spain and Portugal being more innovative than Germany and the United Kingdom.

Gault (2013) cast doubts on using patents as a proxy of innovation since not all firms, mostly industry-sector dependent ones, can protect their intellectual property with patents. Zemplinerova and Hromadkova (2012) mentioned that the intensity of using the number of patents to measure innovation output depends on the industry sector and it is often used in the pharmaceutical industry much more than in other industry sectors. Criscuolo and Haskel (2003) state that importance of patenting against imitation varies between industrial sectors and the type of innovation. Lööf and Heshmati (2002a) also cast doubts on the quality of the data regarding number of patents because firms tend to register patents in general. Mairesse and Mohnen (2010) argue that firms may generate patents that are never introduced on the market, and therefore they did not achieve any economic profit from them. Contrasty, Criscuolo and Haskel (2003) state that not all innovation are patented by firms.

According to the OECD (2010a), the number of patents is useful information to understand a firms' innovation strategy but they cannot be used as a measure to express the full scope of innovation activities.

Innovation Share of Sales: The effect of product innovation on firm performance may appear only at the point of commercial success (Nguyen and Martin, 2010). To adopt the share of innovation sales as a measure for innovation output, researchers used two approaches:

The first approach is to consider the sales of innovative products as a share of sales of all products as in the model of Crepon et al. (1998). Since 2003, the share of sales which are attributable to new products has been captured in CIS data to provide information about the novelty of product innovation (Rammer and Peters, 2013), which weights innovation according to its degree of success (Mairesse and Mohnen, 2010), and makes a more appropriate indicator than patents (Hall and Mairesse, 2006). The share of sales generated by product innovation differs by industry sector due to the product life cycle. For chemical and pharmaceutical industry sectors, products stays on the market for 10-20 years but for the IT and communication industry it does not stay for more than 2-5 years (Aschhoff, 2013). Hall (2011a) argues that the sales share of innovative products covers products and services but does not cover process and organisational innovations, which may lead to a bias towards product innovators and firms conducting R&D.

The second approach is using the share of innovation sales per employee as proposed by Baum et al. (2015). Similarly, Peters (2005) used sales in the last year resulting from products newly developed in the period of the survey as a measure of product innovation, scaled by the number of employees. Johansson and Lööf (2009) used the log innovation sales per employee as dependent variable for innovation output. Nguyen and Martin (2010) uses the percentage turnover of sales resulting from new or improved products as a proxy for product innovation.

Using a mixed approach: As discussed above, the number of patents is noisy and not stable enough to be used as a measure for innovation output. However, the information coming from innovation surveys are qualitative and much noisier than patent data. Furthermore, innovation data is subjective because firms are asked whether or not they introduced innovations (Crespi and Zuniga, 2012; Mairesse and Mohnen, 2010). To improve the proxy measurement of innovation, Hall (2011a) proposed using the share sales of innovative products as a second measure of innovation improves measuring innovation against noisiness and imprecision. Gault (2013) used a mixed approach to proxy innovation based on both R&D expenditure and patents. Criscuolo and Haskel (2003) used both innovation expenditure and number of patents as innovation output. Mairesse and Mohnen (2004) used the share of total sales of innovation and patent protected sales. However, Baum et al. (2015) used the number of patent applications generated in the last year as an input that causes R&D expenditure.

Process Innovation

Process innovations aim to improve economic prosperity by elaborating significant improvements in methods of production, distribution, or logistics that lead successfully to reductions in unit cost or to improvements in product or service quality (OECD, 2005, p.49). This in turn improves the firms' price competition and increases sales opportunities. Process innovation also covers supporting activities such as accounting, purchasing, or computing. Helpman (2004) emphasised that technological change is not restricted to the technology itself but also considers the production processes and product diffusion, taking a long-term view. Furthermore, modifying and developing technologies or buying off-the-shelf technologies that are new to the firm is considered process innovation in CIS data Gault (2013). The quantitative effect of process innovations can be measured by the resulting change in turnover (Aschhoff, 2013).

Process innovations are risky because the expected higher efficiency might not be associated with cost reduction or may result in difficulties in implementation (Peters et al., 2013).

In CIS data there is no separation between expenditure on product and process innovation. Therefore, either R&D expenditure or innovation expenditure can be used as input in the knowledge production function for both product innovation and process innovation. German CIS is the only survey that quantifies process innovation as the percentage of cost reduction in unit cost caused by process innovation. Since 2003 German CIS captures the change in sales achieved by process innovations (Rammer and Peters, 2013). Peters (2005) used reduction in unit cost in the last year caused by introducing new processes in the survey timeframe and scaled by the number of employees as proxy for process innovation.

Finally it is important to mention that process innovation contributes to product innovation by assuring that new technologies and products meet requirements and quality. It also contributes to productivity by reducing production costs. Hall (2011a) argues that process innovation can increase real output at firm level without changing revenue.

Organisational Innovation

Organisational innovation can be defined as 'the creation or adoption of a new organisational method to the firm' (Aschhoff, 2013, p.108). Moreover, it includes introducing new or improved business practices; a knowledge management system; changes in management structure, methods of workplace organisation and different activities; or changes in the organisation of external relations with enterprises or public institutions (OECD, 2005, p.51).

The Oslo Manual (OECD, 2005) defines the main outputs of organisational innovation as a reduction in the response time to customers or suppliers, improvement in the ability to produce technological innovation (product and process innovations), improvement in the quality of goods or services, reduction in costs per unit of output, and improvement in communication and the share of information within the enterprise or among enterprises.

To evaluate whether or not a firm has carried out organisational innovation, CIS asks whether during the survey period the firm has introduced new business practices for organisational task, new ways of organising responsibilities and/or decision making processes, or implemented changes in the way it organises its external relationships with public bodies or other companies (ZEW, 2009). Starting from the year 2005, MIP data has been collected for five indicators regarding introducing organisational innovation in a firm: reduction of reaction time, improvement in product quality, reduction in costs, and improvement in employees job satisfaction. In 2009, the survey asked questions to capture improved organisational capabilities concerning the development of new products and improvement in communication (Gottschalk, 2013).

Armbruster et al. (2008) insists on the high impact of organisational innovation for competitiveness and business performance; however, the success of organisational innovation depends on how the organisational structure responds to the use of new technologies. Furthermore, Armbruster et al. (2008) categorized organisational innovation into two main categories: firstly, structural organisational innovations that change and improve responsibilities, accountabilities, and information flow. Secondly, procedural organisational innovations, which affect routines, processes, and operations within the firm.

Armbruster et al. (2008) view the existing literature as ambiguous in defining the meaning of 'organisational innovation' and summarised three areas of research that handle topics of organisational innovation: in the first area, researchers concentrate on identifying the structural characteristics of organisational innovation e.g. responsibilities and information flows and its effect on technical innovations. In the second area, researchers focus on procedural topics, for instance, how organisations change and evaluating the different types of changes and understanding the resistance to changes within the organisation and how to overcome them. In the third area, researchers focus on the theory of organisational understanding and sustaining creativity in developing organisational innovation at micro level and processes of continuous improvement.

Kangasniemi and Robinson (2008, p.4) defined organisational change as ’the process of organisational reform’, which can raise productivity in two ways: firstly, through introducing adaptation or by improvement of deploying new technologies, and secondly, through optimal utilisation of skilled workers. They argued that organisational changes target to absorb technology and utilise it to improve the firms’ response to growing global competition pressure. These changes may result in internal restructuring of responsibilities and task distribution inside the working team, but they may also result in external restructuring of responsibilities and relationships among firms, or they may result in a combination of internal and external restructuring.

Vickery and Wurzburg (1998) addressed some barriers that may hamper adoption of organisational changes in enterprises due to these conditions:

- External conditions: a lack of demand or inopportune economic outlook.
- Economic conditions: the combination of high costs and a lack of financing.
- Managerial conditions: management has no strategy or sees no need.
- Information related: lack of awareness of whether it works.
- Resource related: lack of skills and management ability to apply change.

Marketing Innovation

Marketing innovation can be defined as ’the implementation of new marketing methods involving significant changes in product design, packaging of goods or services, or changes of sales and distribution methods’ (OECD, 2005, p.49).

Schumpeter (1942) realised that activities such as opening a new market or expanding customer relations are innovation activities and contribute to the firms’ success. Moreover, marketing innovation increases the probability of technological innovations success (Schubert, 2010).

Aschhoff (2013) pointed out that measuring non-technical innovations is challenging due to the lack of a consistent definition, and due to the low frequency of published data for German enterprises. However, a dedicated questionnaire about non-technical innovations has been started and is available. Since 2005, MIP data has covered three indicators of marketing innovation: accessing of new geographical markets, approaching new customer groups, and increasing or maintaining the market share in existing markets (Gottschalk, 2013).

2.3.10 Association between Innovation Types

The interaction among different types of innovation and how this interaction contributes to firm performance is diverse in the literature and may relate to other factors such as industrial sector or firm size. Hall et al. (2012) argued that different kinds of innovations are so highly correlated that they can be modelled as one dimension in the knowledge production function.

Fosse et al. (2013) classify innovation types according to synergies between them: product and marketing innovation as demand innovations, which may lead to better market prices, and process and organisational innovation as supply innovations that result in more efficient and lower-cost production processes. Junge et al. (2012) confirm empirically that demand innovations and supply innovations are two 'distinct and non-overlapping' categories.

Schubert (2010) investigated the complementarity between non-technological innovations and technological innovations using German CIS data and found that marketing innovation affects product innovations by making them more successful on the market and affects process innovation in terms of cost reduction and maintaining a firms' competitiveness. However, Schubert (2010) did not find that organisational innovations had a positive effect on either product innovation or process innovation. Aschhoff (2013) investigated the dependency of occurrence of technological and non-technological innovations in German enterprises and found a strong complimentary relationship between them. Junge et al. (2012) did the same investigation using Danish firms' data and found that product and marketing innovations are complementary inputs that lead to higher productivity growth. However, product and marketing innovations contribute to productivity growth separately from organisational innovation.

Aschhoff (2013) observed the need to combine process and product innovation because firms need advanced processes to implement product innovation. The smaller the firm is, the higher is the consequential need to process innovations. According to OECD (2010b), those firms which implement both product and process innovations achieve around 30% more sales per employee than firms which focus only on product innovation. Polder et al. (2009) found complementarity between product and process innovation in the manufacturing sector but not in the service sector. Polder et al. (2010) tested the relationship between different innovation types using a model from Mohnen and Röller (2003) and found a complementarity of product and process innovations and a substitutability of organisational and product innovation. Furthermore, Polder et al. (2009, 2010) found that product and process innovation lead to productivity gain only in combination with organisational innovation.

2.3.11 Innovation patterns

According to Soskice and Hall (2001), there are two types of capitalism which have different institutional frameworks, inter- firm relations, and length of employment tenures, which result in different innovation patterns:

- Liberal Market Economies (LME) e.g. United States, United Kingdom, Canada, and Australia, which support radically new emerging technologies.
- Coordinated Market Economies (CME) e.g. Germany, Japan, and Sweden, which support incremental innovation.

Soskice and Hall (2001) also found that radical innovation is important in high technology such as semiconductors, telecommunication, and software. However, incremental innovation is important for maintaining competitiveness in the fields of machinery and factory equipments. By analysing data from the European Patent Office, Soskice and Hall (2001) found that Germany specialises in incremental innovation in the fields of mechanical engineering and product handling, in comparison with the United States which specialises in radical innovation in the fields of telecommunications and medical engineering.

Soskice and Hall (2001) noticed that some CME firms may shift some of their activities to LME to ensure access to institutional support for radical innovation. Contrarily, firms in the LME category may shift some of their activities to CME to ensure access to institutional frameworks regarding incremental innovation, quality control, and skill levels.

Hall (2011a) noticed that very few innovations create entirely new markets. Instead, most registered innovations are improvements of existing products. Lööf and Heshmati (2002c) found that the rate of productivity growth is only boosted by radical innovations. Lööf and Heshmati (2002b) found that radical innovations have less correlation with recent investment in R&D and returns to R&D.

Hall (2011a) pointed out the need for a precise definition of 'new' to clarify whether it means 'new to the firm' or 'new to the market' and requested that radical innovation and imitation should be distinguished from one another. Starting from the year 2002, the German CIS survey merged the share of sales resulting from 'new products' with those from 'significantly improved products'. There is no ability to recognize the sale share coming from radical innovation and incremental innovation separately. Alternatively, the sales share of market novelties is captured as an indicator of new products that have not previously been introduced to the market. The definition of 'market' is set by the firm itself, which may be regional, national, or international (Aschhoff, 2013).

2.3.12 Knowledge Accumulation

Analysing sources of productivity growth shows that capital and labour can only explain about half of that growth. The 'residual' referred to technical change and takes the measure of R&D capital stock and the creation of knowledge capital into consideration (Hall, 2011a). Romer (1990) argues that the Solow model of growth with constant exogenous rates of technological change is inadequate to explain long term economic growth.

The economic theory considers knowledge as a stock called knowledge capital (Knell and Nas, 2006). Therefore, Romer (1990) proposed an improved model that considers externalities in the accumulation of knowledge depending on inputs such a labour, capital, and an economy's stock knowledge, which rises over time, in which each firm contributes to its own private knowledge and contributes to the aggregate public stock of knowledge which raise every one's productivity. Therefore, several empirical studies that investigate the relationship between innovation and productivity consider the accumulation of a firms' stock of knowledge capital as a further input in the production function.

The role of knowledge in innovation activities is emphasised in the literature. Stone et al. (2008) addressed two related attributes for innovation that reflect the non-linear relationship between knowledge and innovation:

- Knowledge is a key innovation input. R&D activities may lead to patents but not necessarily innovation because the best ideas come from industrial changes and market structure.
- Knowledge is a key innovation output. Tangible and intangible outputs reflect the firms' knowledge of technologies, resources, and markets.

Abdih and Joutz (2005) states that producing new knowledge leans strongly on the existing stock of knowledge so that ideas developed in the past may supports the creation of new ideas. Knell and Nas (2006) criticised innovation surveys for attributing more importance to newly produced knowledge rather than existing knowledge.

Lööf and Heshmati (2002b) found a positive association between the past and current R&D behaviour, which shows the effect of knowledge accumulation. Lööf and Heshmati (2002a) found that knowledge capital is a key element for firms' performance heterogeneity. Johansson and Lööf (2009) investigated the role of persistent innovation in knowledge accumulation and recognised a correlation between past and current innovation investments. Peters (2007) found correlation between past and current innovation output in patents and innovation sales, in addition to the correlation between persistency of innovation and firm performance.

Griliches (1998) modelled the stock of knowledge and considered the knowledge production as a non-linear equation, because it depends not only on current R&D effort, but also on previous accumulated results. Vieira et al. (2008) found that over time innovation activities have a non-linear effect. Investigating the relationship between activities of innovation, measured by R&D expenditures, and labour productivity in European regions shows a positive exponential relationship in the long term because the return from the investment does not occur immediately but in the medium and long term.

Kwon (2009) argues that learning process of human capital is the most important factor for increasing knowledge accumulation. Rammer and Peters (2013) argue that skilled employees is an additional factor contributing to the accumulation effect of knowledge caused by past innovation experience. Additionally, persistence of innovation behaviour enhances knowledge accumulation because experience in innovation is combined with the effects of learning-by-doing and learning-to-learning, which increase returns (Rammer and Peters, 2013).

Searching for proxies to measure knowledge increments, Pakes and Griliches (1984) investigated the quality of patent counts as an indicator of knowledge increments over time and found that using patents as a quantitative indicator of innovation has two problems: firstly, firms does not patent all new innovations. Secondly, the economic impact of a patent varies strongly from one to another. An advantage is that patents can be used to breakdown the lag in the process of transforming R&D inputs to knowledge and then to benefits into two parts: one which produces patents from research investments, and another which transforms patents (with additional expenditure) into benefits. However they found that patents are a representative indicator of differences among firms in terms of advances in knowledge. Pakes and Griliches (1984) found that the relationship between knowledge and patents is related to different conditions of an economic, technological, and legal nature, and to the cost of the patenting process. Romer (1990) arrived at similar findings and presented a model where firms invest in R&D for developing new products and protect them via patents. The more R&D activities, the larger the knowledge stock, and the lower the cost of future R&D (Helpman, 2004).

Griliches (1994) employed capital data and patent data to estimate the knowledge capital stock. Hall and Mairesse (2006) state that empirical studies of knowledge management at micro level are still at the beginning due to the lack of the detailed data and measurement concepts. Lööf and Heshmati (2002a) used the ratio of innovation sales to total sales as a proxy for knowledge capital.

Knell and Nas (2006) addressed an important aspect of knowledge accumulation. With time the stock of knowledge may become less valuable or obsolete according to the process of creative destruction. Therefore, focus is more on the generation of new knowledge rather than on knowledge accumulation.

Another important aspect that should be considered in this context is **absorptive capacity**, which is 'the ability of a firm to identify, assimilate, and exploit knowledge from the environment' (Cohen and Levinthal, 1990, p.569). The hypothesis of absorptive capacity has attracted empirical research by giving formerly gained knowledge and the learning process an important role. Firms with greater R&D efforts and prior knowledge show a stronger effect of spillover of R&D and knowledge on the firms' innovation output when they cooperate with other firms. Guellec and van Pottelsberghe (2004) emphasise absorptive capability as a source of knowledge that has a significant effect on productivity growth. Harhoff (1990) used German manufacturing data to confirms that firms with high R&D capital stock show higher profits from external R&D and that production experience gives firms the necessary background to absorb new information and processes.

2.3.13 Information and Communication Technology

The economic role of ICT on productivity was presented in the literature firstly in strategic management by optimising the structure of the organisation, introducing new organisational methods, and investing in business processes; and secondly in cost reduction and the improvement of intangible characteristics like variety and quality (Nguyen and Martin, 2010; Bresnahan et al., 2002).

The impact of using ICT was investigated in the 1950s by Solow (1957), as well as in many other later studies. Evaluating the impact of ICT on firm productivity, the literature splits in two main streams:

- One stream considers ICT use to have a direct and positive effect on firm productivity.
- The other stream confirms the importance of ICT for organisational innovation by allowing better quality equipment, enhancing products, processes and organisational structures, improving the relationship between supplier and customer, reducing geographical limitations, and enhancing investment in intangibles, which in turn contributes to the firms' productivity.

In accordance with the first stream, the Eurostat (2008) project underlines the direct impact of using ICT on productivity. Hall et al. (2012) used Italian firm level data to examine the impact of product, process, and organisational innovation on productivity affected by the main factors R&D and ICT. They found that investing in R&D is significant for innovation and ICT investment is significant for productivity. Moreover, Hagen et al. (2008) found a direct link in Swedish data between the use of broadband and ICT, and firm productivity. Van Leeuwen and Van der Wiel (2003) found in Dutch data for the service marketing sector that ICT impacted positively on labour productivity growth by including spillovers of ICT in a production function model. Using Dutch and UK data, Van Leeuwen and Farooqui (2008) found that using e- commerce and broadband connectivity as drivers for ICT directly affects productivity excluding R&D as an input to innovation. Additionally, they found evidence that ICT investments in fast internet depends on ICT capital deepening, whereas the use of e-selling has a direct impact on TFP.

In accordance with the second stream, Polder et al. (2009) argue that ICT investment affects productivity indirectly through organisational innovation. Similarly, Nguyen and Martin (2010) stress the link between organisational innovation and ICT because technological change drives new approaches to organising firms. Nguyen and Martin (2010) found that deploying ICT if combined with organisational innovations leads to the reduction of waste costs, improves product quality, and expands product variety, which may improve labour productivity. Brynjolfsson and Hitt (2000) and Brynjolfsson et al. (2006) established this stream in their empirical work and claim that the organisational change combined with IT investment lead to cost reduction, and therefore, increase productivity.

Considering both R&D and ICT as inputs for different types of innovations, Hall et al. (2012) view ICT as more related to productivity. However, it is argued that ICT alone is not sufficient to impact productivity. Hall et al. (2012) referred to different economic studies that tried to avoid any overestimation of ICT effects on productivity by concentrating on the interactive relationship between ICT, human capital, and organisational change. These studies underline that implementing ICT requires organisational change, management effort, and training cost and time. They assigned two impacts for ICT in this content: the first, is organisational and process related, which leads to cost reduction and therefore contributes directly to the production function. The second impact leads to producing new products and services allows new ways and methods of research, and contributes to knowledge production.

Polder et al. (2009) investigated the relation between ICT and organisational innovation and found that investing in ICT can lead to changes in the organisation of the firm through presenting new services, new approaches of producing goods or services, new approaches of doing business, or new approaches of marketing. Polder et al. (2009) argued that investing in ICT is important for product and organisational innovation in the manufacturing sector and for all types of innovation in the service sector.

Koellinger (2008) argues that using ICT supports innovation because it increases the reduction of transaction costs, the improvement of business processes, coordination with suppliers, and process fragmentation along the value chain across different geographical locations.

Using UK firm panel data, Crespi et al. (2006) found a positive relation between productivity growth and investment in IT if coupled with organisational change. In contrast, they did not find an impact from organisational change with non IT investment on productivity growth. Using Danish data to investigated the impact of ICT on innovation, Fosse et al. (2013) found that ICT induced innovations explaining between 18 and 26 per cent of the productivity growth specifically. Fosse et al. (2013) found that firms which introduced ICT technologies in the year 2007 displayed the highest average growth in productivity over the period 2007-2010 if coupled with organisational changes. Firms which only introduced one of both factors did not achieve statically significant productivity growth.

However, Nguyen and Martin (2010) used a set of indicators to proxy ICT in order to investigate its impact on innovation and found that not all increases in ICT investment lead consequently to new products or new process innovations being introduced nor do they improve the performance of innovation. Koellinger (2008) found using data on European enterprises from the year 2003 that internet use boosts product and process innovation.

Spiezia (2011) found that using ICT enables firms to adopt product and marketing innovation in both manufacturing and service sectors but it did not increases the firms' capability for cooperation or to developing innovations. Spiezia (2011) addressed the question of whether ICT stimulates cooperative effort in innovation and found weak evidence for the association between ICT intensity and cooperation in innovation activities, but he did not find a positive relationship between the intensity of ICT and introducing new-to-the-market products.

Measuring ICT: Challenged by data availability, researchers have tried different proxies to measure ICT. Van Leeuwen and Farooqui (2008) incorporated the share of employees able to work on computers, the use of broadband, and the use of e-commerce as proxies for ICT which were available in CIS data. They confirmed that the use of e-commerce and fast internet connectivity as proxies for ICT capital stock is possible. Polder et al. (2009) used the use of broadband as proxy for ICT investment (as the log per employee). Nguyen and Martin (2010) used a large set of indicators from the annual report on 'ICT usage and E-commerce in Enterprises Survey' on firm level such as the use of intranet, extranet, video-conferencing, e-commerce, usage of electronic forums, and software. Eurostat (2008) argues that these indicators are representative because broadband access allows firms to exchange information internally and with business or development partners, which reflects how advanced the ICT infrastructure in the firm is; the use of e-commerce as an indicator expresses how firms use their ICT infrastructure to sell and purchase goods and services; this affects process innovation in both the manufacturing and service sectors (Polder et al., 2009).

Finally, to avoid overestimation of the impact of capital deepening caused by ICT on labour productivity growth, Van Leeuwen and Van der Wiel (2003) argue that the correlation between the use of ICT in a firm and the increased use of ICT in its environment should not be neglected.

2.4 The Relationship between Innovation and Productivity

The effect of innovation on productivity has been inspected in the empirical literature using different approaches and on different aggregation levels (macro, meso, or micro). Innovations influence firm productivity in two main ways: firstly, innovations create new products and services and may lead consequently to an increased demand for a firms' products. Secondly, process and organisational innovation lead to efficiency gains in production and a reduction in production costs, so that more efficient enterprises will enter the market or maintain their position, and less efficient enterprises will exit the market. Furthermore, innovating firms grow more than others and displace inefficient existing firms. In both cases, the operative institutional and macroeconomic environment plays a deciding role (Hall, 2011a).

2.4.1 Background

The historical trend of studying the relationship between innovation and productivity starts from the classical Cobb-Douglas production function (Cobb, 1928), which details the ratio between inputs and output of production, and the contribution of Solow (1956), which describes the impact of technological change on the production function. Griliches (1979) proposed the knowledge production function to express the relationship between innovation inputs and outputs, similar to the production function. Pakes and Griliches (1984) developed the first model that considers both knowledge production function and the traditional production function to handle the neglected link between inputs and outputs of the innovation process.

The model of Pakes and Griliches (1984) describes the relationship between R&D, innovation, and productivity, presented in three sequential equations: the performance equation is based on the theoretical framework of the Cobb-Douglas production function, which explains the output using the log of standard input variables vector expressed in per employee terms, such as physical capital, human capital, material, and R&D investment. The knowledge production function considers the generation of economically valuable knowledge represented by the number of patents resulting from past R&D expenditure. The innovation input equation represents the investment in innovation. According to Crepon et al. (1998), this model suffers from two issues: selectivity and simultaneity bias. This is due to the fact that explanatory variables are jointly determined along with the dependent variable. They, therefore worked to solve these issues in their own model.

The empirical research is split into two streams. The first stream assumes a direct link between R&D activities and productivity, for example the work of (Griliches, 1979; Harhoff, 1990; Griliches, 1994; Jefferson et al., 2002; Bond et al., 2003; Guellec and van Pottelsberghe, 2004; Belderbos et al., 2004; Khan et al., 2010). The second stream assumes multiple-steps starting from the decision to conduct R&D activities, a link between R&D activities and the production of knowledge, and finally a link between generated knowledge and the production. The second stream is mostly driven by the work of Crepon et al. (1998), which is also known as the CDM model and the improvements made in other later studies.

2.4.2 The CDM Framework

Crepon et al. (1998) reflected on the innovation process starting from conducting R&D activities to the generation of patents and sale of new products, which in turn drive a firms' productivity. The 'Crepon, Duguet, and Mairesse' CDM model demonstrates that firm productivity is driven by innovation outputs such as patents and innovation sales, and not driven directly by innovation input such as R&D investments (Peeters and Pottelsberghe, 2004). Furthermore, the proposed CDM structural four-equation model solves the problems addressed by the model proposed by Pakes and Griliches (1984):

The selectivity issue of innovating firms has been solved by adding a probit modelled selectivity equation to distinguish between firms that do invest in R&D and those which do not invest. By definition, the R&D firms are a selected group, which are committed to investing in R&D, compared to the negative decision of other firms concerning investing in R&D (Bond et al., 1997).

To endognise the R&D and productivity, the disturbances in the four equations are assumed jointly correlated, which means that factors which affect R&D also affect productivity and vice versa. Lööf and Heshmati (2002c) argues that influencing the endogeneity of the R&D variable in the knowledge production function is confirmed by innovation theory and many empirical studies on R&D and productivity. According to Mairesse and Mohnen (2010), endogeneity of innovation outputs in the production function is caused by measurement error rather than by simultaneity.

According to Knell and Nas (2006), the CDM model addresses four econometric issues: the first issue is that of measurement problems due to the relationship between sale as output and material as constant input tending to be inflated. A second issue is that of endogeneity between innovation and productivity, which could be solved using instrumental variables.

A third issue meets the cross-sectional data, which is that the CDM model assumes the error terms in the four equations are not correlated. A fourth issue is that the linearity of the basic CDM model hampers investigation of the non-linearity in the innovation relationship.

The CDM model expresses the relationship between innovation and productivity in three stages. Each estimated output at a given equation is used as an input for the next equation:

- R&D equations: the first equation describes whether or not a firm is involved in R&D activities; the second equation is the R&D intensity equation which relates to the intensity of investment as innovation input to its determinants.
- Innovation equation: also called the knowledge production function, which relates R&D expenditure as innovation input to innovation output.
- Productivity equation: based on Cobb-Douglas production function, which relates innovation output to productivity as log of sales output per worker (labour productivity).

The advantage of the CDM model is that it is expandable to adopt more equations and variables. Therefore, it has been used as a basis for other later empirical works that expanded it to cover other innovation inputs than R&D such as ICT. On the output side, in addition to product innovation the model has been expanded to cover more innovation outputs such as process and organisational innovation have been considered.

Lööf and Heshmati (2002a) presented an intermediate approach between the exogenous knowledge production function of Pakes and Griliches (1984) which neglects possible correlations, and the endogenous model of the CDM approach, which considers full correlation between all residuals. The first two equations of innovation selectivity and R&D investment are estimated jointly, and the other two equations of knowledge production function and production function are estimated simultaneously. Lööf and Heshmati (2002c) argue that the use of the proposed intermediate approach is due to the cross-sectional nature of CIS data. However, in cases of greater availability of CIS surveys to produce panel data, the completely endogenous CDM model is preferred.

Peters (2005) noticed that the CDM model has ignored the process innovation and proposed an expansion of the model to include it, based on the work of Lööf and Heshmati (2002a). Peters (2005) modelled the knowledge production functions with selection function (for product and process innovation), and an input equation using firm size as well as market structure as input determinants in both selection and input equations. To allow non-linearity (a U-Shape between firm size and intensity of innovation), Peters (2005) added the logarithm of number of employees and the squared logarithm of number of employees into the model. To cover the second Schumpeterian theory, which states that market power stimulates innovation, the market structure is captured by a lagged index to avoid endogeneity.

In the knowledge production function Griffith et al. (2006) distinguished product innovation from process innovation in order to identify their respective effect on labour productivity, but they considered all firms and not only innovative firms like the CDM model did. Griffith et al. (2006) and Moreno and Huergo (2010) argued that all firms carry out some innovation activities but they do not report this effort if it is below a certain threshold or because they want to keep their R&D activities secret.

Based on the work of Griffith et al. (2006), Mairesse and Robin (2009) estimated the relationship in a non-linear simultaneous equations model, which consists of five equations. Product and process innovation are modelled in a bivariate Tobit model, which represents the knowledge production function. Each equation includes an endogenous regressor, and several exogenous regressors, but the equations are jointly determined and their errors are correlated.

Polder et al. (2009) modified the model of Van Leeuwen and Farooqui (2008) by using an expanded CDM model to consider ICT as a driver for non-technological innovation, alongside R&D as a driver for technological innovations. Polder et al. (2009) proposed a model in which the the knowledge production function has two input equations one for R&D and the other for ICT, in addition to three outputs for product, process and organisational innovations. Hall et al. (2012) confirmed the model proposed by Polder et al. (2009) and extended the CDM framework by including both ICT investment and R&D expenditure as inputs to the knowledge production function. Nguyen and Martin (2010) expanded the CDM model by considering both R&D and ICT as innovation input that affect the probability of introducing technological and non-technological innovations and consequently the level of productivity growth.

There is a large set of empirical studies which investigate the relationship between innovation and productivity, focusing on technological activities as a source of firms' performance. Appendix A includes an overview of the empirical studies aimed at investigating the relationship between firm productivity and innovation using the CDM model, and the relevant information such as the time period or used data, the estimation approach, proxies used for innovation input, innovation output, and whether or not ICT has been considered in the model.

2.4.3 The Impact of Innovation on Productivity

The impact of innovation on productivity on firm level is somewhat controversial in the literature. Atkinson and Wial (2008) state that labour productivity is the best aggregate indicator of the economic fruit of innovation. However, these studies were done using data from different countries, use different proxies for innovation or productivity, employ different data models (panel or cross-section), different configurations for the econometric model, or different estimation methods. This diversity might explain the different, or in some cases contradictory results of these studies. A summary of the studies is provided in appendix A.

The impact of product innovation was investigated by many works using CDM framework or other approaches. As a proxy for product innovation, the share of sales resulting from new or modified products, or sales per employee, were used. In studies with a panel data model, if further innovations were considered in the same model, a dummy variable was used to proxy occurrence (or not) of product innovation.

Most of the studies used a cross-sectional data model. A positive relationship between product innovation and productivity was found by Griffith et al. (2006), Crepon et al. (1998), and Mairesse and Robin (2009) in French data, Griffith et al. (2006) in Spanish data, Klomp and van Leeuwen (2001) in Dutch manufacturing data, Lööf and Heshmati (2002c) and Janz et al. (2004) in Swedish data, Jefferson et al. (2002) and Mairesse et al. (2012) in Chinese data, Nguyen and Martin (2010) in data from Luxembourg. Griffith et al. (2006) found a positive relationship in UK data, as did Criscuolo and Haskel (2003), though the relationship was weaker. However, Van Leeuwen and Klomp (2006) found no impact of product innovation on productivity in Dutch data. In German data, Janz et al. (2004) and Peters (2007) found a positive relationship between product innovation and productivity, however, Griffith et al. (2006) found no relationship.

Using panel data model, Baum et al. (2015) found a positive effect of product innovation on productivity in Swedish manufacturing data. Similar result was found by Raymond et al. (2013) for Dutch and French manufacturing data, and Zemplinerova and Hromadkova (2012) for Czech firm data. Furthermore, Moreno and Huergo (2010) found positive correlation between the occurrence of product innovation and firm productivity in Spain firm data, Peters et al. (2013) and and Roberts and Vuong (2013) in German manufacturing data.

However, Parisi et al. (2006) found in Italian manufacturing data no impact. Polder et al. (2009) found in Dutch manufacturing positive impact on product innovation on productivity only if combined with organisational innovation. Smilrat finding resulted by Hall et al. (2012) for Italian manufacturing data.

Mairesse and Mohnen (2010) and Mairesse and Robin (2009) noticed that the impact of process and product innovation differs from one country to another. They stated that process innovation leads to the reduction of production costs, but the fruit of product innovation takes more time to appear in the productivity statistics.

Using a panel data model, a positive effect of process innovation on productivity was found by Moreno and Huergo (2010) in Spanish data, Parisi et al. (2006) in Italian data, Peters et al. (2013) and Roberts and Vuong (2013) in German manufacturing data. In Dutch manufacturing data, Polder et al. (2009) found a positive impact on process innovation on productivity only if it is combined with organisational innovation. Similar findings was recorded by Hall et al. (2012) in Italian manufacturing data.

Using a cross-sectional data model, Janz et al. (2004) and Peters (2007) found a correlation between process innovation and productivity in German firm data. Similar findings were identified by Janz et al. (2004) in Swedish data, Huergo and Jaumandreu (2004) in Spanish data, Griffith et al. (2006) in French data, and Nguyen and Martin (2010) in data from Luxembourg.

However, Griffith et al. (2006) found no relationship between process innovation and productivity in Spain, the UK, or Germany. Mairesse and Robin (2009) also found process innovation had no impact on productivity in French data. Koellinger (2008) did not find a strong relationship between process innovation and profitability. The reason is that process innovation may take longer to generate returns than product innovation. Furthermore, process innovation might be independent of other technologies and resources, which were not advanced enough to yield returns.

Hall (2011a) investigated the relationship between innovation and productivity and found process innovation had a negative effect on productivity for two possible reasons: The first reason may be that these firms do have market power but they operate out of the elasticity of demand curve, therefore productivity falls when they become more efficient. A second reason may be the high measurement error in the innovation variables, which is significant in the productivity equation.

The impact of organisational innovation on productivity is positively polarised in the literature. Kangasniemi and Robinson (2008) argue that organisational change may increase productivity through optimising the usage of new technologies to enhance the horizontal structure of enterprises and through more efficient utilization of skilled labour. Vickery and Wurzburg (1998) argues that workforce skills and management ability, as well as the effective use of new technologies in enterprises need an appropriate organisational structure to result in productivity.

A positive impact of organisational innovation on productivity was found in panel data model by Hall et al. (2012) using Italian manufacturing data and Polder et al. (2009, 2010) using Dutch manufacturing data. A similar effect was found by Nguyen and Martin (2010) in a cross-sectional data model from Luxembourg. Kangasniemi and Robinson (2008) found organisational change had an impact on productivity in highly skilled industry in the UK, but this impact was very small.

2.4.4 The Dynamic Causality between Innovation and Productivity

According to Schumpeter (1943), a recursive link from productivity to innovation activity does exist because a firms' previous performance will positively affect future investment in innovation activities. Investigating the dynamic behaviour is related to the availability of panel data because it cannot be done by analysing cross-sectional data (Raymond et al., 2013).

The dynamics of production function was investigated in the literature. The concept of treating productivity as a function to the expected future productivity was developed by Olley and Pakes (1996). Furthermore, Peters et al. (2013) estimates a dynamic structural model using MIP data to measure the firm's profit. In this model, productivity is a function of expected future productivity, in addition to the assumed shift of the distribution of future productivity caused by innovation. Similarly, Roberts and Vuong (2013) develop a structural dynamic model to measure the expected firm's profit resulting from its investment in R&D. The R&D leads to a higher probability of generating innovations, and a higher level of productivity next year, which raises future profits. However, Roberts and Vuong (2013) state that predicting the impact of the innovation process on a firm's profit is difficult and uncertain because the output of the innovation process is not fully predictable.

Raymond et al. (2013) emphasised the importance of the dynamic relationship between innovation and productivity for firm heterogeneity for the following reasons. Firstly, there is a time lag between the decision to invest in innovation and the fruit of innovation success. Secondly, successful firms that conduct innovation tend to innovate more in the future. Thirdly, firms with a good performance show more persistence in conducting innovation activities. Fourthly, firms tends to utilise productivity revenue to fund innovation instead of asking for external funding. Raymond et al. (2013) focused on the direction of the causal relationship between investment in R&D, innovation, and productivity. They found uni-direction causality of past R&D which affects innovation, and of past innovation that affects productivity without evidence of a recursive link from productivity to future R&D investments.

Zemplinerova and Hromadkova (2012) investigated the problem of simultaneity in the equation of knowledge production function and the production function. Innovation output is expected to raise performance, but at the same time higher productivity improves innovation output. Similarly, Baum et al. (2015) investigated the dynamic behaviour in the relationship described in the CDM model and used a generalised structural equation model for estimation on Swedish data. They argue that prior behaviour of firms' productivity affects the amount of investment in R&D, but they did not consider it an issue but rather as a relationship to be investigated. To solve this, Zemplinerova and Hromadkova (2012) proposed to use the data on barriers of innovations available in the CIS data as an instrument which does not directly affect either the output of innovation or firm performance at the same time.

Lööf and Heshmati (2002a) addressed serious challenges when they modelled the relationship between innovation and productivity. Firstly, it is difficult to untangle the separate effects of the variables because most variables change together over time. Secondly, establishing causality is hard because innovations are themselves affected by output and by the past profit and productivity of the firm. Trying to solve these issues, Lööf and Heshmati (2002a) formulated simultaneous equation models and used more complex estimation approaches.

2.5 Determinants of Innovation and Productivity

In the previous section, the relationship between innovation and productivity was presented in different stages. This section addresses the main determinants that affect innovation and productivity over these stages of the relationship.

2.5.1 Firm Size

The German economy consists not only of big companies such as Siemens, Daimler, or Volkswagen, but also includes SME called 'Mittelstand', which make a significant contribution to German economic success (Hamilton and Quinlan, 2008).

The influence of firm size on innovation and productivity has been widely investigated in the economic literature in terms of three major aspects. Firstly, the impact of firm size on a firms' decision to invests in innovations and ICT. Secondly, the relationship between firm size and producing innovations. Thirdly, the effect of firm size on the relationship between innovation and productivity. The conclusions of these empirical studies differ regarding magnitude and significance, or even signs of the relationship. However, most studies agree about the important effect of firm size when studying innovation. In CIS German data, 86 per cent of all firms are small enterprises that have between 5-49 employees, and a further 13 per cent of firms have between 50-499 employees (Rammer and Peters, 2013).

Firm size and the decision to innovate: The economic literature agree that firm size positively affects the decision to invest in ICT, and in activities that may lead to innovation. Nguyen and Martin (2010) find that firm size positively affects the decision to invest in ICT. Schumpeter (1943) claims there is a positive relationship between firm size and innovation decisions for two reasons. Firstly, large firms have the advantage of ensuring financial resources to innovate especially concerning R&D projects with higher risk. Secondly, firm size is correlated with the stability and availability of internally generated funds that reduce the risk associated with prospective returns from innovation. Gault (2013) finds that SMEs are less likely to conduct R&D and innovate than large firms, but large firms have a correlation between R&D performance and innovation activities. Furthermore, Gault (2013) argues that small firms have limited resources, so by making just one mistake in their business strategy they may go out of business as a victim of creative destruction.

Morris (2008) states that Schumpeter's hypothesis of 'creative destruction' was driven by the analysis of various social and economic systems and has no empirical foundation, especially in terms of the relationship between firm size and a firms' ability to innovate. People involved in failed businesses may create new businesses which learn from their mistakes. By making mistakes, large firms learn and accumulate knowledge that they then use to dominate the market.

Firm size and producing innovation: Discussions on the effect of firm size on producing innovations can be split into two contradictory groups. On the one hand, Hall (2011a) found that the rate of presenting process or product innovation to the market is much higher for large firms than SMEs because they are involved a wide range of activities and they could innovate in at least one of them. Nevertheless, we cannot say that large firms are always more innovative than small firms based on these findings. Using German data and for all industry sectors, Aschhoff (2013) found a significant positive relationship between the firm size and the presentation of technological innovations; this effect is also available for non-technological innovation but is less sharp.

On the other hand, Cohen (2010) presented the counter opinion that firm size has a negative effect on innovation for two reasons: firstly, because large firms may suffer from a loss of managerial control or extreme bureaucratic control, and secondly, because the incentives for scientists may be blunted as they are not able to benefit from their creative motivation or their individual results, or they are frustrated because of the conservative hierarchies of large firms, which Schumpeter himself argued in his early work. Zemplinerova and Hromadkova (2012) found in Czech CIS data a negative relationship between a firms' size and innovation output as innovation sales per employee, which means that smaller firms are more efficient in transforming innovation inputs to outputs. Acs and Audretsch (1988) found the relationship between firm size and innovation to be negative using U.S. small business administration data on new products processes and services. Cohen (2010) found a negative relationship between the productivity of R&D and firm size. Small firms, especially new firms, are more capable of innovating than large firms. Cohen and Klepper (1996) argued that R&D returns increase with the sales output because the fixed costs of innovation will spread and reconcile the positive relationship between R&D expenditure and firm size with both declining R&D and productivity in large firms, and the higher probability of incremental innovation and process innovation in large firms.

Considering the innovation patterns and whether firm size affects the production of incremental versus radical innovations, the key findings are that larger firms tends to pursue more incremental innovation (Henderson, 1993) and more process innovation (Cohen and

Klepper, 1996). However, smaller firms generate radical innovation (Cohen, 2010).

Firm size as a determinant has an issue of 'reverse causation', which is the impact of innovation output on the change in the firm size. To avoid that effect, Zemplinerova and Hromadkova (2012) fixed the employment level at the starting point of data.

Rammer and Schubert (2016) analysed German firms' behaviour in the last two decades and noticed that while innovation expenditure grew, the number of firms conducting innovation activities fell, hence innovation expenditure is concentrated on fewer firms. Furthermore, most firms that tend to refrain from innovation activities are small firms and this behaviour occurred during and after the financial crisis of 2008 but was unrelated to it.

Firm size and the relationship between innovation and productivity: The economic literature agrees about the positive impact of firm size on the relationship between innovation output and productivity. Schumpeter (1943) argues that the returns from R&D are higher if the volume of sales is larger, because the fixed costs of innovation will spread. Furthermore, R&D activities are more productive in large firms because of the possible complementarity between R&D and other business activities like financial planning and marketing. Peters (2005) found a high positive impact of firm size on the output of process innovation driven by higher cost reduction. Aschhoff (2013) argues that firm size plays an important role since large firms are more able to reach the next geographical market than a small or medium size firms.

Hall (2011a) finds a weak correlation between productivity and firm size, which could be due to the regulatory and financial environment dominant in the country. Bartelsman et al. (2009) found that the relationship between firm size and productivity differs across countries and industries. Therefore, Bartelsman et al. (2009) developed a calibratable model to check the inefficient allocation of resources in firms. The model is based on the covariance between firm size and productivity to prevent firms exceeding their optimal size and therefore being less productive.

General observations: Cohen (2010) addressed three important limitations in studies done on the relationship between firm size and innovation. Firstly, most of the samples used are non-random and have sample selection bias. Secondly, studies did not examine whether the relationship is due to firm characteristics in correlation to firm size, such as cash flow or degree of complementary and diversification, which give firms the ability to spread R&D

costs over output. Thirdly, due to the aggregation of businesses engaged in a variety of industries outside their primary industry, especially in large firms, it is not easy to control for the industry effect.

2.5.2 Market Structure

Different theories to describe the relationship between market structure and the firms' ability to innovate were found in the economic literature, but all of them agree that an intermediate market power may maximise the incentive of firms to innovate.

The first theory of Schumpeter (1943) emphasises the role of market power on innovation activities for two reasons:

- Firms require the expectation of market power to invest in R&D, supported by patent law and based on market power caused by post-innovation, which in turn ensures financial resources for firms to invest in R&D.
- The possession of market power linked to a monopolistic market structure encourage firms to innovate.

Grossman and Helpman (1991) modelled the interaction between innovation and imitation and argues that imitations shorten the time period that an innovator can enjoy monopoly rents and reduce the motivation to innovate. Grossman and Helpman (1991) distinguish between two groups of firms: leaders (Northern), who undertake innovation and R&D to develop the first generation of the product, and followers (Southern), who do not pay research costs and imitate. Successful innovators get monopoly profits for a while because they introduce new products on the markets before other suppliers. However, successful imitators get rents because their manufacturing costs are lower than those of competitors. Grossman and Helpman (1991) investigated three types of market structure in which firms can earn rents: Northern firms which innovate to compete on quality with other rival Northern firms; Northern firms which innovate to compete on quality over a rival Southern firm; and Southern firms which imitate the product of a Northern firm and compete with them on price.

The second theory argues that firms operating in a competitive environment have greater incentive to invest in R&D and innovation than a monopolist firm that wants to maximise the profit from its old technologies (Arrow, 1962). Crepon et al. (1998) stated that market share and diversification have a positive impact on a firms' R&D effort; demand pull and technology push also have a positive impact on a firms' R&D activities. Furthermore, market competition allows better resource allocation and more efficient companies because non-productive firms get out of the market and new production units will be created. The share of turnover due to market novelties can have an impact on innovation firms and indicates their potential for 'creative destruction' (Aschhoff, 2013). Scherer and Ross (1990) argue that the absence of competitive pressure increases bureaucracy and discourages innovation.

Nguyen and Martin (2010) found that operating in a highly competitive market increases the probability that a firm will introduce organisational innovations, however, this do not affect the probability that a firm will introduce product or process innovation. Crespi et al. (2006) found that competition stimulates organisational change.

Mairesse and Mohnen (2004) found that demand pull and market share plays a significant role only in low technology sectors but not in high technology sectors. Mairesse and Mohnen (2010) found that demand pull is more important for innovation than technology push.

A third theory presents an intermediate opinion. Scherer (1967) investigated the regression of R&D intensity against market concentration and found an 'inverted U-shape' relationship. The relationship between market competition and innovation was examined by Aghion et al. (2005) using panel data and a similar relationship was identified. Aghion et al. (2005) argued that firms trade-off two effects: firms with low competition try to mitigate competition, while firms with high competition may increase their profits from innovation because they lower rents before innovation compared with those post innovation. Furthermore, they formalised the relationship between intensity of competition and the count of innovation. Schubert (2010) used German CIS data to investigate whether innovations are more likely to occur in monopolistic or competitive markets and the degree of competition, by which innovations are maximised. He found that firms with extremely weak or leading position on the market tend to conduct only market or organisational innovation, whereas firms with a moderate market share are much more likely to have an extensive innovation strategy that covers all four types of innovation.

Mairesse and Mohnen (2010) found that data from innovation surveys contributes to explaining innovation because it allows renewal of the Schumpeter hypotheses on firm size and monopoly power, and discussion of demand pull and technology push. In absence of direct measures for demand pull and technology push, several indicators have been used to express the market structure. Mairesse and Mohnen (2010) proposed using the objective of increasing or maintaining market share to indicate demand pull, and using the importance of government and university research labs as source of information which leads to innovations as a proxy for technology push. Monjon and Waelbroeck (2003) considered the market share as indicating market power and diversification. Belderbos et al. (2004) used the sum of scores on the significance of the innovation as a 'demand-pull' factor and the sum of scores on its significance for cost reduction as 'cost-push' factor.

Cortright (2001) states that the Neoclassical economy assumes that the markets are very competitive rather than tending toward monopolises. However, the knowledge-based economy tends towards to 'monopolistic competition', in which businesses try to maintain their monopoly position and compete on product characteristics such as new products, new features, variety, and quality rather than competing based on cutting prices.

Javorik (2004) found that if a higher productivity firm enters the market, other firms from the same industry can be encouraged to improve their performance by adopting new technologies or by hiring skilled workers from foreign owned firms, especially in the high technology sector. Consequently, domestic firms may be forced to exit the market.

2.5.3 Geographic Area of Operation

The area of operation of an enterprise has an impact on its innovation behaviour. OECD (2010b) results show that companies operating in international markets have between 40% and 70% more likelihood of introducing innovations than other companies. Furthermore, investing in foreign countries has a positive impact on innovation because it transfer technological knowledge to the firm by the spillover effect. However, this effect is not unanimous (Javorik, 2004). According to Aschhoff (2013), firms may expand their innovation activities to foreign countries due to the following possible motives:

- Market: being closer to local customers and reacting faster to their requests and needs.
- Cost reduction: reducing personnel cost, the low cost of setting up laboratories or to staying geographically close to the reduced-cost production.
- Technology: using resources in foreign locations, such as knowledge or skilled personnel, which are not available in the same quantity or quality in the firms' country.

MIP data covers firms that have their headquarters in Germany and German subsidiaries of firms which have their headquarters abroad. This allows the degree of internationalisation as a relationship to the industry sector and firm size to be captured (Aschhoff, 2013).

The use of ICT affects the size of the geographical area in which business operates. Polder et al. (2009) argue that competing in a foreign market pushes firms to be innovative and communicative, which enhances investment in both R&D and ICT. Koellinger (2008) argues that using ICT facilitates cooperation and communication across wider geographic areas.

Investigating the relationship between the decision to enter the export market, innovation, and firm performance shows that two types of relationships have been addressed in the literature. In both of them, the problem of simultaneity exists.

Firstly, the literature examines the link between export and productivity. Firms engaged in export are most productive and firms that focus only on domestic market are the least productive (Kimura and Kiyota, 2006). For German and Austrian data, Hansen (2010) found that exporter firms are around 40 per cent more productive than non exporter firms. However, the relationship between export and productivity seem simultaneous because firms with higher productivity have a higher probability of exporting and being active on foreign markets than firms with lower productivity (Helpman et al., 2004; Hansen, 2010; Jienwatcharamongkhol and Tavassoli, 2014).

Secondly, the link between export and innovation is discussed. Cassiman and Martínez-Ros (2007) investigated data from exporting and non-exporting firms and found that only product innovation drives the export decision of small non-exporting firms. Van Beveren and Vandenbussche (2009) found a 'self-selection' effect that means only firms which are highly likely to export tend to conduct product and process innovation before entering the export market. However, the relationship has causality because better performing firms tends to enter the export market and exporting firms tends to have a good performance.

Wagner (2012) analysed the relationship between export, conducting R&D and productivity for German firm data. He found that the productivity of exporter firms that conduct R&D is higher than that for those do not conduct R&D, which in turn are more productive than firms who neither export nor conduct R&D. Furthermore, export results in competition and spillover effects that support a firms' innovation effort and correlate with the firms' level of technological skills (Crespi and Zuniga, 2012). Jienwatcharamongkhol and Tavassoli (2014) found in Swedish data that the export firms, driven by their past R&D activities, are productive and generate innovations.

As mentioned in the discussion about innovation patterns, some firms in coordinated market economies may shift some of their activities to liberal market economies to increase radical innovation. Contrarily, some firms in liberal market economies may shift some of their activities to coordinated market economies to improve quality control and skill level. This may help to explain why some German pharmaceutical firms have research labs in the United States and why General Motors has located its engine plant in Germany (Soskice and Hall, 2001).

Additionally, the local area of operation plays an important role. German CIS data between 2002 and 2010 shows that the share in turnover due to product innovation and the effectiveness of process innovations in Western Germany have higher rates of success than in Eastern Germany despite the large expenditures of East German firms and supporting policies (Aschhoff, 2013).

2.5.4 Firm Age and Employee Age

The empirical literature disagrees about the effects of firm age on innovation. A general observation is that the negative impact has a link to the technological innovation, but that the positive impact is related to non-technological innovations. The assumption was that firms learn during their life course and as they get older they gain experience, which may make work more efficient and facilitate innovations. According to the framework of Schumpeter (1942), firm age plays a deciding role in innovation, in terms of novelty and imitation.

On the one hand, Coad et al. (2012) investigated how firm age impacts on innovation and firm growth using Spanish CIS data for the period 2004-2010; they found that firm age shows a significant negative impact among young firms and a non-significant impact among old firms. Balasubramanian and Lee (2007) analysed firms' patent data to examine how firm age relates to innovation quality for different industry areas; they found that firm age is related negatively to innovation quality and that this impact increases for fields which can be considered more technologically active. Huergo and Jaumandreu (2004) analysed the effect of firm age and process innovation on the growth of productivity and found that young firms have above-average productivity growth, which is progressively weakened, for many years.

On the other hand, there is the positive impact of firm age linked to the organisational innovation. In the context of organisational learning, Levitt and March (1988) state that new firms face difficulties due to the lack of market recognition, alliances with business partners and managerial knowledge.

When discussing the average age of employees, Germany's demographic development plays a key role. The participation rate of the older working population in Germany is increasing with time (Meyer, 2008) due to the demographic change in the society (Bertschek and Meyer, 2008). Schubert and Andersson (2013) investigated the relationship between average employee age and firm innovativeness using CIS Swedish data and found that employee age negatively affects both propensity and success of product innovation. Similarly, for German data from knowledge-intensive service firms Meyer (2008) found that the probability of technology adoption is related negatively to the age of employees. The result may be due to the fact that younger employees are more acquainted with ICT. However, older employees are experienced and more familiar with the firms' structure and processes. Schubert and Andersson (2013) found that the optimal level of turnover is lower for firms with older employees.

However, Meyer (2008) found that firms that have advanced teamwork and more older employees than younger employees have a greater likelihood of absorbing new technologies than other firms. Ciriaci et al. (2012) goes beyond this result and using Spanish firm data, found that younger innovative firms are likely to grow faster in terms of sales and employment, but have difficulty to innovating at later business stages and do not show an advantage in their innovative sales growth.

Schneider (2007) analysed the impact of the age structure of employees on product innovation in German manufacturing data and found an inverse U-shaped relationship. Using German manufacturing and service data, Bertschek and Meyer (2008) investigated whether the age of the workers affects firms' labour productivity driven by IT. They found that the productivity of workers older than 49 is not less than workers of the prime age between 30 and 49, but workers who are younger than 30 have less productivity than those of the prime age. Furthermore, the study found that IT has a positive impact on the productivity of older workers.

Another determinant which may affect innovation is the rate of employee staying in the firm. Schubert and Andersson (2013) found that the relationship between employee retention rate in a firm and innovation shows an inverted U-shape.

2.5.5 Protection Measures

Firms employ different approaches to protect their intellectual property, including informal strategic approaches such as secrecy and the advantage of lead time, but also formal approaches such as employing patents and trademarks. Mostly, the usage of protection mechanisms is related to firm size. A manufacturing sector firm with more than 1000 employees uses a combination of formal and informal mechanisms. However, small firms relay more on informal mechanisms, which may due to the higher costs of using formal protection (Aschhoff, 2013). The protection approach used is a determinant that affects both the generation of innovation outputs and the productivity function.

Romer (1990) found that the return from R&D depends on some external conditions and institutional features such as the length of patent protection and the coverage of trade mark protection, the efficacy of protection, and the regulatory framework of the business operation. Activities producing inadvertently disembodied knowledge cannot be kept as a trade secret and will be available to other inventors, hence will lead to a cost reduction in their R&D activities.

Peeters and Pottelsberghe (2004) argues that protecting intellectual property is an incentive for firms to motivate them to invest more in innovation activities. Mairesse and Robin (2009) also emphasised the significant role of intellectual property protection in product innovation but not in process innovation.

Aschhoff (2013) found in German R&D intensive manufacturing between 2008 and 2010 that the most efficient strategies were informal protection measures more than the formal measures. Nguyen and Martin (2010) found that using formal methods to protect innovation positively affects the probability of presenting all kinds of innovations. Coe et al. (2008) examined the impact of strong patent protection on a macroeconomic level and found it accompanied by higher levels of TFP, higher returns from domestic R&D activities and higher foreign R&D spillovers. However, Schmidt (2005) found a negative impact on using formal methods of protection like patents on R&D cooperation especially in Germany. Peeters and Pottelsberghe (2004) found no effect of protecting firms' intellectual property on labour productivity, but it is important for a firms' innovation process to cover the costs of innovation activities.

2.5.6 Public Subsidies

The government plays a major role in fostering innovation through regulation of domestic activity and trade (OECD, 2010b). OECD (2010b) results show that firms which receive public financial support are greater investors in innovation activities than those which do not. Even though subsidies are prohibited by an European Union (EU) treaty because they harm free trade and competition rules, competition authorities are still providing exemptions and EU and national governments continue to provide R&D subsidies in order to boost innovation and thus stimulate economic growth.

National and EU subsidies impact negatively on the relationship between innovation inputs and outputs. Firms that get national and EU subsidies invest more in innovation but their innovation outputs are below those of firms that are not subsidised by national or EU bodies (Zemplinerova and Hromadkova, 2012; Almus and Czarnitzki, 2002; Gonzalez et al., 2005; Hashi and Stojcic, 2010).

Zemplinerova and Hromadkova (2012) notice that large companies have a better chance of receiving subsidies because they have more political power. Kemp et al. (2002) found that national innovation policies can encourage smaller firms to innovate more and affects the amount of innovation expenditure. Moreover, they noticed that small firms claim mostly national subsidies, while medium-sized firms obtain European subsidies.

Based on several studies evaluating the impact of government support on firm behaviour, Mairesse and Mohnen (2010) conclude that government support for R&D activities leads to more R&D expenditures in the firm without crowding-out public R&D. Rammer and Peters (2013) argues that public subsidies stimulate innovation activities and have a sustained effect by inducing a permanent change in firm innovation behaviour.

The elements of the national innovation system, such as R&D policy, financial system, tax policy, educational system, competitiveness, and international integration are significant factors in productivity performance. Lööf and Heshmati (2002b) found that the source of high productivity in Finland is on macro level rather than the micro level, but in Sweden it appears to be on both. Macro variables are not focused on directly in this study because the investigated firms do business in the same macro conditions, which is the German macro environment. Further research investigating several national innovation systems may outline the strengths and weaknesses of each of them in supporting innovation. Fernandes and Paunov (2012) conclude that credit constraints force firms to take a higher risk of innovation and therefore they see policy makers in the role of offering support to deal with failed innovations in order to reduce firms' exposure to innovation risk.

Public subsidies seem to play a minimal role in Germany. Only about 8 per cent of German innovative firms use public subsidies to finance their innovation activities; this may be due to the higher costs of applying for public subsidies for small firms. However, 80 percent of small firms and 96 percent of large firms finance innovation activities using their own cash flow and bank loans to finance big projects (Aschhoff, 2013).

Rammer and Schubert (2016) state that public support for innovation can improve the level of innovativeness in firms. Peters (2006) and Aschhoff (2013) found that among firms receiving public funding, small firms have a higher share of innovation expenditure financed through public sources. However, large firms have a very small share of innovation expenditure.

MIP data includes information regarding whether a firm has received public funding from any government agencies to finance innovation activities.

2.5.7 Knowledge Spillover, Cooperation and Collaboration

Innovation cooperation is a 'formalised and target-oriented exchange of knowledge, in which partners do not benefit commercially from collaboration' (Aschhoff, 2013). Cooperation partners may be universities, public research institutions, other enterprises, supplier, customer or competitors. As declared by firms, universities are the most used and the most valuable cooperation partner but are less considered as an important information source for innovations(Aschhoff, 2013).

Cooperation in R&D activities among development partners and enterprises is motivated by the need to share knowledge, exchange complementaries, and share risk and cost (Cassiman and Veugelers, 2001). This need is due to the lack of knowledge and resource constrains, due to absorptive capacity, or because of the financial resources needed for innovation projects, which are mostly expensive and long-term in nature (Schmidt, 2005). Furthermore, firms internal knowledge is often insufficient to generate innovations, which leads them to expand their internal knowledge through cooperation and collaboration with external partners (Aschhoff, 2013). Guellec and van Pottelsberghe (2004) insist on the role of public institutions as an important source of knowledge that affects productivity growth.

Franco et al. (2012) found in firm data from Germany, Italy and Spain that the absorptive capacity increases if the firm interact with research organisations such as universities and laboratories. The knowledge production process incorporates a large number of individuals and organisations, and requires coordination and communication (OECD, 2010a). Such complex networks or 'collective intelligence' can be followed as a part of an 'innovation measurement framework' including the exchange of technology between universities and industry (OECD, 2010a).

Mairesse and Mohnen (2010) found that cooperating firms have higher R&D expenditure. Similar results have been found by OECD (2010b) showing that collaborating firms invest in innovation 20% to 50% more than firms not involved in collaborations.

Mairesse and Mohnen (2010) found that cooperation in R&D depends on the technological system. A system with low appropriability (the ability of the firm to protect its internal knowledge) discourages collaboration with suppliers; demand pull encourages collaboration with customers and supply dominated firms and discourages collaboration; and science-based systems innovate and collaborate with universities. Schmidt (2005) used the importance of strategic protection methods for internal knowledge as a proxy for appropriability.

Belderbos et al. (2004) investigated the impact of several types of R&D coordination partnerships (supplier, customer, competitor, and university or research institute) on two indicators of firm performance (labour productivity and the growth in sales per employee resulting from products newly launched on the market) using CIS Dutch data. They found that all cooperation types contribute positively to firm performance. Furthermore, they found that cooperation with competitors and suppliers is important for incremental innovations, and cooperation with universities and competitors increases product novelty on the market and is thus a significant source of knowledge which drives radical innovation.

Mairesse and Mohnen (2010) found that cooperation is more intensive among small firms than large firms. (Mairesse and Mohnen, 2004) found that cooperation in high-tech sectors is not a significant factor in defining the intensity of R&D, in contrast to the low-tech sector. Schmidt (2005) drew corresponding findings that the probability that a firm cooperates rises with its size but also depends on the industrial sector.

Monjon and Waelbroeck (2003) defined technological spillovers as 'the non appropriate amount of knowledge that is created by an innovative firm'. They tried to quantify spillover using some proxies from a CIS survey for knowledge transfers and research collaboration such as:

- Information sources that firms use to innovate as a direct measure for knowledge flow between universities and innovative firms instead of proxies like R&D expenditures or patents.
- The source of national and international collaboration may decompose the knowledge flow into formal collaboration and spillovers.
- The measure of innovation novelty, where the degree of innovativeness depends on whether the product is new to the market or new to the firm.

Harhoff (1990) found in German manufacturing data a stronger positive productivity-enhancing effect for spillovers in high-technology sectors than in other industry sectors, and that spillover encourages R&D investments.

Mairesse and Mohnen (2010) summarised the work of Monjon and Waelbroeck (2003) and Belderbos et al. (2004) and established a relationship between innovation patterns and collaboration: incremental innovators in close proximity to universities benefit from spillovers of intra-industry knowledge through suppliers and customer collaboration. However, radical innovators collaborate only with universities. Lööf and Heshmati (2002b) found that cooperation with domestic universities has positively impacts on innovation outputs. Monjon and Waelbroeck (2003) investigated the contribution of information flow from universities to innovative firms through international formal cooperation and knowledge spillovers using French CIS data. Firms that imitate existing technologies or that implement incremental innovation benefit mostly from spillovers. However, highly innovative firms obtaining benefits from collaborative research but only with foreign universities.

Monjon and Waelbroeck (2003) argue that the knowledge generated in universities diffuses among firms through formal cooperation, or informally through knowledge spillovers in conferences and publications. Universities conduct basic research, which is the generator of radical innovations and firms gain the most benefits from formal collaboration with universities. A weakness in the data is that universities and firms are covered by indicators, yet currently it does not consider the role of individuals, consumer or government in the innovation process. The OECD (2010b) addressed the need for definitions for public sector innovation, as well as for metrics to measure innovation efforts.

Peeters and Pottelsberghe (2004) stated that the informal interaction between a firm and its customers, suppliers, consultants, and competitors results in a significant contribution to labour productivity. However, a formal R&D partnership with research institutes and universities to create knowledge leads to a positive impact on labour productivity.

Robin and Schubert (2010) investigated the impact of a firm's cooperation with public research institutes on product and process innovation using German and French CIS data. They found a positive impact on product innovation but no impact on process innovation.

Peters (2005) proposed measuring spillovers in terms of the importance of conferences, journals, exhibitions as the given source of innovation. Schmidt (2005) measured spillover by the importance of external knowledge available to each firm. Mohnen et al. (2006) argues that firms show a more innovative performance when they diffuse non-confidential technologies via publication.

Using ICT facilitates collaboration and communication among business partners (Koellinger, 2008) and positively affects spillover by expanding the development network of academics, international researchers and the collaborative groups (Gretton et al., 2003).

German data shows that 42 percent of product innovations were developed in collaboration with other partners and 46 percent of process innovations were developed with third parties (Aschhoff, 2013). In MIP data, cooperation and collaboration have been captured since 2007 using three indicators: information source, involvement in innovation cooperation, and the final result of collaboration with external partners. Aschhoff (2013) summarised the MIP captured sources of information needed for innovation in four groups. The first two sources of information are considered as most important and most widely used by firms:

- Internal sources, within firms or the firm group.
- Market sources such as customers, suppliers, competitors, consulting and knowledge providers.
- Institutional sources, such as universities and public research organisations.

- Other sources, such as conferences, fairs, scientific publications and patent specifications.

2.5.8 Human Capital

To maintain a competitive advantage in the knowledge-based economy and technical evolution, firms rely on people with higher levels of skills and individual competence, who are becoming valuable assets known as human capital (Kwon, 2009). The relationship between human capital, innovation, and productivity shows two streams in the literature. The first is the 'labour force', in the classical economic perspective, similar to other production inputs such as physical capital as one of the production elements that generates added-value. This line of thought links human capital directly to the Cobb-Douglas production function. The second is the 'human as creator' who frames knowledge, skills and experiences to connect 'self' and 'environment' (Kwon, 2009). This view links human capital to the innovation process to generate economically valuable knowledge.

The first stream emphasising the direct contribution of the accumulation of human capital on labour productivity growth is dominant in the literature (Donselaar et al., 2004). Crepon et al. (1998) found that the rate of skilled worker in a firm has a very significant impact on productivity. Van Ark et al. (2009) insisted on human capital as a key source of economic performance. Abowd et al. (2002) found a strong positive relationship between skilled workers and productivity. Fallahi et al. (2010) investigated the impact of the education level of the work force on labour productivity and found a significant positive effect. Coe et al. (2008) found that human capital is a significant determinant of TFP. Chiswick (2005) argues that more high-skilled professionals improve the productivity of low-skilled workers who assist them and increase the productivity of capital as well. Helpman (2004) found that workers are more productive in an environment with more educated co-workers because of the interaction learning effect. However, high productivity per worker produced by expulsion of less skilled workers from the labour market will subsequently cause productivity problems. Therefore, institutions should retain a framework in which the options of both the quantitative and qualitative potential of human capital are available (Van Winden and Reitsma, 2004).

The second stream emphasises the impact of human resources on innovation. Donselaar et al. (2004) found that the effect of human capital on the ability of a firm to innovate is not clear in the empirical studies. Nevertheless, human capital is a main factor influencing the innovation process. Janz et al. (2004) and Crespi and Zuniga (2012) pointed out that human capital may have endogeneity issues with innovation expenditure because it includes personnel working in R&D. Furthermore, Crespi and Zuniga (2012) found that the skill level of the work force is correlated with costs of labour in innovation activities.

Furthermore, Lucas (1988) presented three models endogenous to the neoclassical theory of economic growth. The first model emphasises technological change and the accumulation of physical capital. The second model emphasises the accumulation of human capital through training and schooling. The third model emphasises the accumulation of specialised human capital through learning-by-doing. Lucas (1988) found that the accumulation of human capital raises both labour productivity and physical capital productivity. In this model, growth is driven by investing in human capital: splitting time between work and training activities is necessarily a compromise as by sacrificing some of their income to train but they increase their future productivity at the same time.

Many researchers have modelled the accumulation of human capital depending on years of schooling. Because the lifetime of individuals is finite, an increase in human capital does not represent a permanent source of economic growth (Helpman, 2004). Arrow (1962) underlines the role of the learning process in acquiring knowledge and skills through study and experience. Rammer and Peters (2013) argued that skilled employees are an additional factor to the accumulation effect of knowledge caused by past innovation experience. OECD (2010b) stated that knowledge interaction in a workplace environment can lead to innovation if effective management ensures that the talents of individuals are tapped.

To measure human capital stock, Kwon (2009) proposed a three prong approach:

- Based on output, such as school enrolment rates, or average years of schooling.
- Based on cost, such as calculation of costs paid for obtaining knowledge.
- Based on income, linked to each individual's benefit obtained by investments in education and training.

OECD (2009) states that information about human capital in innovation surveys is still limited and the impact of human capital on innovation is not sufficiently discussed in the literature.

Lööf et al. (2001) found that the availability of qualified personnel affects innovation in different ways in different countries. The OECD (2009) found that a shortage of qualified personnel hampers innovation. Using German CIS data, Schubert (2010) found that the availability of human capital increases the probability of technological innovations and supports implementing innovation strategies if firms have an intermediate market. Bennett and McGuinness (2009) examined high-tech firm data from Northern Ireland and found that highly productive firms are more sensitive to skill shortages, especially under labour market constraints, and productivity may reduced between 65-75 percent and waste the productivity advantage of firms that have the best performance. A similar finding was raised by Forth and Mason (2004) using data from the UK. Hagen et al. (2008) found that education and staff qualifications are a very important factor for uptake of ICT.

Using data from 21 industrial and 82 developing countries between 1970 and 1995, Kwark and Shyn (2006) found that human capital has a significant influence on R&D activities targeted at absorbing foreign technologies. Franco et al. (2012) found in firm data from Germany, Italy and Spain that absorptive capacity into actual innovation is favoured by human capital. Junge et al. (2012) found that product and marketing innovation contribute to productivity growth significantly more quickly in skill-intensive firms compared with other firms.

2.5.9 Physical Capital and Financing Innovation Activities

In addition to its role as an input in the production function, physical capital and investment sources, together with human capital and knowledge capital significantly effect innovation. The relationship between firms that innovate and the financial market is that the capital market gives firms opportunities to invest in projects for future realisation, which are expected to create returns that exceed the return required by financial markets and the investment risk. This will promote new technology and increase competitive pressure on other firms to innovate. Rammer and Schubert (2016) states that financial capability can improve the level of innovativeness.

The ability to finance innovative projects influences labour productivity. Peeters and Pottelsberghe (2004) addressed two factors in this context: Firstly, firms able to finance innovation projects with adequate funds tend to innovate more, and therefore will impact their performance positively. Secondly, using external funds to finance innovations means, relatively speaking, that more internal resources are kept for activities related to value creation by employees.

The share of capital expenditure in total innovation expenditure represents machinery, equipment, software and other intangibles which are obtained because they are needed for innovation activities. Part of this expenditure leads to innovation output, so that financing of R&D is problematic because of low returns is to be expected, coupled with associated risks (Zemplinerova and Hromadkova, 2012).

Lööf and Heshmati (2002a) found that the lack of innovation investment sources has a negative impact on productivity. Lööf et al. (2001) found that the availability of financial sources is an important factor for innovation but it affect innovation in different ways in different countries. Johansson and Lööf (2009) state that a firm's past productivity affects R&D investments and that its previous economic performance is important to finance investments in R&D. To avoid simultaneity, Johansson and Lööf (2009) proposed to differentiate by measuring labour force between regular labour and R&D labour that is associated with R&D expenditure.

Aschhoff (2013) argues that investing in innovation activities has a worse credit crunch compared with investment in buying new production facilities because innovation activities are dominated by human capital wages to generate a specific know-how, such as non-physical goods or intangible assets. Therefore, in the case of bankruptcy it is difficult to liquidate them. In the 2007 MIP survey, firms were asked whether they suffer from credit constraints by financing innovation projects. The survey shows that firms in the R&D intensive sector, knowledge intensive sectors, and electrical and electronics industry, invest in further innovation projects in the case of sufficient funding (Aschhoff, 2013). Bond et al. (2003) found that cash flow for German firms is pretty expensive and not informative of the decision to invest in R&D, and investments in Germany are sensitive to cash flow for firms performing R&D. However, in the UK the financial constraints are more significant regarding participation in R&D activities and less conductive to long term investment than in Germany.

Mairesse and Robin (2009) emphasised the role of physical capital in process innovation, which represents the acquisition of new equipments or production lines that the process innovation may need. Peters (2005) considered physical capital to be the expenditure on physical investment per employee. Polder et al. (2010) found that capital intensity represented as depreciation per full-time employee positively impacts on productivity in both the manufacturing and service sectors to an important extent. Supan (2008) mentioned some examples of how capital affect productivity. In terms of the automotive industry, the reduction of downtime increases output for both workers and capital, as does the manufacturability design, in order to reduce the number of steps in the process of car assembly, so that less capital services and less worker hours are needed to produce one car.

According to Raes et al. (2004), capital deepening leads to higher productivity growth, especially if the labour market is less flexible. Using US data, Rao et al. (2001) found that 10 percent increase in investment in machinery and equipments contributed a 4.4 percent increase in labour productivity, but a 10 percent increase in human capital leads only to a 0.3 per cent increase in labour productivity.

2.5.10 Firm Ownership and Membership

Being a member of a group has a positive effect on innovation. Firms that are members of a multinational group with foreign affiliates have the advantage of accessing global knowledge that firms with only domestic affiliates do not (Johansson and Lööf, 2009). Furthermore, these firms have facilities for internal access to finance and knowledge resources, and group membership allows synergies that enhance the carrying out of R&D and investment in ICT (Polder et al., 2009). A similar positive effect of being a member of national or international group was identified by Knell and Nas (2006), as increasing the free exchange of knowledge and information. In this case, it is difficult to identify clear inputs and outputs because the innovations produced can be utilised globally (Knell and Nas, 2006).

Crespi et al. (2006) found that organisational innovation depends strongly on the firm's ownership; for example, US owned firms tend to introduce organisational changes more than domestic UK firms. Crespi and Zuniga (2012) found that the impact of foreign ownership is not clear. Using Chinese data Mairesse et al. (2012) found that foreign firms have less innovation input and output, though they have a higher level of productivity.

2.5.11 Firm Behaviour in Financial Crisis

Schumpeter (1942) argues that recessions eliminate firms that are unable to reorganise and innovate. To find out what the impact of business cycles is on the dynamic of innovation behaviour, firms were asked in an MIP data survey whether they remain engaged and invest in innovation projects during an economic crisis. Filippetti et al. (2009) investigated the impact of economic downturn on a firm's decision to increase or decrease investment in innovation in response to recession. The study was not able to find uniform behaviour, but creative destruction was the dominant pattern. This means that small start-up firms led investment in innovation, but big established firms reduced their investments.

2.6 Conclusion

2.6.1 Summary

The huge body of literature available on innovation and productivity and the distribution of the empirical contributions worldwide confirms the importance of the topic for micro and macro-economics, but at the same time creates a challenge due to the diversity of the investigated factors and the conflicting statements issued by different researchers.

Most studies investigate the uni-directional relationship between innovation and productivity using the CDM approach, but few studies investigated the simultaneity of the relationship in the reverse direction, which means considering whether previous productivity affects the propensity of a firm to innovate, and the amount of investment that a firm dedicates to innovation activities.

2.6.2 The Conceptual Framework

The theoretical analysis of the research questions was mainly based on reviewing the available literature on the topic, which led to the structuring of the conceptual framework as shown in figure 2.4. It aims to represent the relationship between innovation and productivity and assign the explored determinants that impact this relationship. The conceptual framework serves as a basis for development of the empirical model presented in the next chapter, which will be used for data analysis to answer the research questions.

Modelling the relationship between innovation and productivity reflects the theoretical understanding of investigated economical phenomenon, which will be tested through this research using the data. The model comprises the most important determinants of innovation activities as discussed in detail in this chapter and based on the acknowledged CDM model approach but also takes relevant improvements and expansions into consideration.

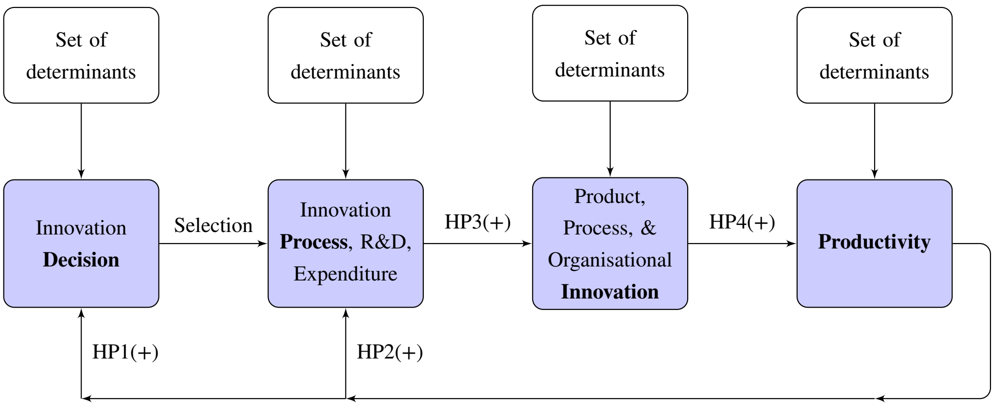


Fig. 2.4 Conceptual framework

The conceptual framework illustrated in figure 2.4 was developed based on the literature review. It draws on the CDM model proposed by Crepon et al. (1998). The first box *Decision* represents the firm's decision whether or not to take part in innovation activities. The second box *Process* represents the amount of expenditure that a firm spends on innovation in the case of positive decision, which is the input of the innovation process. The third box *Innovation* describes the outcome resulting from the innovation process. The last box *Productivity* represents the expected firm's performance resulting from innovation output. As seen in the third box *Innovation*, the model considers the extension of Peters (2005) and Parisi et al. (2006) which added process innovations and the extension of Polder et al. (2009) which added organisational innovations. The knowledge production function, in which input is the innovation expenditure and the outputs are product, process, and organisational innovations is described in HP3. The production function, in which inputs are the resulting product, process and organisational innovations and its output is labour productivity, is described in HP4.

The recursive relationship between productivity of prior time and innovation decision and the input of the knowledge production function is adopted from the work of Raymond et al. (2013) and Baum et al. (2015). Firstly, productivity affects the firm's decision to innovate or not, which results in HP1. Secondly, productivity affects the amount that firms which decided to innovate spend on innovation, and results in HP2.

2.6.3 Research Hypotheses

Based on the conceptual framework constructed in figure 2.4, the following four hypotheses can be established, which may contribute to the acquisition of greater understanding of the character of the relationship between innovation and productivity and answer the research questions stated in section 1.5:

- HP1: Labour productivity positively affects the firm's decision to engage in innovation.
- HP2: Labour productivity positively affects the firm's level of innovation expenditure.
- HP3: The level of innovation expenditure positively affects the generation of different types of innovations.
- HP4: Innovation positively affects a firm's labour productivity.

These hypotheses will be tested using German manufacturing data and according to the research strategy described in the chapters that follow.

Chapter 3

Research Methodology

3.1 Introduction

Chapter 1 outlined the research aim, questions and objectives. Chapter 2 analysed and evaluated the most relevant empirical studies that deal with the relationship between innovation and productivity. Most of these studies showed a link between innovation and productivity, however they differ as to the nature and the direction of the link. Moreover, a set of key determinants may influence how this relationship has been addressed and investigated. The results of various studies on how these determinants affect the relationship were different. Section 2.6 of chapter 2 presents the conceptual framework as a model for the preliminary relations among investigated variables, in addition to the research hypotheses, which will be tested later on.

The objectives of this chapter are to present the research methodology followed in this study and to explain the methods which have been used to achieve the research aim. Saunders et al. (2009) underline using the terms *methods* while *methodology* precisely. The difference is that 'the term *methodology* refers to the theory of how research should be undertaken and 'the term *methods* refers to techniques and procedures used to obtain and analyse data'. Hence, research methods depicts the particular research techniques, regression analysis, tools or procedures applied in the research to achieve the research objectives.

To cover these objectives, section 3.2 includes an introduction to the research philosophy and justification of my ontological, epistemological, and axiological position. Section 3.3 presents the research approach. Section 3.4 describes the research design. Finally, section 3.5 describes the research process followed in this work.

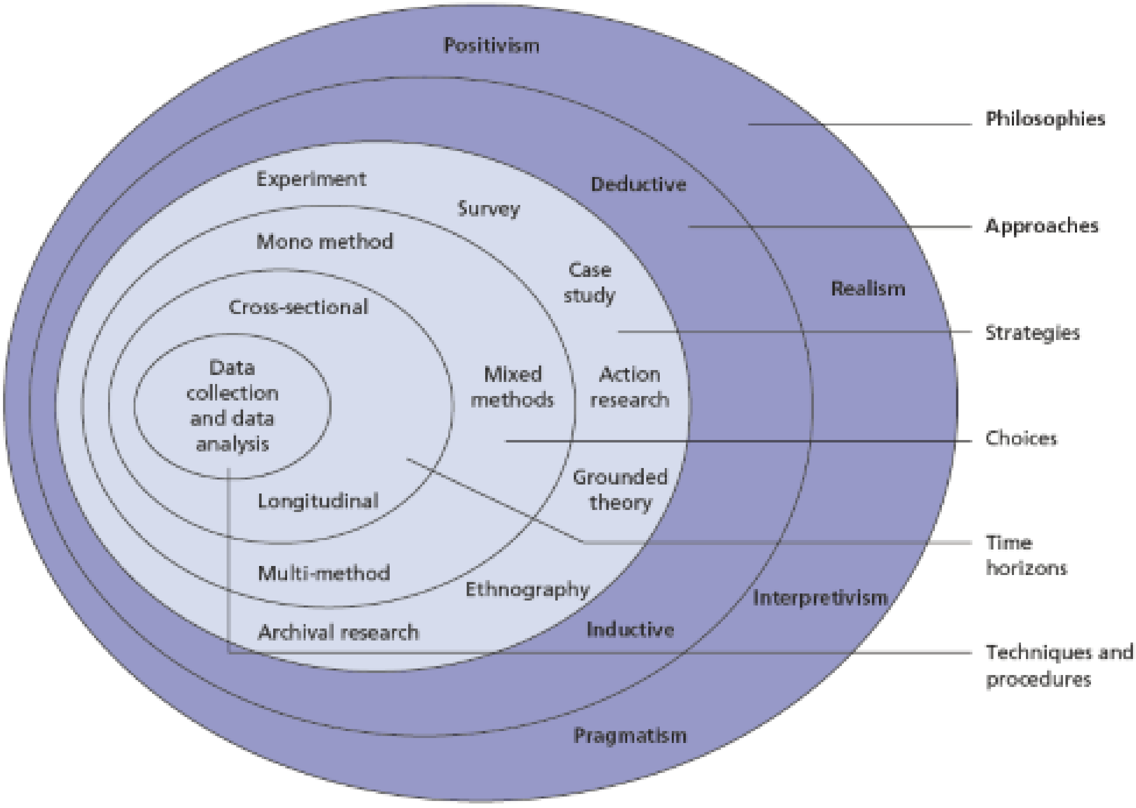


Fig. 3.1 The research onion (Saunders et al., 2009)

Figure 3.1 shows the 'research onion' as proposed by Saunders et al. (2009) including my own position for this work, which will be justified in details in this chapter. The first two outer layers represent the research philosophy and research approach, and the next three layers represent the research strategy, research choices and time horizon, which will be covered in the research design in section 3.4.

3.2 Research Philosophy

The research philosophy concerns the type of knowledge and how it is developed (Saunders et al., 2009). It holds assumptions about how the world is viewed and the influence of the researcher's own values, beliefs and assumptions on the research design. There are three key ways of thinking about research philosophy, all of which impact the research process: ontology, epistemology, and axiology. Table 3.1 presents a summary overview of the addressed philosophical positions.

Table 3.1 Research philosophies, adapted from (Saunders et al., 2009)

	Positivism	**Realism**	**Interpretivism**	**Pragmatism**
Ontology	Objective and independent of social actors	Objective. Exists independently of human thoughts and beliefs (direct realist), but is interpreted through social conditions (critical realist)	Subjective, socially constructed, may change, multiple	Multiple, view chosen to best enable answering of research questions
Epistemology	Only observable, measurable facts can provide credible knowledge	Observable phenomena provide credible knowledge, Focus on explaining within context	Focus on details of situation, reality behind these details. Subjective meanings motivate actions	Dependent on research questions, either both observable phenomena and subjective meanings can provide acceptable knowledge
Axiology	Value-free	Researcher's values bias findings	Researcher is a part of what is being research	Researcher's values plays a large role in interpreting results

3.2.1 Ontology

According to Saunders et al. (2009), ontology is a view of the nature of reality and expresses the researcher's beliefs about the way the **world** works. Ontology answers the central question of whether social entities are considered objective or subjective. Hence, we can distinguish between two contrary ontological aspects: objectivism and subjectivism.

Objectivism is the notion that 'social entities exist in reality external to social actors concerned with their existence' (Saunders et al., 2009, p.110). Alternatively, this position 'asserts that social phenomena and their meaning have an existence that is independent of social actors' (Bryman, 2012, p.33).

Subjectivism , is also known as constructivism, and represents the position that 'social phenomena are created from the perceptions and consequent actions of social actors' (Saunders et al., 2009, p.111). Alternatively, this position 'asserts that social phenomena and their meanings are continually being accomplished by social actors' (Bryman, 2012, p.33).

3.2.2 Epistemology

According to Saunders et al. (2009), epistemology is a view regarding the nature of **knowledge**, how knowledge is generated and what constitutes acceptable knowledge. Carson et al. (2001) defines it as the relationship between the researcher and the real world and how reality is captured in the research.

Positivism 'advocates the application of the methods of the natural sciences to the study of social reality and beyond' (Bryman, 2012, p.28). Hence, valid knowledge is scientific knowledge based on facts, sensory experience, or observable phenomena. When adopting positivism as a position, the researcher should be independent from his research and purely objective (Wilson, 2014).

Positivism has influenced economics by focusing on knowledge about values as being positivistic knowledge when they are quantifiable and/or demonstrable and highlights the importance of objectivity in economic research. However, economists cannot fully adopt positivism because many things that are not concrete are nonetheless real. Senses are not the only means through which experience leads to knowledge. An existing value could be also seen through a logical process of conceptualisation (Ethridge, 2004).

Empiricism describes a general approach to studying the reality that suggests that knowledge is acceptable only if it is gained through sensory experiences. Experience and evidence are considered fundamental in the genesis and development of ideas (Bryman, 2012, p.23). It is the view that only perception by means of the physical senses is really there, and that these perceptions are accurate reflections of it. Hence, truth for an empiricist, is based on evidence.

By derogation from positivism, empiricism includes measurement but extends into quantification and estimation of relationships (Ethridge, 2004). Caldwell (2003) characterised logical positivism as evolution toward logical empiricism, which is a merging of logical positivism and pragmatism. The philosophy of logical positivism has influenced economics by emphasising measurements and quantification, and influenced new methods and methods in statistics and econometrics.

Realism relates to a scientific approach to collecting data and developing knowledge, however it believes that there is a reality, which exists independently of the human mind (Saunders et al., 2009, p.114). Two major forms of realism can be recognised:

- Empirical realism asserts that reality of nature can be understood using appropriate methods (Bryman, 2012). Because 'what you see is what you get', it is also called 'native' realism or 'direct' realism to express the correspondence between description and reality (Saunders et al., 2009).
- Critical realism asserts that reality of nature can be understood by identifying structures of the work (Bryman, 2012). It is based on experiencing the world through 'images of the things in the real world, not the things directly' (Saunders et al., 2009, p.115).

Interpretivism appreciates differences between human being and the rest of the natural world. It requires researchers to recognise that actions in the social realm have subjective meaning (Bryman, 2012, p.30). Research is based on understanding of human behaviour and differences between humans rather than explaining this behaviour, as in the positivism approach.

Pragmatism sees the research question as the most important determinant in research. It recognises that adopting one philosophical position is in some cases unrealistic, hence a variation of ontological and epistemological considerations are needed (Saunders et al., 2009, p.109). Pragmatism believes that interpreting the world can be done in different ways so that multiple point of views are needed to represent a complete picture of reality.

3.2.3 Axiology

According to Saunders et al. (2009), axiology is the role of a researcher's personal **values** on the research process. Values reflect the personal beliefs or feelings of the researcher. Here we can distinguish between two extreme positions: value-free and value-laden. The chosen philosophical approach reflects how large or small a role the researcher values play on the research results.

Value-free insists that only objective and unbiased research is valid. Researchers are independent from the data they deal with and must maintain a position of objectivity. However, this view vanishes with time. It is acknowledged that research cannot be unbiased due to the self-reflection, which plays an important role in some research (Saunders et al., 2009).

Value-laden accepts that a researcher carries an inherent bias due to his world views, cultural background, and assumptions, which will influence the research. The depth of that influence may vary from slight bias in interpreting results to being a part of what is being researched with outstanding subjectivity (Saunders et al., 2009).

3.2.4 Rationale of Philosophical Position

Generating new knowledge will be done by interpreting reality because I try to interpret and understand the world with its local meaning in time and place (Thietart, 2001). However, I use the causal approach, which leads me to examine the reasons for the presence of 'facts' through data but also accepts the possibility of multiple or circular causality. Unlike pure positivists, I seek understanding 'verstehen' the real world and the meaning actors give to reality.

Knowledge gained is not purely positivistic because it is not knowledge of conditions nor directly observable or measurable, but a normativistic conditionally prescriptive.

This knowledge is value-free, not done to solve a specific problem but concerned with good and bad in a particular situation and inherently embodies judgement, hence it helps to make decisions and take actions (Ethridge, 2004).

Due to my engineering background, I trust knowledge based on numbers and feel comfortable working with variables and models. However, I believe that the theories, background, knowledge and values that I have obtained over years of experience have an influence on what I observe. This constructed view of the world based on personal perceptions will generate conjectural knowledge. Nevertheless, I still approach objectivity by recognising the possible effects of biases even if it will not be achieved perfectly. Based on these arguments, I see myself as 'post-positivist'.

3.3 Research Approach

The design of the research strategy is based on the research approach selected, which defines the relationship between theory and research. Two main research approaches can be distinguished, which differ in their starting point and the desired outcome: inductive or deductive reasoning.

By linking these research approaches to the research philosophies discussed above, it can be asserted that the deduction approach draws more on positivism and induction than on interpretivism, though this cannot be a general statement (Saunders et al., 2009).

3.3.1 Deductive Approach

The deductive approach represents a scientific approach, in which theory guides research. Hence, it starts with the development of hypotheses based on the existing theory and ends by testing these hypotheses against observations (Wilson, 2014). The deductive approach follows a structured methodology in explaining causal relationships between variables and ensures quantitative concepts to measure facts (Saunders et al., 2009).

Figure 3.2 shows the 'top-down' process of deductive research, in which the existing theory is used to build a research hypotheses to be tested and confirmed.

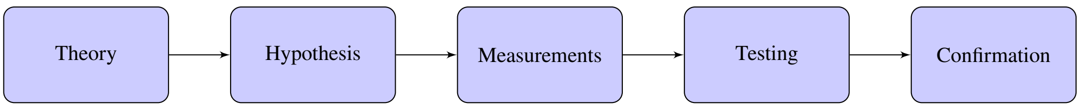


Fig. 3.2 Deductive approach

3.3.2 Inductive Approach

The inductive approach starts with detailed observations of the real world, searches for patterns and relationships from these observations, develops an explanation for these patterns through a series of hypotheses, and ends by drawing conclusions and generating a theory as an outcome of the research (Cooper and Schnindler, 2013). However, the inductive approach does not prevent researchers from using the existing theory to construct research questions (Saunders et al., 2009). The inductive approach is more interested in understanding the nature of the problem. Figure 3.3 shows the 'bottom-up' process of inductive research, in which detailed observations are used to build theoretical abstraction.

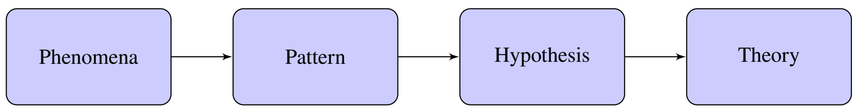


Fig. 3.3 Inductive approach

3.3.3 Abductive Approach

The abductive approach combines both inductive and deductive approaches within the same research. It overcomes weaknesses in applying a deductive or inductive approach by embracing pragmatism (Bryman, 2012), to explain 'incomplete observations' or 'surprising facts' if the researcher is faced with empirical phenomena that cannot be explained using the existing theories (Saunders et al., 2009).

Ethridge (2004) addressed several reasons for the need for an abductive approach. Firstly, each isolated approach leaves a gap in our understanding. On the one hand, deductive reasoning enables organisation of existing knowledge and deduction of new relationships, which are insufficient to be considered as new knowledge. On the other hand, inductive reasoning fails to use prior knowledge, so it is inefficient. Secondly, deduction and theorising alone are fallible when studying the real world because of the lack of evaluation. Induction and empiricism alone are fallible because of the probability of errors. Because no theory could be confirmed without facts and vice versa, the scientific approach must rely on both deductive and inductive approaches in constant interaction with each other.

Rationale of Research Approach

Based on the nature of the research questions and the abundance sources of theory, I intend to use a deductive approach. In general, economic theory is developed based on logical reasoning and a set of plausible assumptions to explain how variables are determined in the real world. The traditional approach to conducting empirical analysis in economics is the positivist position based on a deductive approach and using quantitative data (Seddighi, 2012).

I am interested to know how well deductive models explain the real world relationship and to find empirical evidence to confirm or reject a particular model.

3.4 Research Design

Vogt (2005, p.196) describes research design as the 'science and art of planning procedures for conducting studies so as to get the most valid findings'. It is the framework that brings the various component of the research project together (Thietart, 2001, p.111) and 'involves the intersection of philosophy, strategies of inquiry, and specific methods' (Creswell, 2013). Saunders et al. (2009) refined the research design methods within three layers, each of which focuses on a different aspect as illustrated in figure 3.1. However, Bryman (2012) used a compact approach, which I will use in describing my research design.

Research Purpose

Based on the research purpose, a research design can be exploratory, descriptive or causal, which differ in their main characteristics as summarised in table 3.3.

Exploratory Research aims to gain better insight into specific phenomena of the research area to understand the nature of a problem without having information about how such a problem has been solved in the past (Sekaran, 2003). Exploratory research can be conducted via searching in the literature, interviewing experts, or interviewing groups following an inductive approach. A researcher may change the direction of the research according to the new evidences gained during the research process.

Descriptive Research aims to describe the characteristics of an existing situation or phenomena (Sekaran, 2003). Descriptive research is formalised and structured, has clear hypotheses or research questions, and can be either qualitative or quantitative (Cooper and Schnindler, 2013). Furthermore, it is possible that descriptive research is used as a part of causal research (Saunders et al., 2009).

Causal Research is also called explanatory research and aims to establish causal relationships between variables or to explain cause-effect relationships between variables studying the situation in which they emerge. It is concerned with answering the question 'why' (Wilson, 2014).

Cooper and Schnindler (2013) distinguish between three degrees of causality: absolute, conditional, and contributory. Absolute causality 'means the cause is necessary and sufficient to bring about effect', while conditional causality 'means that a cause is necessary and not sufficient to bring about effect'. Finally, contributory causality 'means that a cause need be neither necessary nor sufficient to bring about an effect'. Nevertheless, causality may be argued using these three factors, which can be clarified by testing the relationship.

Table 3.3 Main characteristics of research designs (Zikmund et al., 2009)

	Exploratory	**Descriptive**	**Causal**
How uncertain is the decision situation	Highly ambiguous	Partially defined	Clearly defined
Key Research Statement	Research question	Research question	Research hypothesis
When Conducted?	Early stage of decision making	Later stages of decision making	Later stages of decision making
Research Approach	Unstructured	Structured	Highly Structured
Nature of Results	Discovery oriented	Can be confirmatory	Confirmatory oriented

3.4.1 Research Strategy

The research strategy outlines the general orientation for conducting research, which can be qualitative, quantitative or a mix of both (Bryman, 2012). The terms quantitative and qualitative are widely employed to 'differentiate both data collection techniques and data analysis procedures' (Saunders et al., 2009, P.151). The historical trend evolution of the research approach shows that quantitative approaches were dominant in social science until the mid 20th century. Later, qualitative approaches gained more interest. Figure 3.5 shows an overview of the main differences between quantitative and qualitative research strategies, creating a link between research philosophy and research strategy.

Table 3.5 Differences between research strategies (Bryman, 2012)

Orientation	**Quantitative**	**Qualitative**
Approach	Deductive	Inductive
Role of theory	Testing of theory	Generation of theory
Epistemology	Positivism	Interpretivism
Ontology	Objectivism	Subjectivism

Qualitative Approach describes techniques and procedures for data collection or analysis that produce or use non-numerical data which are hard to quantify or deals with information (Saunders et al., 2009). It aims to explore and understand an individual or group phenomenon from the participants point of view (Creswell, 2013), in which the researcher is himself part of the research. The term qualitative is a synonym for categorical or nominal (Vogt, 2005). Qualitative data are obtained by interviewing individuals who may help the researcher to understand phenomena.

Quantitative Approach describes techniques and procedures for data collection or analysis that produce or use data and can be handled numerically (Saunders et al., 2009). It aims to test objective theories by examining relationships among the measured variables of a research area (Creswell, 2013), in which the researcher is an objective observer who does not influence the research. Quantitative data are numbered data obtained by questionnaires, which can be analysed using statistical procedures.

Mixed Approach is a term that is used to express combining or associating methods associated with both quantitative and qualitative approaches. Hence, the study's overall strength outweighs that a study employing only a quantitative or qualitative approach (Creswell, 2013).

3.4.2 Methods for Research Design

There are different methods and choices to structure the research design. Saunders et al. (2009) describes various forms of research including experimental, survey, archival, ethnography, case study, action research, and grounded theory. Further, the choice of method can be a mono method, mixed method, or multi-method. Finally, the time horizon of the study can be cross-sectional or longitudinal. Bryman (2012) used a mixed approach for these aspects to address the fact that a research design can be experimental, comparative, cross-sectional, longitudinal, or based on a case study. Vogt (2005) provides some valuable definitions:

Cross-sectional: Survey or structured observation based on samples taken at one, specific moment in time.

Longitudinal: Survey or structured observation based on samples taken over time, such as a panel study.

Experimental: Study of causal links between variables by controlling some conditions and checking whether manipulation of one independent variable results in changes to another dependent variable. It is employed in exploratory and explanatory research to answer questions of 'how' and 'why' (Saunders et al., 2009).

Case study: Survey on a single case on a small number of cases as a way of studying broader phenomenon. This allows intensive empirical investigations about individuals, groups, societies, or events in a real life context.

Action research: is applied research that can find the most effective way to make a desired change or to solve an organisational issue within a specific context and clear purpose.

Comparative: Studying more than one event, group, or society to isolate variables that explain patterns. However, in a general sense almost all systematic research can be considered comparative.

Figure 3.7 brings the research strategy together with the methods for the research design. The mixed strategy of combining quantitative and qualitative methods is also possible if both strategies are set as typical forms.

Table 3.7 Linking research strategy with research design methods (Bryman, 2012)

Design Method	Quantitative	Qualitative
Experimental	Typical form	No typical form
Cross-sectional	Typical form	Typical form
Longitudinal	Typical form	Typical form
Case study	Typical form	Typical form
Action research	No typical form	Typical form
Comparative	Typical form	Typical form

3.4.3 Rationale of Research Design

In accordance with my epistemological position, it is my intention to use a quantitative approach to analyse data and test the constructed model. My computer programming knowledge will help me to work with software tools for quantitative analyses. The research design is based on longitudinal because it covers the behaviour of German firms over time. Longitudinal data analysis combines time series analysis with regression analysis. Time series analysis allows investigation of the dynamic aspect of the studied relationships. However, regression analysis allows statements to be made about the measurement by controlling other variables (Frees, 2004).

Two kinds of the above mentioned research designs will be used in this work: causal and descriptive. Causal design will be used to examine the cause-effect relationship between innovation and productivity and the impact of determinants on this relationship. However, descriptive design will be used as part of this research to characterise the available quantitative data such as frequencies, mean, standard deviation and other characteristics that describe phenomena related to the main topic.

3.5 Research Process

The process followed in this research consists of the following stages. Each stage is based on my assumptions about the source and nature of the obtained knowledge. Each chapter of this work deals with one stage or more of the research process. The link between each research stage and this work is also provided in the chapter number. Figure 3.4 shows an overview of the stages of the research process followed in this work.

- Stage 1: Identifying the problem and clarifying the importance of the research topic and the motivation of the work. This is covered in chapter 1.
- Stage 2: Stating the research aim, formulating research questions and research objectives to provide an explicit statement about what this research wants to achieve. This is covered in chapter 1.
- Stage 3: Conducting a critical examination of the literature review using a wide range of existing books, studies, articles, analyses and technical reports regarding the topic. This is covered in chapter 2.
- Stage 4: Constructing the conceptual framework and developing the research hypotheses. This is available in chapter 2.
- Stage 5: Defining the philosophical position and research approach. This is covered in chapter 3.
- Stage 6: Secondary data will be used, therefore it is necessary to prepare panel data for analysis and ensure that all model variables are represented and available in it. This is done in chapter 4.
- Stage 7: Conducting descriptive statistics to deliver key information on data. This is done in chapter 4.
- Stage 8: Describing the econometric model and justifying the estimation strategy. This is done in chapter 5.
- Stage 9: Conducting data analysis to test the research hypotheses and interpreting results. This is done in chapter 6.
- Stage 10: Reaching a conclusion and justifying the achievements of the research aim and objectives. This is covered in chapter 7.

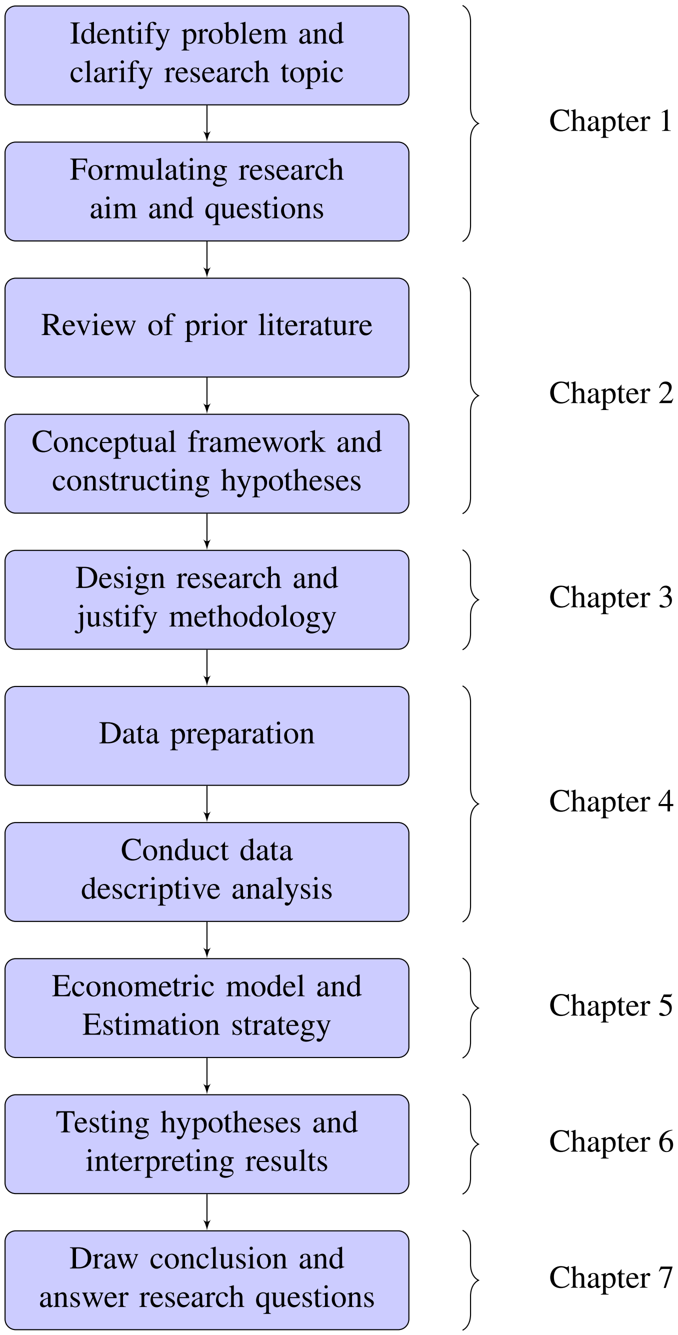


Fig. 3.4 Plan of research process

Chapter 4

The Data

4.1 Introduction

In chapter 2 the empirical literature shows that innovation may have an impact on labour productivity. Additionally, a set of relevant determinants that may affect this relationship have been defined. In chapter 3, the rationale for applying a quantitative approach using longitudinal data is presented. This chapter describes the secondary data of the MIP that will be used in this research to test the relationship between innovation and productivity and the impact of addressed determinants.

Section 4.2 provides the rationale for using secondary data. Section 4.3 contains advantages of using panel data and issues resulting from it. Section 4.7. describes the dataset of MIP. Section 4.6 summarises the main characteristics of the dataset. Section 4.7 addresses known issues of MIP. Section 4.8 describes the process carried out to prepare panel data for analyses in order to keep a high number of samples and records in the data. Finally, section 4.9 presents some key information and descriptive statistics.

4.2 Using Secondary Data

According to Saunders et al. (2009), survey-based secondary data is data gathered using questionnaires, which are analysed according to the survey's original aims. Wilson (2014) mentioned that the topic plays a role in that some research projects are entirely based on secondary data if large data is planed to be used in the research design, as is the case in this work. The need for longitudinal survey data on productivity and innovation over several years made using primary data unfeasible.

Using secondary data to conduct research has some advantages and disadvantages. The main advantage is that the data already exists. The collection of a huge database covering such a large amount of firms on national level, and based on a detailed regular surveys, data evaluation, data maintenance and error corrections is possible only by institutional working, which involves a number of researchers and workers could not be collected by individuals. This kind of data made a longitudinal study design feasible.

(Wilson, 2014) addressed the concern that using secondary data may have some disadvantages. Firstly, access to reliable and high quality data may be difficult and time-consuming. A special agreement with ZEW was needed to access the MIP data in this case. Secondly, data may not match the research problem, though this was not the case in this study because the data serves very well to answer the research questions. Thirdly, it may be difficult to verify the reliability of secondary data. Therefore, data quality and data issues are discussed in the next paragraphs in detail. Fourthly, the available data is not in a manageable form. The delivered format of MIP data is easily manageable. Fifthly, a possible issue is comparability.

Because the main intention was to investigate the impact of ICT on innovation and productivity, additional data on ICT was needed. It was not possible to investigate this because sample firms within ICT survey data have different identifiers for those in the innovation survey, which made considering ICT data unfeasible and led to cancelling of this objective.

4.3 Using Panel Data

Panel data, also called longitudinal data, is a dataset based on repeatedly surveying the behaviour of the same entities or individuals over time (Frees, 2004), which combines time series and cross sectional data. Panel data allows constructing and testing of realistic behavioural models, which cannot be analysed using only cross sectional data or time series data. The information technology revolution in the last decades has facilitated saving and processing such huge data and offered tools to analyse them.

According to Baltagi (2005) and Hsiao (2014), panel data offers in general several advantages:

1. The increasing number of observations creates more variability by introducing combinations of variations across the micro entities over time. Hence, it increases the degree of freedom and reduces collinearity among explanatory variables, which leads to improvement in the efficiency of parameter estimates.

2. The ability to analyse complex behavioural models cannot be investigated using cross sectional or time series data such as dynamic behaviour of variables with long-term characteristics, which is the case for innovation.

3. In micro panel data, biases caused by aggregation over observed entities may be reduced or eliminated.

4. Potential to account for omitted variables and cross-sectional heterogeneity bias.

However, using panel data also brings a few challenges due to the limitation addressed by Baltagi (2005), which follows. These limitations will be discussed later and a solutions will be proposed to minimise their impact on the results:

1. Problems in design, collecting, and managing data may arise.

2. Potential for measurement error due to responses to unclear questions or misleading information.

3. Selectivity problems such as self-selectivity, non response, and attrition, which may lead to estimation bias.

4. Heterogeneity bias due to the difference across cross-sectional entities, which may not be reflected in the available data.

5. Autocorrelation of the error term if observations in one time period are dependent on error terms from the prior time period (Washington et al., 2011).

6. Heteroskedasticity due to the fact that the error term's variance is not constant across observations.

Two kinds of panel can be distinguished, balanced and unbalanced. Panel data is considered as balanced when each entity has one observation in each time period. Unbalanced panel data is when entities are not observed in all time periods e.g. due to missing values.

4.4 Mannheim Innovation Panel

The Mannheim Innovation Panel (MIP) data extracted is the German part of the Community Innovation Survey (CIS). The Centre of European Economic Research (The Centre for European Economic Research (ZEW)), commissioned by the German Federal Ministry of Education and Research The German Federal Ministry of Education and Research (BMBF), has been gathering data on innovation behaviour and the activities of the German economy since 1993 on a yearly basis.

The innovation survey provides qualitative and quantitative data on innovation activities using the enterprise as statistical unit. It allows policy makers to monitor innovation, evaluate innovation policies, and gain a better understanding of the innovation process. Innovation surveys were developed to collect information about innovation types in firms, reasons to innovate or not, cooperation with the public research sector, flow of data, and quantitative data on sales product innovation (OECD, 2010a).

The survey is conducted annually using methodology based on the Oslo Manual. The structure of the survey supports panel data generation because it includes the same firm identifier and refreshes samples each year to capture new firms or substitute firms which have left the market (Gottschalk, 2013). It facilitates comparison of firms across regions or industrial sectors. The last available survey is from 2013.

Using panel data supports investigation of the behaviour of investment in innovation and productivity because these are long-terms variables, which are not observable in cross sectional data. The content of the MIP data serves to answer the research questions because it contains proxies for the main variables and the indicators proxy the analysed determinants addressed in the literature review.

Data is 'factually anonymised', which mean that the dataset has been changed to make a re-identification of the participant firms possible only by investing excessive amounts of time, money and work (Gottschalk, 2013). Access to the aforementioned dataset has been achieved by signing a confidentiality agreement with ZEW. The deliverables are: a cross sectional dataset for each year readable with various statistical software tools, which includes the source questionnaire of the dataset, and a guide to the survey for external users.

4.5 Methodology of Data Collection

The Community Innovation Survey CIS draws on the methodology laid out in the Oslo Manual, a joint effort of the Directorate General of the European Commission (Eurostat) and the Organisation for Economic Co-operation and Development OECD to standardise innovation surveys internationally. The Oslo Manual defines the ways of quantitative measuring innovation's inputs and outputs in surveys, and provides a conceptual background for analysing innovation in firms. Furthermore, it presents an economic framework based on elements from Schumpeter and a subject approach for creating and diffusing knowledge (OECD, 2005).

The Oslo Manual presents itself as a set of 'Guidelines for Collecting and Interpreting innovation Data' (OECD, 2005, P.4) and is considered as a framework to standardise measurement of innovation internationally. It is used in 34 OECD member countries, 27 states of the European Union and a set of other countries world-wide. It emphasises the systems approach focused on aggregate indices to understand innovation and enhance a continuous learning process of data collection. Innovation survey data has become widely used as a source of innovation indicators. The resulting CIS indicators are used to generate national, European and OECD reports about innovation (OECD, 2010b).

The Oslo Manual first appeared in 1992, the second edition in 1997, and the third and last revision was published in 2005. Starting from the revision of 2005, Oslo Manual distinguishes between four kinds of innovations: product, process, organisational, and marketing innovation.

Mairesse and Mohnen (2010) and Arundel et al. (2013) differentiate between two data gathering approaches for innovation data, the object-based approach and the subject-based approach:

- The object approach considers the individual innovation as a unit of analysis, which collects information on specific types of innovations by evaluating announcements in media and trade journals. The advantage of this method is that it evaluates the quality of the innovation because firms announce only important differences from competitors (Arundel et al., 2013). However, this method collects less information about firms' innovation strategies and misses process and organisational innovation.
- The subject approach collects data for all innovation activities and outputs at firm level. The subject-based surveys collect a wide spread of information on innovation activi-

ties and cover all types of innovation at the level of decision making, which enables accounting and financial data (if available) to be obtained, merged with the innovation data. The disadvantage of this method is that it provides limited information about the characteristics of innovations (Arundel et al., 2013).

The methodology of the Oslo Manual and Community Innovation Survey follow the subject approach.

4.6 Data Characteristics

In the Community Innovation Survey enterprises are asked whether they have innovated during the period of the survey, differentiating different types of innovation. Additionally, enterprises are asked to give information on their R&D expenditure, knowledge inputs and behaviour, such as cooperation with other firms or institutions. It is important to mention that participation in CIS surveys is voluntary for firms. The survey captures a large set of indicators containing information in different data types: general information data, dichotomous data (yes/no), and qualitative categorical data, in addition to qualitative data and financial data, such as:

- General information: firm identification number, industry branch, region, scope of the market, physical assets, number of employees, personnel qualifications, and labour productivity.
- Indicators of innovation outputs: information on type of innovation, introducing new products or new processes, new marketing concepts, organisational changes, and share of sales for new products.
- Innovation expenditures: R&D expenditures, acquisition of patents and licenses, and personnel training.
- Innovation output: information on new products, processes, and services within firms and the degree of success in introducing them to the market.
- Information about the way of carrying out innovation: firms' behaviour like cooperation partners, the sources of information for innovation, protection policy, the reasons for innovating, obstacles to innovation, public funding, and research cooperation and partnership.
- Factors that hamper innovations: personnel, funding, information, etc.

Data includes a generic industry sector identifier aggregated into four categories from 2011 onward: research-intensive industry, other industry, knowledge-intensive services, and other services. Additionally, a more-detailed industrial sector identifier based on the ZEW indicators of the innovation panel (Gottschalk, 2013, p.136) is still available in the dataset. For records before 2011, the generic industry sector identifier has been calculated according to the transformation table defined in (Gottschalk, 2013, p.137).

Hence, CIS data supports studying the relationship between innovation inputs, innovation outputs and firm performance using different definitions. German data does not contain data from other sources. It contains all needed indicators to investigate firm performance in relationship with innovation, which give MIP data large potential for carrying over empirical analysis (Rammer and Peters, 2013).

Data distinguishes two types of enterprises. The first includes innovative enterprises, which can be of three kinds: successfully implemented, on-going, and abandoned before implementation. The second are non-innovative enterprises, which carried out no innovation activities during the survey period. Innovators are enterprises that presented new products or new processes where 'new' is defined as significantly improved or completely new, but 'new' can also mean newly introduced to the market or only new to the firm itself.

4.7 Data Issues

Identifying data issues and considering them in the research method and econometric model may minimise their impact on research findings. Data issues have been categorised in three levels: the first level concerns general issues, the second level considers issues on raw cross-sectional data, and the third level considers issues arising from the generation of panel data.

4.7.1 General Issues

Economists agree on the need for more information than the CIS data offers. The first issue is that the lack of financial data in CIS surveys hampers investigation of the important relationship between financing and innovation (Lööf, 2002). It may be due to the fact that innovation expenditure is not specified in firms' financial accounts (OECD, 2005). A second issue is that the lack of questions on outputs of process innovation is an important limitation of innovation surveys (Arundel et al., 2013). However, this is not relevant for German CIS data, which captures a variable of reduction in production costs caused by process innovation. A third issue is the bias towards product innovators and firms doing R&D because the survey question about what percentage of sales derive from newly introduced products does not collect data on process or organisational innovation (Hall, 2011b). A fourth issue is the complexity of merging information from other sources with the CIS data from a practical point of view.

Other data issues which have been addressed in the previous literature using CIS data is that surveys do not provide information about the general institutional environment, like the education system or job market (OECD, 2005). However, this issue has no impact on my work at national firm level because all observed firms operate in the same institutional environment and all investigated firms enjoy the same advantages and suffer from the same disadvantages. Hence, this effect is crucial for comparative studies across countries. Nevertheless, regional differences can be investigated in the MIP data.

4.7.2 Cross-Sectional Data Issues

Since CIS is a voluntary survey with no guarantee of participation or completeness, the quality of the cross-sectional data suffers due to the low response rate among firms (Mairesse and Mohnen, 2010). Rammer and Peters (2013) addressed data quality issues due to responding rate and participation. Furthermore, most variables resulting from innovation surveys in cross-sectional data are qualitative, subjective, censored, and selected (Mairesse and

Mohnen, 2010).

Whether a variable is qualitative or quantitative has advantages and disadvantages. Although qualitative data has less information than quantitative data, it is also less subject to errors in measurement (Mairesse and Mohnen, 2010).

The subjective nature of many of the quantitative and qualitative variables in CIS data is attributable to the fact they are based on personal judgement. Lööf (2002) argues that firms tend to exaggerate their self-reporting of innovativeness in order to achieve a better ranking, which may affect data quality. An example is the share of sales due to new products, which tends to be rounded to the nearest 5 per cent; this value is subjective because it is up to the reporter to define whether a product is new or improved, new to the market or new to the firm. Survey participants did not understand the definitions for specific types of innovation, especially non technical organisational and marketing innovations. Second, participants might misinterpret the required level of novelty of innovation and the difference between 'new', 'significantly improved', 'new to the market' and 'new to the firm'. Hall (2011b) addressed this issue because the term 'new' can be interpreted differently, which may lead to questionable results. Arundel et al. (2013) argues that due to its high subjective nature, the innovation survey could fail to correctly define innovative firms. Cognitive testing was used on innovation data from the state of Tasmania in Australia obtained by asking open-ended question to investigate the quality of reported innovation. The study found through open-ended questions that 35.3 per cent of self-reported non-innovators most important product describes a valid innovation, and around 19 per cent of self-reported innovators most important innovation was not in fact an innovation. This issue may be resolved by using a mix of patents and R&D expenditure outputs as an indicator for innovations. Arundel et al. (2013) argue that managers of large firms reject the Oslo Manual definition that 'new to the firm' is an innovation.

Another issue in MIP data set is that many variables of interest are not observed above a certain magnitude due to a censoring mechanism. For individual cases, in which firms show extreme behaviour e.g. R&D intensity over 25 per cent, values are censored to prevent recognition of these firms on the basis of these intensities. The upper limits used are selected according to the distribution of these intensities (Gottschalk, 2013). However, these values can be recognised by evaluating another variable included in the dataset to indicate censoring, which has the same variable name with an 'x' added to the end of the name.

Another kind of censoring carried out in MIP data is that some variables are ordinal categorical variables, in which a scaled range is given instead of a value e.g < 10, < 20, < 30. A variable is considered censored when all values within a certain range are converted into a single value (Guo and Fraser, 2009).

An additional issue in CIS data is sample selection, which means that samples are not randomly selected. The selectivity in the data is due to questions that have been asked only to a subset of firms, which are innovative firms. Information about non-innovating firms is therefore minimal. The selectivity bias should be considered and corrected using a selection equation. Rammer and Peters (2013) addressed a methodological issue due to the potential selection bias of responding firms with specific innovation behaviour. Non-responding firms are biased towards non-innovating firms may be due to the high costs of answering the questionnaire. On the other hand, there are innovative firms which do not conduct R&D. Arundel et al. (2013) found that 59 per cent of firms acquired their most important innovation externally with little internally-made creative effort.

The Oslo Manual defines a firm as innovative if it is 'one that has implemented an innovation during the period under review' (OECD, 2005, p.152). Identifying innovating firms thus involve identifying the percentage of firms that have brought in at least one new or significantly modified product or process within the three years before the end of the reference year. Nevertheless, sectors show strong annual fluctuation of this indicator due to the unbalanced nature of data caused by the 52 per cent non-response rate in the two succeeding years. This phenomena is may be associated with business cycle fluctuations (Rammer and Peters, 2013).

Mairesse and Mohnen (2010); Aschhoff (2013) reported the use of imputation in cross-section data for variables that were not available in previous years to substitute missing values in a firm's sample.

ZEW avoids the confidentiality problem characterising most other OECD countries databases by creating firm level data where the firm identification number is decoded. Decoded data containing anonymous firms is available for purely scientific and non-commercial purposes. In that way, bias can be reduced if a firm does not profit from over estimating its innovation activities.

The Oslo Manual describes innovation expenditures as the sum of R&D spending, engineering expenses, personnel training, capital expenses, and marketing expenses, which are related to the product or process innovation. (Mairesse and Mohnen, 2010) claim that innovation expenditure has a low quality or is not even answered in the surveys and only R&D expenditures are reported.

Hall (2011b) argues that defining and measuring real outputs and inputs and associated price deflators is a challenging task. Hence, inputs and outputs at firm level may increase or decrease owing to shifts in market power, which conventional price indices do not represent.

According to ZEW, the quality of MIP data has been improved by contacting enterprises to clarify unexpected answers for data on innovation expenditures, in response to specific unanswered questions. A considerable effort by MIP to achieve good data quality has been made in an attempt to improve accuracy, ensure logical consistency, reduce errors, or identify and correct them, validate data during and after collection, and avoid imputation by re-contacting enterprises to correcting errors. Improving sampling errors was done to ensure that there are enough units in the industry domain. Coverage errors are minimal due to inclusion of units that no longer exist. Limitation of measurement errors has been achieved by continuous testing of questionnaires, and training of the team involved. Non-response errors have been reduced by re-contacting enterprises to remind them to respond (Rammer and Peters, 2013).

4.7.3 Panel Data Issues

By generating the panel data from the annual cross-sectional data, sample representativeness at the industry branch level may be affected by the following problems:

- Innovation surveys and the methodical framework of the Oslo Manual, are being continuously improved by adding important new indicators, a new definition of innovation, or new survey methodology, which has led to comparability issues among different cross-sectional data.
- There is no guarantee that a firm's data may still be available for all survey waves; some existing firms may go out of business, or new firms may enter the data; here it is important to consider data bias caused by cancelling data for 'unsuccessful' firms that go out of business and to consider their boundary conditions before they exit the market. Hall (2011a) says that the CDM model uses cross-sectional estimates

and ignores the contribution made by the timing of innovation issues in relation to productivity. This issue is related to the nature of data because firms may appear in one survey and disappear in another. Hall (2011a) feels that the absence of such an overlap makes the study of time series impossible.

- In the MIP data, the sample will be cancelled for small and medium sized firms (from 5 to 499 employee) if they did not respond within four consecutive surveys. However, large firms remain in the sample even they did not respond. The sample will also be cancelled if the firm changes its main economic activity sector or if the number of employees drops below 5 (Aschhoff, 2013).
- Since the survey is expanding, each new survey may include new indicators or some indicators may have the same name but a different meaning compared with previous surveys. The possibility of solving this issue depends on the case. In some cases, transforming the old indicator to the newest one is possible, in other cases not.

These issues creates two challenging trade-offs: The first is between expanding the panel data in the time domain and using new variables in a way that does not cause the relevant cancellation of data records. The second trade-off is between the chosen time period of the panel data and the availability of firm data over the chosen time period, which may lead to cancelling the observation, hence reducing the statistical power of the results. In the generation of panel data, a minimum continuity of observed variables shall be ensured without using imputation, otherwise the data record has to be cancelled. In this sense, there is a trade off between the quality of the data and the quality of the analysis.

4.8 Data Preparation

4.8.1 Process of Data Preparation

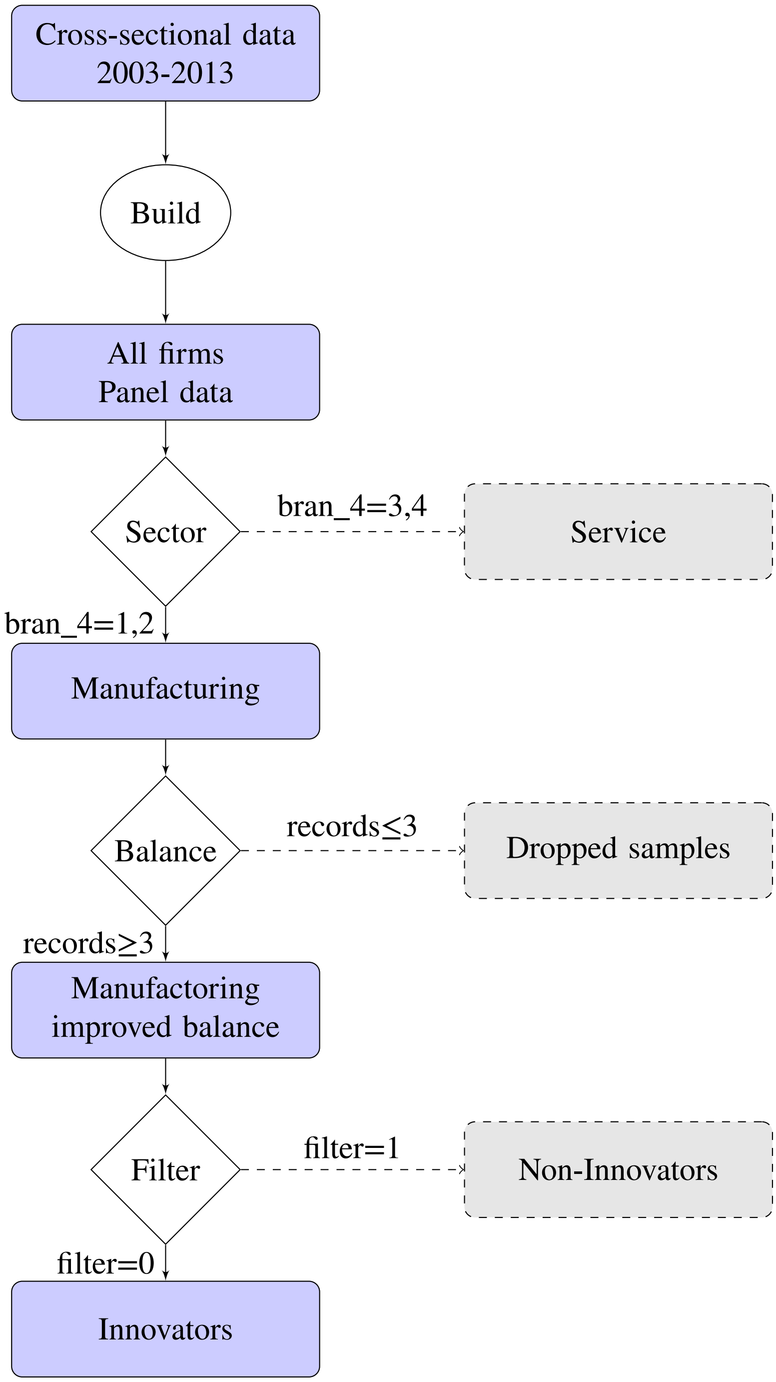


Fig. 4.1 Process of data preparation

To prepare the MIP dataset, several steps should be carried out before starting analyses. Figure 4.1 shows the process of work steps needed for data preparation and the selection of samples used to test the model.

4.8.2 Building Panel Data

The first step in data preparation is to generate the panel data from cross-sectional available datasets between the year 2003 and 2013. Many empirical studies such as Griffith et al. (2006), Lööf and Heshmati (2002c), Hall et al. (2008), Polder et al. (2009), Klomp and van Leeuwen (2001), Hall (2011a), Parisi et al. (2006), and Janz et al. (2004) have used cross-sectional data to analyse the CDM model. However, Lööf and Heshmati (2002c) addressed the need for more innovation surveys to generate panel data because the completely endogenous model of CDM is preferred if more surveys are available. In this study I intend to use the panel data approach due to the limitations addressed for using cross-sectional data for such an analysis, which are:

1. OECD (2005) emphasised the difficulty in capturing timing of innovation activities in periodical surveys because innovation is a continuous process. Therefore, the period of time that panel data covers is important for two reasons:
   - Innovation and its determinants are long term variables. Over a long period of time panel data reduces biased estimates caused by determinants with long term sensitivity.
   - Creating a panel data over a long period of time helps in studying the dynamics of innovation, correcting individual heterogeneity and in drawing conclusions about causality and addressing difficulties arising over time. Investing in R&D and ensuring financial support are strategic decisions which are taken simultaneously and are dependent on other environmental factors, which may not have been observed. The analysis of causality with these variables requires a structural model that considers the availability of panel data.
2. Knowledge accumulation in previous years and the long-term impact of innovation expenditure to get its fruit shall be considered. Crespi and Zuniga (2012) state that the nature of 'knowledge capital' is that it is based on the accumulated stock of R&D activities and investment in innovation activities in the previous years and cannot be captured in cross-sectional data.

3. The dynamic link in the recursive relationship between productivity and innovation cannot be investigated using cross-sectional data (Raymond et al., 2013).

4. Using panel data in addition to controlling selectivity and endogeneity bias, yields a lower estimated elasticity of innovation output (Rammer and Peters, 2013).

Differently to many other countries, the MIP German community innovation survey is conducted each year, which enhances generation of panel data. Before generating the panel dataset, it is necessary to check whether there are any repeated observations for firm identifier, industry identifier or year. If any are found, the observation shall be dropped.

4.8.3 Choosing Industry Sectors

MIP distinguishes between two main industrial sectors: the manufacturing sector and the service sector. The manufacturing sector is split into R&D-intensive manufacturing and other manufacturing. The service sector is split into knowledge-intensive services and other services. Table 4.1 shows an overview of the available data records for each year and for all years.

To limit investigated firm samples on manufacturing firms and exempt service sector from the analysis, an aggregate industrial identifier is used. Since 2011, an indicator has been available that contains four aggregate industrial categories: research-intensive industry, other industry, knowledge-intensive services, and other services.

MIP data shows that innovation indicators vary greatly from one industry to another and have various degrees of novelty; e.g. the share of innovation sales differs due to the product life cycle, which is pretty short in the IT industry compared to the chemical or pharmaceutical industries (Aschhoff, 2013).

A variable is generated for the years before 2011 based on a mapping table provided in the data guideline (Gottschalk, 2013, p.144-147). All detailed industrial sectors have been transformed into these four categories. Nevertheless, analysis can also be done on a detailed industrial sector.

Figure 4.2 shows the distribution of available firm samples in the industrial sector compared with all sectors (industry and service).

Table 4.1 Overview of firm distribution over industry sectors and year

Year	Number of samples	Industry	Service
2003	3549	1646	1903
2004	3249	1508	1741
2005	3928	1804	2124
2006	3436	1640	1796
2007	5207	2252	2955
2008	4459	2134	2325
2009	5546	2578	2968
2010	4778	2156	2622
2011	4218	1954	2264
2012	4154	1898	2256
2013	4426	1936	2490
All	20930	9008	11922

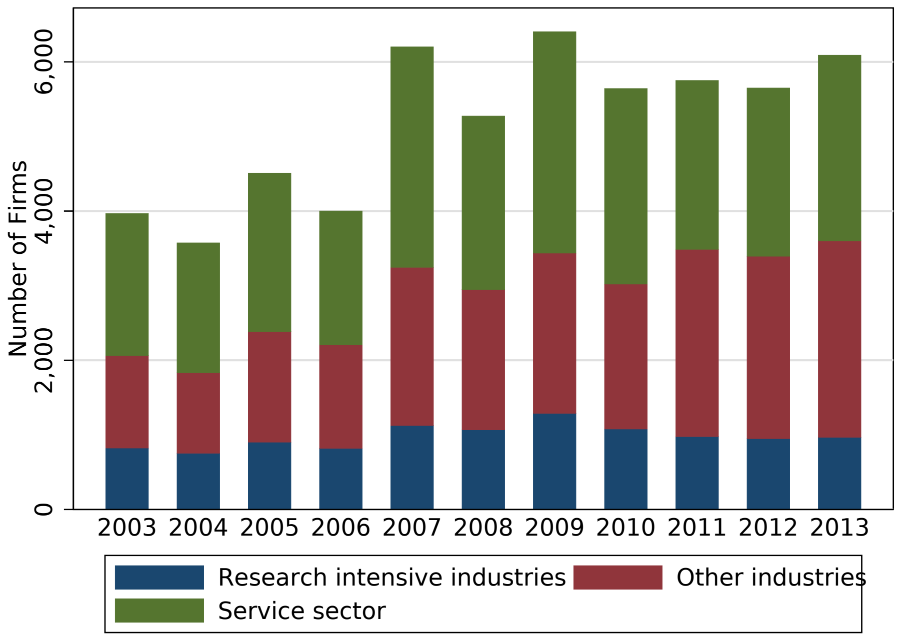


Fig. 4.2 Sample distribution of industry sectors over years

4.8.4 Balance Check

By definition, balanced panel data has the same number of time observations for each unit or sample. On the one hand, if a dataset is missing time frames for some cross-sectional units in the sample, the dataset is considered an unbalanced panel (Wooldridge, 2013). On the other hand, working with unbalanced data offers two clear advantages. Firstly, it allows broader observations, hence it ensures more accurate estimation. Secondly, it reduces bias because firms are allowed to enter and exit the samples at any year (Raymond et al., 2013).

Due to the conditions discussed above, the MIP data is highly unbalanced. Improving the balance of data without either non-random sampling bias or losing a large number of sample firms is a challenging task. Nevertheless, this may be achieved by analysing the frequency distribution of the samples to define a threshold of the required number of records available for that sample firm in the dataset.

The trade-off between improving the balance of the panel data on the one hand, and in the process not losing a large number of samples and reducing the statistical power on the other hand. A quantitative balance check is needed to investigate the availability of a firm sample in the cross-sectional data of each year. Based on this analysis, a threshold will be defined and if a firm sample has missing records in the panel data higher than that threshold, the sample has to be substituted.

A balance check as presented in figure 4.3 shows the data structure based on how many samples are available over the years of the panel. Each sample available for less than three years in the panel data will be dropped.

Table 4.3 represents the number of samples before and after the balance improvement, and the percent of samples dropped due to balance improvement.

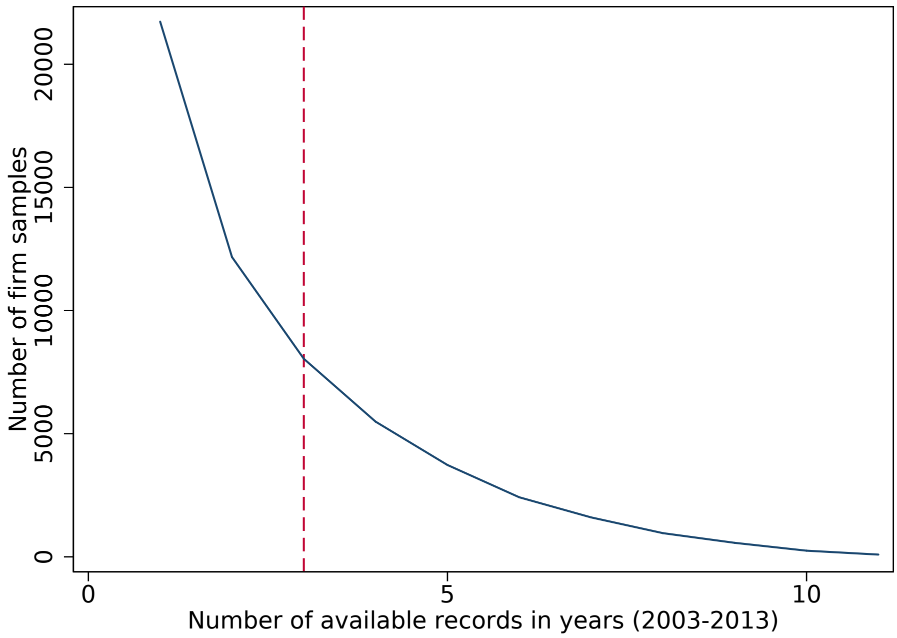


Fig. 4.3 Number of samples over panel period

Table 4.3 Number of samples before and after data balancing

Year	Before balancing	After balancing	Dropped samples	Share of dropped samples
2003	2064	1379	685	33.19
2004	1833	1181	652	35.57
2005	2385	1803	582	24.40
2006	2205	1746	459	20.82
2007	3246	1928	1318	40.60
2008	2949	2258	691	23.43
2009	3436	2358	1078	31.37
2010	3021	2335	686	22.71
2011	3487	2474	1013	29.05
2012	3394	2472	922	27.17
2013	3600	2168	1432	39.78
All	12356	4968	7388	59.79

4.8.5 Response Rate

In order to check whether missing observations are shaped by a certain mechanism or structure that might bias analyses, a comparison of the likelihood of responding to firm characteristics, mainly the firm size has been done. Figure 4.4 depicts the response rate for three different classes of firm size (small, medium-sized, and large firms). On the left side, the response rate for raw data and on the right side the response rate after balancing are presented.

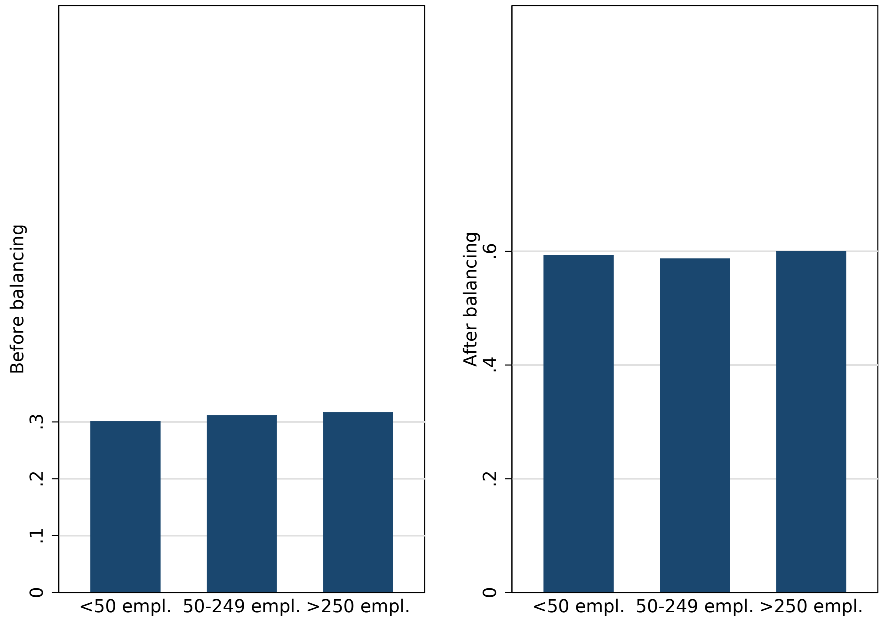


Fig. 4.4 Response rate before and after data balance

The left figure suggests that before balancing the response rate of large firms is slightly higher than that of medium-sized firms, which in turn tend to respond more often than small firms. A statistical test indeed rejected the null hypothesis of no dependence between response rate and firm size $\chi^2 = 83.3175$. However, the difference is small and may not systematically influence the results. The analyses will be performed with adjusted weights when checking robustness.

After balancing, the response rate is consequently much higher for all firm size categories compared to that for raw data, but the relationship between response rate and firm size category is almost the same. Hence, by balancing the sample, more small firm samples have been dropped from the data but the proportion among categories remained almost the same.

4.8.6 Missing Values within a Sample

If a single variable has a high percentage of missing values within the dataset, it is recommended that it be eliminated from the analyses to avoid losing a high number of samples from the data and hence reducing the statistical power. Stata removes such samples automatically while running analysis without deleting the sample itself from dataset.

On the one hand, in order to prevent a massive loss of observations due to a relatively high proportion of missing values, there is a need to determine which variables have a lot of missing values so that they can be excluded form the analyses to keep the number of samples as high as possible. On the other hand, some variables are important for this study and cannot be dropped or ignored from the analysis but these represent a pretty low number of values within the data.

To improve the availability of these important variables such as the intensity of physical capital (INVS), missing values are simply replaced by the mean of the existing values in previous and next year. This approach is justified by the assumption that the intensity of physical capital is a rather sticky value. As seen in table 4.5, the statistical impact of this improvement did not affect the statistical variance of the data.

Table 4.5 Descriptive statistics for physical investment before and after improvement

	count	mean	sd	min	max
invs (before)	8896	.0593268	.1162952	0	1
INVS (after)	19400	.0571678	.1038631	0	1

4.9 Summary Statistics

Before discussing the main topic of this work i.e. the relationship between innovation and productivity, this section provides at a glance the first results on data characteristics and relevant observations.

The raw MIP data considered in the analyses starts from the year 2003, when changes were carried out on the structure of the data. Considering previous years was combined with enormous effort without guarantee of success because types of indicators differ significantly. The last year available in scientific-file data is 2013.

Table 4.6 shows an overview of the available data records in each year, and for all years, and the number of innovator firms and non-innovator firms, classified according to whether or not a firm has presented innovations in the last three years.

Table 4.6 Overview of innovator firms distribution over year

	Innovators		Non-Innovators	
Year	Count	Share	Count	Share
2003	877	65.59 %	460	34.41 %
2004	821	70.65 %	341	29.35 %
2005	1328	76.06 %	418	23.94 %
2006	1147	67.43 %	554	32.57 %
2007	1343	70.61 %	559	29.39 %
2008	1478	66.22 %	754	33.78 %
2009	1584	68.39 %	732	31.61 %
2010	1501	65.23 %	800	34.77 %
2011	1563	67.75 %	744	32.25 %
2012	1301	56.69 %	994	43.31 %
2013	1249	61.92 %	768	38.08 %
All	3852	61.09 %	2453	38.91 %

Figure 4.5 shows the number of firm samples available in the dataset and the portion classified as innovator and as non-innovator:

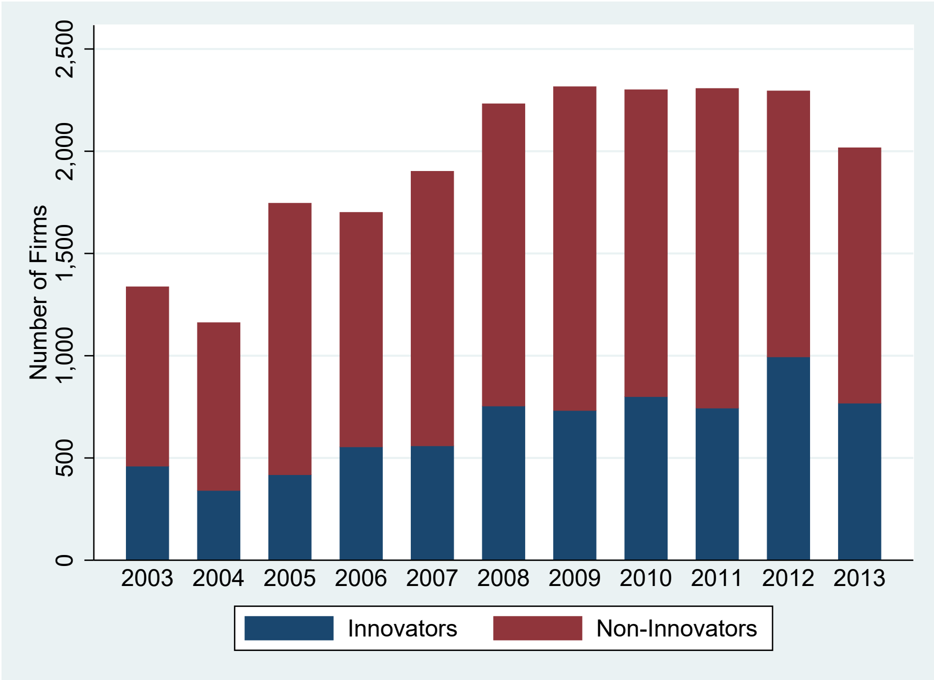


Fig. 4.5 Sample distribution of innovator firms over years

Figure 4.6 shows an overview of the distribution of firm size categories within innovator and non- innovator firms.

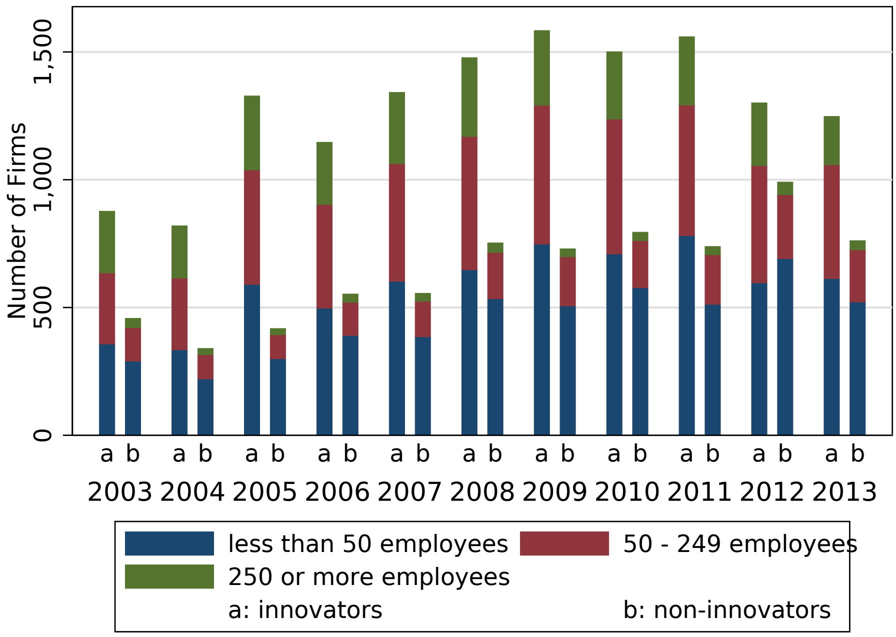


Fig. 4.6 Firm size of innovator and non-innovator firms

Figure 4.7 shows the distribution of available firm samples between innovative and non-innovative firms over investigated industry sectors.

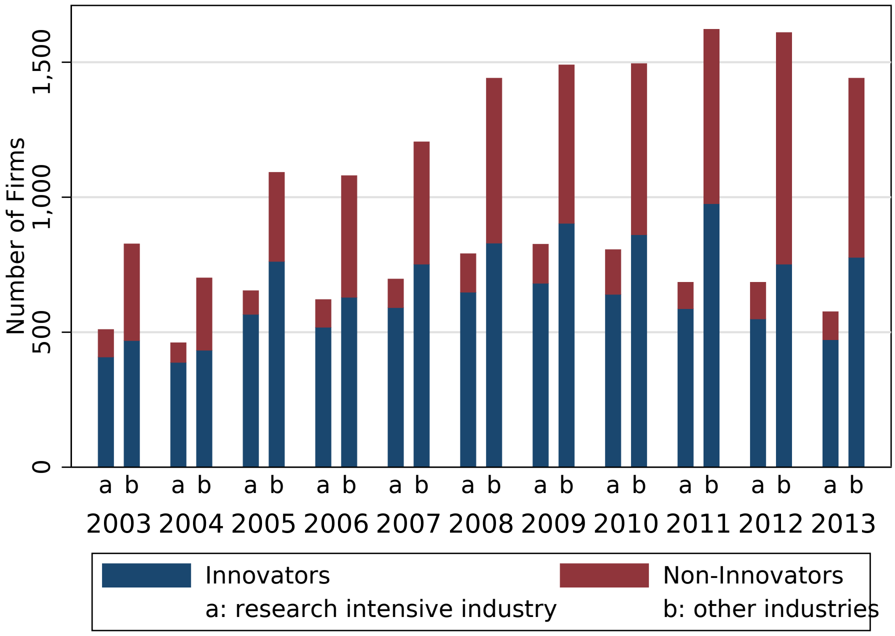


Fig. 4.7 Innovator firms samples over industry sectors

Figure 4.8 shows an overview of the distribution of firm size categories within research intensive industries and other industries. According to MIP dataset definition, firm size is categorised into three main groups based on the number of employees: ≤ 50 employee, from 50 to ≤ 249 employees, and ≥ 250 employees.

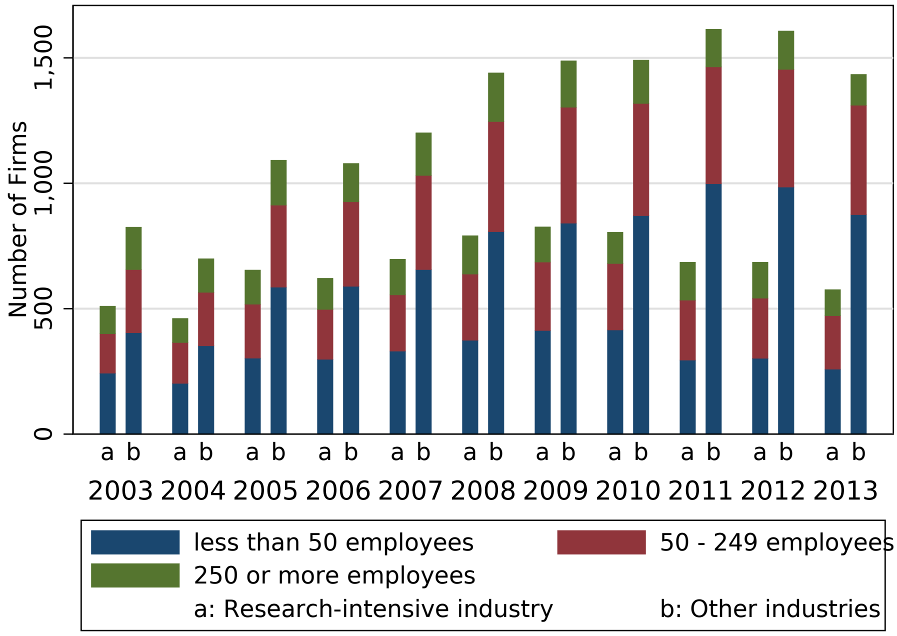


Fig. 4.8 Firm size of research intensive industries and other industries

Chapter 5

The Econometric Model

5.1 Introduction

The econometric model is a theoretical construct that reflects the predictions and expectations regarding what will be found in the data and represents a causal relationship between the investigated economic variables.

This chapter describes the econometric model and deals with the research methods and techniques needed to obtain the empirical evidence about the real world for theoretical understanding, and how to test the research hypotheses using German manufacturing data. It addresses relevant econometric issues that this study should deal with. Additionally, in chapter 4 some data issues were addressed, which need to be solved here.

Section 5.2 introduces the main econometric issues related to this study and proposes a remedy to deal with them. Section 5.3 describes the structural model and the mapping of variables and indicators available in the data. Section 5.4 describes the estimation strategy used in this work.

5.2 Handling Econometric Issues

There are some econometric issues which may seriously affect the parameter estimates in this study. Moreover, as mentioned in chapter 4, the usage of panel data is associated with potential issues such as heterogeneity. Other issues are those related to the use of MIP data such as censored variables. Another category are those related to the topic itself such as selection bias and endogeneity. Additionally, this section describes how to account for each issue to reduce its effect on the estimated parameters.

5.2.1 Heterogeneity

Heterogeneity refers to the fact that the observed micro entities across cross-sectional data are 'all different from one another in a fundamental unmeasured way' (Kennedy, 2008, p.282). These unobservable variables may affect the behaviour of each utility in cross-sectional data with the same intercept. Ignoring such heterogeneity may lead to bias and inconsistent parameter estimates unless the influence of these omitted variables is uncorrelated with the explanatory variables.

To deal with this issue and improve estimation results, there are two different ways to model how the intercept varies in the data by exploring the relationship between predictor and response variables. Hence, variation of the intercept in panel data can be classified into two types associated with the presence of fixed effects and random effects:

1. Variation from observation to observation **within** an individual entity (fixed effects).
2. Variation **between** individual entities (random effects).

Because using an appropriate model for panel data affects the **consistency** and **efficiency** of the estimator, the model should be chosen based on the variation within or between data.

A consistent estimator is an estimator that tends to provide more precise estimates as the number of observations increases. An efficient estimator is an unbiased estimators that has the smallest variance (Vogt, 2005). Assume that we have a basic regression model such as:

$$y_{it} = \lambda_i + \beta X'_{it} + \epsilon_{it} \tag{5.1}$$

where λ_i captures the individual characteristics of the entity that differs from one entity to another but is constant over time, x'_{it} is the vector of explanatory variables, and ϵ_{it} is the error term.

Fixed Effects: The Fixed Effects (FE) model assumes that every entity has its own unique characteristics that could potentially have an influence on the predictor variables. Therefore, the fixed effect model controls for omitted variables with λ_i, which is correlated with the explanatory variables $E(\lambda_i|x_i) \neq 0$. However, the explanatory variables and the variable of individual characteristics are assumed to be exogenous $E(\epsilon_i|x_i, \lambda_i) = 0$.

Random Effects: The Random Effects (RE) model assumes that the variation of intercepts across entities is random and considered as a part of the error term, which is not correlated with the explanatory variables. The random effect model assumes that the variable that expresses the individual characteristics is purely random $E(\lambda_i) = 0$ and not correlated with the explanatory variables $E(\lambda_i|x_i) = 0$. However, the explanatory variables are assumed to be exogenous $E(v_{it}|x_i) = 0$, where $v_{it} = \epsilon_{it} + \lambda_i$.

Hausman Test: To ascertain whether the fixed effects model or the random effects model is appropriate, the Hausman test can be applied. Hausman (1978) derived a test to check whether the constant that captures the individual characteristics of the entity over time is uncorrelated with the explanatory variables. In this case, the suitable model is the random effects model, otherwise the fixed effects model is appropriate. In other words, the Hausman test checks whether the random effects estimate differs to an insignificant extent from the unbiased fixed effect estimate (Kennedy, 2008). The null hypothesis H_0 of Hausman test is:

$$H_0 \,:\, E(\lambda_i|x_{it}) = 0 \tag{5.2}$$

where x_{it} is the vector of explanatory variables for each entity and time period, and λ_i is the unobserved effects of the entity.

$$W = \chi^2(k) = (\beta_{RE} - \beta_{FE})' \widehat{\sum}^{-1} (\beta_{RE} - \beta_{FE}) \tag{5.3}$$

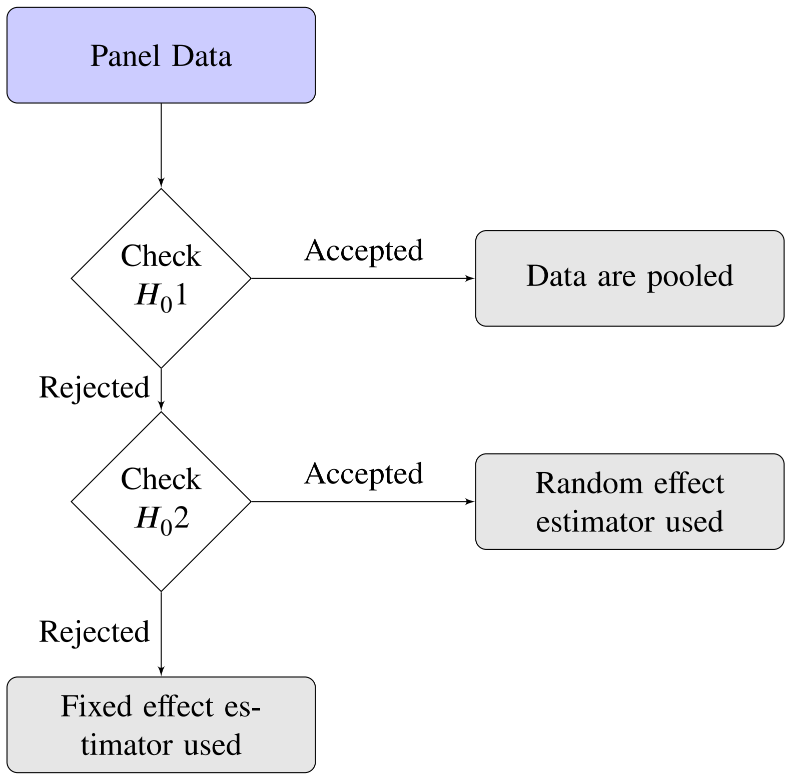
where k is the number of regressors, hence $k - 1$ is the degree of freedom. If W is significant, the null hypothesis is rejected and the random effects estimator should not be used.

Table 5.1 shows the relationship between the unobserved effects and the regressor to choose an appropriate estimator in terms of consistency and efficiency.

Table 5.1 Choosing appropriate estimator

H_0: Explanatory variables and unobserved effects are uncorrelated		
H_0	**Fixed effects**	**Random effects**
True	consistent	consistent
	inefficient	efficient
False	consistent	inconsistent

Figure 5.1 illustrates the procedure of the Hausman test according to Kennedy (2008) used for estimation of panel data. It starts with the null hypothesis that intercepts are equal. The data are pooled if the null is accepted, while if rejected a Hausman test should be applied. The test calculates the distributed Chi-squared for the null hypothesis that the random effect estimator is unbiased, and therefore appropriate. If the null hypothesis is rejected, the fixed effect estimator is used.



$H_0 1$: Intercepts are equal
$H_0 2$: Random effect estimator unbiased

Fig. 5.1 Procedure of Hausman test, own creation based on (Kennedy, 2008)

5.2.2 Multicollinearity

The multiple regression analysis determines the separate effects of the regressors on the dependent variable. If two or more of these regressors are highly correlated, multicollinearity exists and makes the determination difficult or impossible (Vogt, 2005). This does not depend on any theoretical linear relationship among regressors but is essentially a data problem caused by the existence of a linear relationship in the available data set. This could be due to several reasons such as independent variables having the same time trend, variables varied together due to a limited sample collection in the data, or one variable being the lagged of another (Kennedy, 2008).

The consequences of multicollinearity are that it leads to a large variance in the Ordinary least square (OLS) regression and inflates standard errors, which reduces confidence in reliable coefficient estimates so that their value is statistically insignificant. However, this undesirable high variance may also be caused by inadequate variation of the regressors in the data (Kennedy, 2008).

Multicollinearity can be reduced by collecting more data but solving it for a given data set can be worse than the problem itself because dropping other regressors that belongs in the model to reduce multicollinearity may cause a substantial change in the model and consequently lead to a bias in parameter estimates (Wooldridge, 2013). However, the existence of multicollinearity does not necessarily indicate that the coefficient estimates have unacceptably high variance (Kennedy, 2008).

Detecting multicollinearity can be done by using the condition index (the square root of the ratio of the largest to the smallest characteristic root). A condition index which is higher than 30 indicates collinearity. A more popular approach is to detect multicollinearity using the inverse of correlation matrix between all pairs of independent variables, which diagonal parameters called Variance Inflation Factor (VIF) given by the inverted tolerance $(1-R_i^2)^{-1}$, where R_i^2 is the R^2 resulting from the regression of the ith independent variable on all other independent variables (Kennedy, 2008). This indicates the degree to which the standard errors are inflated by a factor of VIF associated with collinearity.

The acceptable value of VIF follows the rule of thumb and varies in the literature. According to Kennedy (2008), having $VIF_i > 10$ indicates a problematic collinearity, which is the most used threshold. However, Rogerson (2001) recommends a maximum VIF value of 5 to indicate a problematic collinearity.

5.2.3 Endogeneity and Simultaneity

The process of transforming innovation activities into productivity is a multi- channel one that involves a simultaneous framework (Janz et al., 2004). Firstly, is very likely that innovative firms achieve higher productivity levels which motivate them to invest more in innovation activities, which in turn make them innovate further. If this causal relation has not been analysed simultaneously, the results would possibly be biased because some explanatory variables are not exogenous but jointly determined with dependent variables. Secondly, the dynamic nature of innovation requires consideration of a time lag in describing the process of generating innovation, which creates a endogeneity issue by neglecting it.

Endogeneity arises if an explanatory variable x is correlated with the error term u, which gives $Cov(x, u) \neq 0$. The difference between endogenous and exogenous variables is that endogenous variables are jointly dependent variables whose values are determined in the model, whereas exogenous variables are predetermined variables whose values are determined outside the model. Possible sources of endogeneity are omitting relevant variable from the model, measurement error in the right hand side variables of the model, and simultaneity between variables.

Simultaneity is an important form of endogeneity, in which there are at least two endogenous variables in one equation a dependent variable and in the other equation an explanatory variable. In this case, the estimation of the structural equations suffers from 'simultaneity bias' that has to be solved. A system approach offers solutions for the issue of simultaneity bias, which can be fulfilled by using simultaneous equation modelling.

Simultaneous Equations Modelling

A Simultaneous Equation Model (SEM) consists of a system of equations, also called a 'structural' form, representing a set of relationships among variables, and each equation has a causal interpretation (Wooldridge, 2013). The impact of labour productivity on the propensity to innovate and on innovation input is an important aspect of the assessment.

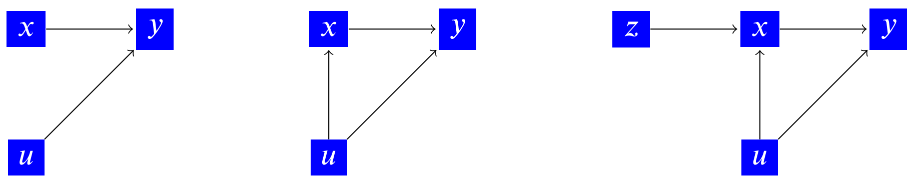
The advantage of the SEM approach is that it supports the testing of this simultaneity and reverse causalities between previous productivity which affect the firm's decision to engage in innovation activity and the level of expenditure on innovation (Kemp et al., 2002). According to Wooldridge (2013), the basic approach to estimating SEM with panel data may be carried out in two steps: eliminating unobserved effects as described above, and finding instrumental variables.

Instrumental Variables

As mentioned above, the possible correlation between independent variables and the error term may lead to biased and inconsistent estimation results. An instrumental variable is a variable used to replace an independent variable in the regression equation if that independent variable is highly correlated with the error term (Vogt, 2005). Hence, a good instrumental variable is highly correlated with the independent variable that it replace but uncorrelated with the error term. Wooldridge (2013) mentioned that the method of Instrumental Variables

IV is the leading method for estimating simultaneous equation models because it can eliminate possible bias caused by three sources:

1. Simultaneous causality bias from endogenous explanatory variables.
2. Omitted variable bias caused by an important variable that is correlated with the explanatory variable but is not observed, hence it cannot be considered in the regression.
3. Bias resulting from measurement errors.



Exdogeneity Endogeneity IV

Fig. 5.2 Exogeneity, Endogeneity, and IV

Figure 5.2 shows path diagrams of the relationship among the dependent variable y, the explanatory variable x and the error term u in case of exogeneity, endogeneity, and their relationship to the instrumental variable z (Katchova, 2013a). According to Wooldridge (2002), an instrumental variable should fulfil certain conditions:

1. The instrumental variable z is associated with the explanatory endogenous variable x, which means $Cov(x,z) \neq 0$.
2. The instrumental variable z is not associated with the error term u, which means $Cov(z,u) = 0$.
3. The instrumental variable must not be a direct cause of the dependent variable.
4. Error terms are assumed normally distributed

In this study, two measures will be carried out to reduce the potential of endogeneity: Firstly, each stage of the structured model will use the estimated value of the dependent variable from the previous stage as instrumental variable of main dependent variable. Additionally, a one-year lag of the main independent variable will be used.

The Identification Problem

A simultaneous equation system is considered as complete if there is a structural equation for each endogenous variable explaining its behaviour in the system. To correct the simultaneous bias, the endogenous variable will be replaced with an instrumental variable, which are the predictors obtained by regressing the endogenous variables on a set of exogenous variables in the equation system. The minimum set of exogenous variables needed to ensure consistency of the estimation is called the order condition for identification (Baltagi, 2011).

In this content, the identification problem is associated with a simultaneous equation system and is concerned with obtaining meaningful estimates of the parameters (Kennedy, 2008). A model (or the system of equations) is considered as identified if all equations are identified, otherwise estimating parameters of these equations is not possible. Vogt (2005) describes the possible status of an econometric model:

- Under-identified model: the model contains too many unknowns to be solved. Hence, it is not possible to estimate all of its parameters. This causes an identification problem.
- Just-identified model: the number of variables and the number of parameters to be estimated are equal.
- Over-identified model: the model contains more knowns than those necessary to estimate regression coefficients.

As mentioned above, instrumental variables should fulfil two assumptions: the independence between instruments and error term and the correlation with the endogenous regressor. A test is needed to check if the equation system is over-identifying restriction i.e. the number of instruments is larger than the number of variables on the right hand side. Further, the test relates to the first assumption and checks whether all instruments are relevant and exogenous (Baltagi, 2011).

5.2.4 Selectivity Bias

Innovative firms invest in innovation activities, hence excluding the non-innovative firms from the estimation will cause a sample selectivity issue and bias the estimation results.

A selected sample is a non-random sample caused by a 'selection mechanism' such as sample design or the behaviour of the sampled units so that population is carefully specified (Wooldridge, 2002), which can lead to an erroneous conclusion. In our case, selectivity bias is a relevant issue because this study investigates the behaviour of a specified group of (non-random) firms, which are innovative firms. However, non-innovative firms carry out innovation or conduct R&D even though they neither participate in the survey nor respond to it. Dropping these non-innovative samples will cause a selectivity bias in the estimation.

To correct this selection bias, the sample selection model proposed by Heckman (1979) will be employed in this work, which is widely used in several studies related to the topic such as Griffith et al. (2006); Mohnen et al. (2006); Raymond et al. (2013); Griliches (1994), and Robin and Schubert (2010).

The Heckman model procedure consists of the following steps:

1. Create a participation equation to indicate the propensity to innovate, i.e. whether or not the firm has presented innovation over a specific period or not. The estimation of a firm propensity to innovate will be carried out using a probit model.

2. Calculate the inverted Mills' ratio (IMR) in the first equation, which is a correction term to deal with the fact that investment intensity in innovation can be considered only for innovative firms.

3. Create the intensity equation to indicate that the intensity of investment in activities may lead to innovation. This equation can be consistently estimated only for firms that innovate by running a linear regression that includes both calculated IMR and firm's propensity to innovate.

5.2.5 Data Censoring

As discussed in the section on data description, to ensure anonymity, data of some dependent variables are either left- or right-censored. A variable is censored if sample values higher or lower than a specific threshold are suppressed at this threshold, which biases the estimation (Wooldridge, 2013). In some empirical literature, the terms 'censored' and 'truncated' are used synonymously. However, censored refers to incomplete measurements but truncated refers to incomplete samples (Vogt, 2005).

To minimise the potential bias caused by censoring of data, a Tobit model can be employed. The Tobit model sets parameters around the censored dependent variable to calculate the probability of being observed.

However, the Tobit model is only valid for random effects and therefore, executing a Hausman test to confirm consistency is not possible. Nevertheless, doping the censored samples may bias the results especially if the robustness test shows significant impact of the censoring indicator. Tobit model still the best estimator to use in this case.

5.3 The Structural Model

In order to test the relationship between innovation input, innovation output, and productivity, a structural equation model consisting of three stages and four equations is employed. Structural equation modelling is an important statistical technique which allows testing and estimation of a causal set of relationships between one or more independent variables and one or more dependent variables using quantitative data and qualitative assumptions (Alavifar et al., 2012). Moreover, the structural modelling approach supports the planed longitudinal data analysis because it offers more information about the investigated relationship which achieves better representation of the data (Frees, 2004).

The proposed model draws on the CDM model proposed by Crepon et al. (1998), which is in turn based on the Cobb-Douglas production function (Cobb, 1928) and on the knowledge production function (Griliches, 1979; Pakes and Griliches, 1984). The model considers Peters (2005) and Parisi et al. (2006) extension which added process innovation as outputs from the knowledge production function and inputs to the production function, and the extension of Polder et al. (2009) which added organisational innovation similarly. It also considers the proposed recursive relationship of labour productivity of prior time into the knowledge production function proposed by Raymond et al. (2013) and Baum et al. (2015).

The structural model consists of three stages. The first stage describes the firm's decision whether or not to take part in innovation activities, and the amount of expenditure on innovation. The second stage describes the knowledge production function in which inputs are the innovation expenditure and the outputs are product, process, and organisational innovations. The third stage describes the production function in which inputs are the resulting product, process and organisational innovation and the output is labour productivity.

In each equation two identifiers for panel data are used. The first identifier i is the firm identification number, where $(i = 1, .., N)$. It maintains the same identification for each firm sample of the panel data over time. The second identifier t is the year in which the observation has been captured, where $(t = 2003, .., 2013)$.

5.3.1 The Decision and Expenditure Stage

The first stage of the model is the decision and expenditure stage. Firms take an important decision regarding whether or not to take part in innovation activities, and based on the decision they define the amount of expenditure that they spend on innovation activities.

Because not all firms are engaged in innovation activities, there is a potential selection bias caused by using not randomly selected samples, which are the innovative firms. Hence, using these samples to estimate a relationship may lead to an erroneous conclusion. This issue could be solved using the model developed by Heckman (1979), which proposes three steps to express the determinants of the outcome using two equations: the participation equation and the regression equation.

- Step1: Using all observations (innovator and non-innovator), estimate a probit model where the decision to innovate is the dependent variable and the conditions that hamper or support innovation are the explanatory variables.
- Step2-a: Based on the parameter estimates, the IMR for each observation will be calculated.
- Step2-b: Run the regression using the selected samples, where the decision to innovate is the dependent variable and the IMR is an additional explanatory variable that has been corrected for sample selectivity.

In the first equation, the distinction between innovative and non-innovative firms may be made in two ways. Firstly, firms have been asked whether they produced innovations in the last three years, so they can be considered as innovators. Secondly, based on the amount of expenditure exceeding a defined threshold in the time period under investigation. In this study the second indicator is chosen to cover some firms that may have invested in innovation but without obtaining relevant output.

$$D_{it} = \begin{cases} 1 & \text{if} \quad D^*_{it} = \alpha_1^{'} x_{1\,it} + \gamma_1 P_{i,t-1} + u_{1\,it} > C \\ 0 & \text{if} \quad D^*_{it} = \alpha_1^{'} x_{1\,it} + \gamma_1 P_{i,t-1} + u_{1\,it} \leq C \end{cases} \tag{5.4}$$

where D^*_{it} is a latent dependent variable which expresses the firm's propensity to invest in innovation or not, $P_{i,t-1}$ is labour productivity in the previous year, $x_{1\,it}$ is a vector of explanatory variables affecting the firm's decision of expanding in innovation, C is the threshold of innovation expenditure which equals 0, α_1 is a vector of unknown parameters that reflects the impact of explanatory variables on the decision to involve in innovation or not, γ_1 is an unknown parameter that reflects the impact of previous labour productivity on the decision whether or not to take part in innovation activities, and $u_{1\,it}$ is the error term.

The vector of explanatory variables $x_{1\,it}$ includes the firm size as number of employees, and a dummy variable for the research-intensive industry. The additional variables are: a dummy variables to express whether the firm received public subsidies from federal state, government, or the EU, a set of obstacles that may hamper innovation, a set of protection mechanisms to improve firm's competitiveness, a set of market characteristics, a set of firms cooperation partnerships with firms and research institutes for innovation, and the source of information for innovation.

The IMR_{it} variable is the IMR used to correct the selection bias, which is calculated as:

$$IMR_{it} = \frac{\phi(D^*_{it})}{\Phi(D^*_{it})} \tag{5.5}$$

where $\phi(D^*_{it})$ and $\Phi(D^*_{it})$ is the Probability Density Function (PDF) and the Cumulative Distribution Function (CDF) of the propensity of a firm to innovate consequently (Heckman, 1979).

In the second equation, the amount of expenditure on innovation will be assessed, which reflects human and financial resources dedicated to the innovation process as inputs. As seen in table A.1 and discussed in chapter 2, the empirical literature has mainly used two approaches to express the input of innovation process: R&D expenditure or innovation expenditure. In this work, innovation expenditure will be used because it includes R&D expenditures in addition to other expenditure relevant to generating innovation.

$$INE_{it} = \begin{cases} INE^*_{it} = \alpha_2^{'} x_{2\,it} + \beta_2 IMP_{it} + \gamma_2 P_{i,t-1} + u_{2\,it} & \text{if} \quad D_{it} = 1 \\ 0 & \text{if} \quad D_{it} = 0 \end{cases} \tag{5.6}$$

where INE^*_{it} is the dependent variable that expresses the intensity of a firm's expenditure on innovation as share of turnover, $P_{i,t-1}$ is labour productivity in the previous year, $x_{2\,it}$ is a vector of explanatory variables that affect the intensity of innovation expenditure, α_2, β_2, and γ_2 are vectors of unknown parameters to be estimated that reflect the impact of determinants on the intensity of innovation expenditure, and $u_{2\,it}$ is the error term. The vector of explanatory variables $x_{2\,it}$ contains the same variables of $x_{1\,it}$.

5.3.2 The Knowledge Production Stage

The second stage of the model describes the knowledge production function as proposed by Griliches (1979). It includes the expansion proposed by Peters (2006) to cover process innovation, and the expansion proposed by Polder et al. (2010) to cover organisational innovation without considering ICT as direct input due to the fact that ICT are not available in the MIP dataset. Nevertheless, the intensity of physical capital is assumed to proxy investment in ICT.

Three equations express the relationship between innovation expenditure and the potential results of the innovation process in different forms: product innovation, process innovation, and organisational innovation:

$$PRD_{it} = \beta_1 \widehat{INE}_{i,t-1} + \alpha_3^{'} x_{3\,it} + u_{3\,it} \tag{5.7}$$

$$PRC_{it} = \beta_2 \widehat{INE}_{i,t-1} + \alpha_4^{'} x_{4\,it} + u_{4\,it} \tag{5.8}$$

$$ORG_{it} = \beta_3 \widehat{INE}_{i,t-1} + \alpha_5^{'} x_{5\,it} + u_{5\,it} \tag{5.9}$$

where PRD_{it} product innovation, PRC_{it} process innovation, and ORG_{it} organisational innovation are dependent variables. $INE^*_{i,t-1}$ is innovation expenditure in the previous year. Because innovation projects are long-term projects and do not fruit directly, a time lag of one years is chosen, which also covers the effect of knowledge accumulation. Additionally, using the lag of the independent variable will reduce potential endogeneity and reverse causality. $x_{3\,it}$, $x_{4\,it}$, and $x_{5\,it}$ are vectors of explanatory variables that may affect the relationship between innovation inputs and outputs, α_3, α_4, and α_5 are vectors of unknown parameters to be estimated that reflect the impact of determinants on innovation outputs in each equation.

Determinants that may influence the firm's generation of product innovation, process innovation, and organisational innovation are $x_{3\,it}$, $x_{4\,it}$, $x_{5\,it}$ consequently. Each vector includes the firm size as number of employees, and the intensity of qualified human capital as the proportion of all employees who are educated to degree level or hold another higher education qualification. The additional variables are: a dummy variables to express whether the firm received public subsidies from federal state, government, or the EU, a set of protection mechanisms to improve firms' competitiveness, a set of market characteristics, a set of firms' cooperation partnerships with firms and research institutes for innovation, and the source of information for innovation.

5.3.3 The Production Stage

The third stage of the model depicts the link between innovation outputs and labour productivity using the Cobb-Douglas production function. The equation expresses the added value of presenting innovations to improve a firm's economic performance proxied by labour productivity:

$$P_{it} = \pi_1 \widehat{PRD}_{i,t-1} + \pi_2 \widehat{PRC}_{i,t-1} + \pi_3 \widehat{ORG}_{i,t-1} + \alpha_6^{'} x_{6\,it} + u_{6\,it} \tag{5.10}$$

where P_{it} is the labour productivity as dependent variable, $PRD_{i,t-1}$, $PRC_{i,t-1}$, and $ORG_{i,t-1}$ are sequentially the past product, process, and organisational innovation as explanatory variables. π_1, π_2, π_3 are vectors of the unknown parameters that reflect the impact of product, process, and organisational innovation on labour productivity consequently, $x_{6\,it}$ is the vector of unknown parameters to be estimated which reflect the effect of determinants on productivity.

Determinants that may influence the firm's labour productivity are given in the vector $x_{6\,it}$, which includes firm size as number of employees, and the intensity of physical capital as share of turnover. The additional variables are: a dummy variables to express whether the firm received public subsidies from federal state, government, or the EU, a set of protection mechanisms to improve firm's competitiveness, a set of market characteristics, and a set of firm's cooperation partnership with firms and research institutes for innovation.

5.3.4 Mapping Variables to Data

The dataset includes a large number of captured indicators. Table 5.3 shows the main dependent and explanatory variables used in the structural equations.

Table 5.3 Main independent and explanatory variables

Variable	Description	Format
D	Latent variable expresses the decision to innovate or not	dummy
INE	Innovation intensity (share of turnover)	censored
PRD_IMPR	Proportion of total turnover from new or significantly improved products to the firm	ordinal
PRD_NCHG	Proportion of turnover from products that were not changed or changed only slightly	ordinal
PRD_MNOV	Share of turnover from market novelties	ordinal
PRC_COST	Reduction of average costs by means of process innovations in per cent	ordinal
PRC_QUAL	Increasing turnover as result of quality improvement by process innovation in per cent	ordinal
ORG_TIME	Effect on reduction of reaction time	ordinal
ORG_QUAL	Effect on improvement of quality	ordinal
ORG_COST	Effect on reduction of average of costs	ordinal
P	Labour productivity (turnover / number of employees	censored)

5.3.5 Mapping Determinants to Data

In this section, determinants for each equation of the three-stages model will be identified and linked to the correspondent proxy from captured indicators available in the dataset. The following tables list the relevant determinants for each stage of the model and show the mapping of indicators captured in the dataset.

Table 5.5 Main determinants

Variable	Description	Format
SIZE	Firm size	censored
BRANCH	Research intensive industry	dummy
INVS	Intensity of physical capital (share of turnover)	censored
EMPL_UNI	Intensity of qualified personnel (per cent)	ordinal
PUB_SUBS	Public subsidies from federal state, government, or EU	dummy

Table 5.7 Innovation constraints

Variable	**Description**	**Format**
H_ECO_RISK	High economic risk	ordinal
H_HIG_COST	Innovation costs too high	ordinal
H_INT_FUND	Insufficient internal funding sources	ordinal
H_EXT_FUND	Insufficient external funding sources	ordinal
H_ORG_PROB	Internal organisational problems	ordinal
H_INT_RESI	Internal resistance/opposition	ordinal
H_NQA_EMPL	Insufficiently qualified employees	ordinal
H_TEC_INFO	Missing technological information	ordinal
H_MKT_INFO	Missing market information	ordinal
H_ACC_CUST	Insufficient acceptance by customers	ordinal
H_LEG_INDS	Legislation and industry standards	ordinal
H_ADM_PROC	Long administration and approval processes	ordinal

Table 5.9 Market characteristics

Variable	**Description**	**Format**
M_POS_THRE	High threat to the own market position due to entrance of new competitors	ordinal
M_CMP_UNPR	Unpredictable activities of competitors	ordinal
M_OUT_DATE	Products/services are quickly out-of-date	ordinal
M_PRO_SUBS	Products of competitors can easily substitute own products	ordinal
M_DEM_UNFS	The development of demand is unforeseeable	ordinal
M_FOR_PRES	Great pressure due to foreign competitors	ordinal
M_EXS	Presence on foreign markets	censored

A logarithmic transformation will be use for two reasons: Firstly, to handle the the non-linear relationship exists between independent and dependent variables. Secondly, to transform a highly skewed variables into approximately normal one (Benoit, 2011).

Table 5.11 Cooperation partnership for innovation

Variable	Description	Format
C_GROUP	Other firms within the same group of companies	ordinal
C_CSTMR	Customers from the private sector	ordinal
C_CSTPB	Customers from the public sector	ordinal
C_SUPLR	Suppliers	ordinal
C_COMPT	Competitors	ordinal
C_CNSLT	Consultants, consulting engineers	ordinal
C_UNIVR	Universities or higher education institutions	ordinal

Table 5.13 Source of information for innovations

Variable	Description	Format
I_GROUP	Sources inside the firm or within the group of companies or related companies	ordinal
I_CSTMR	Customers	ordinal
I_SUPLR	Suppliers	ordinal
I_COMPT	Competitors	ordinal
I_CNSLT	Consultants, consulting engineers	ordinal
I_UNIVR	Universities or higher education institutions	ordinal
I_RDINS	Research institutions	ordinal

Table 5.15 Protection mechanisms to improve competitiveness

Variable	Description	Format
P_PATNT	Patents	dummy
P_REGDS	Registered design	dummy
P_TRMKT	Trade marks	dummy
P_CPYRT	Copyright	dummy

5.4 The Estimation Strategy

5.4.1 Regression Analysis

This section presents the methods and techniques used to test the econometric model and the rationale for selecting them. As mentioned above, this research is causal research that aims to investigate the relationship between different variables.

This work employs the software tool Stata for data management and analyses. The main advantages and characteristics of this tool have been summarised in appendix D.

The target of the statistical tests in to determine the likelihood of a value in the sample, for which the null hypothesis is true. Nevertheless, the likelihood of rejecting the null hypothesis is a strength of the statistical theory. For structural equation modelling, the null hypothesis is defined by the fixed and free elements specified in the parameters of the model equation (Alavifar et al., 2012).

Regression analysis is a statistical method that predicts the values of one dependent variable using information from one or more explanatory variables. Hall (2011a) views the regression analysis as a good strategy for testing the empirical model because it allows isolation of the correlation between two variables, while holding other explanatory variables constant.

There are several techniques for conducting regression analysis, which are mostly driven by characteristics of the involved variables such as the number of independent variables, the type of the dependent variable, the nature of data, and the expected shape of the regression line.

To test the relationship between innovation inputs, innovation outputs, and productivity, and to analyse the impact of the determinants on this relationship, there is a need for more advanced econometric techniques that go beyond the traditional linear models. To test this model using the MIP data, two broad estimation techniques can be utilised: single-equations methods or system estimation methods.

5.4.2 Single-equation Estimation

In this approach, each equation in the system of simultaneous equations will be estimated separately. However, the variables in the other equations will still be considered using the estimated variable from the previous stage as an instrument in the next stage. Hence, the approach replaces the endogenous variables on the right hand side in the equations of the equation system with an instrumental variable (Washington et al., 2011).

This approach is also called the 'limited information' approach because in estimation only knowledge of the restriction in the particular equation can be provided (Kennedy, 2008). Single-equation estimation techniques are ordinary least square OLS, indirect least square Indirect Least Square (ILS), instrumental variables IV, two-stage least squares Two-stage least square (2SLS), and limited information maximum likelihood Limited Information Maximum Likelihood (LIML). Table 5.17 provides an overview of the estimation techniques used for the single equation approach.

Ordinary Least Square: OLS is a widely used statistical method for linear regression analysis to determine the relationship between variables. It is based on drawing a regression line based on using the mean to get the smallest possible sum of squared deviations scores from distribution results.

Two-stage Least Square: The 2SLS method involves two consecutive stages, as described by Wooldridge (2013):

1. Regress each endogenous variable on all exogenous regressors.
2. Use the estimated value of the endogenous variable as an instrument to the estimate equations with OLS.

Limited Information Maximum Likelihood: In this method, estimates of parameters are created by identifying the maximum probability of the observed value that would have occurred if it were the true value of the parameter (Vogt, 2005), assuming normally distributed error terms.

Table 5.17 Single-equation Estimation Methods (Washington et al., 2011)

Method	Resulting Parameter Estimates
Indirect least squares (ILS)	Consistent but not unbiased
Instrumental variables (IV)	Consistent but not unbiased
Two-stage least squares (2SLS)	Consistent but not unbiased. Generally better small sample properties than ILS or IV
Limited information maximum likelihood (LIML)	Consistent but not unbiased. Has same asymptotic variance–covariance matrix as 2SLS

5.4.3 System Estimation

In this approach, the full set of system equations defined will be estimated simultaneously with more than one dependent variable and the independent variables. It also called the 'full information' approach because estimation provides knowledge of all parameter restrictions in the equation system (Kennedy, 2008). To estimate a system of equations, a full information technique is needed, in which all the model's parameter will be estimated simultaneously (Mukherjee et al., 1998).

If the error term is correlated with the explanatory variables, classical estimators such as OLS are is not an appropriate approach to estimating a simultaneous equation model. This will result in bias and inconsistency in estimating a simultaneous system (Wooldridge, 2013). Therefore, full information system estimation methods such as seemingly unrelated regressions seemingly unrelated regressions (SUR), three stage least squares Three-stage least square (3SLS), or Full Information Maximum likelihood (FIML) are appropriate solutions. Table 5.19 provides an overview of the estimation techniques used for the system approach.

Three-stage Least Square: The 3SLS method proceeds in three consecutive stages (Mukherjee et al., 1998):

1. Estimation of the reduced form equations using OLS with only exogenous regressors.
2. Substitution of the fitted values from the first stage into the structural equations model and estimating them using OLS such as the second stage regression of 2SLS.

3. Calculation of the residuals from the second stage to obtain an estimate of the error variance covariance matrix and apply Generalized Least Square (GLS).

Full Information Maximum Likelihood: The FIML method finds the estimate of all the structural parameter rather than the endogenous variables by maximising the log-likelihood function of the model with a priori restrictions given by the structural parameters (Kennedy, 2008).

Table 5.19 System Estimation Methods (Washington et al., 2011)

Method	**Resulting Parameter Estimates**
Three-stage least squares (3SLS)	Consistent and more efficient than single-equation estimation methods
Full information maximum likelihood (FIML)	Consistent and more efficient than single-equation estimation methods. Has same asymptotic variance–covariance matrix as 3SLS

5.4.4 Justification of the Estimation Approach

As discussed above, two approaches can be used to estimate the structural equation model: the single-equation approach or system approach. Although the system approach is preferred to the single-equation approach because it considers all the parameter restrictions and error term correlation across the equation system and reduce potential over-identification (Washington et al., 2011). However, Wooldridge (2002, p.252) states that the single-equation approach is more robust against misspecification. The single-equation approach has been employed in this work due to the following advantages:

1. Solving selectivity: as mentioned above, the selectivity issue caused by investigation of only the innovative firms will be solved using a Heckman model in the first stage.

2. Availability of data: some data are not available for each sample and year. If these are estimated in a system, Stata will drop these samples, which may lead to a drastic reduction in the analysed samples and loss of statistical power. Using a single-equation enables the relevant variables for each stage to be controlled.

3. Nature of data: Most relevant variables in the data are censored, ordinal, or binary, which requires specific estimation models such as a Tobit model, ordered probit, or other alternative solutions as discussed above.

4. More robustness against potential misspecification of any equation of the system, which may result inconsistent parameters.

In the **first stage**, the first equation of the Heckman model will be implemented using a logit function because the decision of innovation D is a dummy variable. For the second equation, the Tobit function will be used because the dependent variable, which is innovation expenditure INE, is censored in the data. Tobit model is a random effects estimator.

In the **second stage**, the dependent variables, which are innovation outputs in form of product PRD, process PRD, and organisational ORG innovation, are ordinal variables. Therefore, the equation will be estimated using the ordered probit function, which is a random effect estimator.

In the **third stage**, the dependent variable, labour productivity P, is censored in the data. Therefore, this equation will be estimated using the Tobit function, which is a random effects estimator.

5.4.5 Principal Component Analysis

One additional challenge has popped up because the third stage equation incorporates eight variables (three for product innovation, two for process innovation, and three for organisational innovation), which are predictions of a previous stage with identical regressors. Consequently, these eight variables are no longer linearly independent and may suffer from multicollinearity, which makes it impossible to use them together as explanatory variables in the same equation of the regression model.

To solve this issue and reduce the dimensionality of data, PCA or Factor Analysis (FA) can be employed. The difference between them is that FA assumes the existence of a few common factors that drive the variation in the data, while PCA does not make such an assumption. The PCA approach was invented by Karl Pearson in 1901. It is based on summarising original variables with a small set of linear combinations of the covariates which are uncorrelated with each other but capture the maximum possible variation of the dataset in an iterative process (Jolliffe, 2002).

The target of PCA is to find components $z = [z_1, z_2, .., z_n]$ which are a linear combination $u = [u_1, u_2, .., u_n]'$ of the original variables $x = [x_1, x_2, .., x_n]$ that achieve maximum variance of $z = ux$, such that $u'u = 1$. The first component z_1 accounts for maximum variance, and the second component z_2 captures most of the information not captured by the first component. Each successive component explains the a maximum of the remaining variance in the data (Suryanarayana and Mistry, 2016). However, these components are uncorrelated to each other.

Therefore, PCA performs an eigenvalue decomposition of the correlation matrix, by solving the equation $(R - \lambda I)u = 0$, where R is the correlation matrix of the original variables, λ is the eigenvalue of the associated component z, and u is the eigenvector. The diagonal covariance matrix of the components is $D = diag(\lambda)$. Factors represents the correlations between the original variables x and the components z, so that $F = cor(x, z) = uD^{1/2}$. To simplify the structure of the factor matrix, components can be rotated (Katchova, 2013b).

Nevertheless, PCA creates a trade-off between the use of as few components as possible to reduce complexity and the desire to explain as much data variation as possible. The Kaiser's rule recommends keeping only components with an eigenvalue λ exceeding one, which indicates that the retained components z account for variation at least much as those available in the original variables x (Katchova, 2013b). Because these components do not contain all information in the data, there will be an unaccounted variances in the variables, which equal the sums of squares of the loadings in the dropped components, weighted by their eigenvalues.

The next step of the analysis is the factor rotation to facilitate interpretation. It transforms those factors into new factors which can more easily be interpreted by defining clusters of variables that are highly correlated with only one factor (Bryan et al., 2016). There are two approaches for factor rotation: orthogonal or oblique. Orthogonal rotation is the most used approach such as Varimax rotation which assumes that the interpretation of a factor can be measured by the variance of the squares of the factor loadings, and maximises the sum of these variances for all of the factors. Oblique rotation such as Promax rotation aims to align the factor axes as closely as possible with the clusters of the original variables.

In this study, the different types of innovations under study are highly correlated. To avoid multicollinearity in the regression of the last equation of the econometric model, it might be useful to transform the original set of variables to a set of uncorrelated principal components.

5.4.6 Estimation of Cross Sectional Pooled Data

The investigated literature assumes a nonlinear relationship between innovation and productivity; this section inspects data in pooled form. In the following figures a regularised spline fit is deployed in order to gain an initial insight into the single relationships that will be analysed in detail using the structural model approach in section 6.6.

A spline fit is a flexible method for investigating a nonlinear relationship between two variables. This widely used approach is based on expanding the independent variable by means of B-spline basis functions, which are non-linear local polynomial transformations of the independent variable. Regularisation in this case is obtained by choosing five basis functions (the local polynomials fourth order) with equidistant knots over the independent variable's domain. The fit is obtained via least squares (Hastie et al., 2009). The grey-coloured areas around the black curves reflect point-wise 95% confidence intervals obtained via Stata's prediction method.

5.4.7 Nonlinear Regression

If the relationship between a dependent variable and a predictor is non-linear, neither a single linear model nor a nonlinear model may provide an adequate description to the relationship across the entire data range. Gordon (2015) differentiate between three commonly used approaches for modelling non-linear relationships: Transforming the dependent variable Y or the predictor X using a natural logarithm, using a quadratic form of X, or using a dummy variable for X.

The approach of the spline of piecewise linear regression model aims to simplify the relationship and improve the understanding of the patterns available in the data. It allows multiple linear models to be fitted to the data for the different ranges of the explanatory variable. The values of the explanatory variable where the slope of the linear function changes are called cut-off (or join) points, which may not be known before analysis. There are two cases of piecewise regression: The first is continuous, in which the regression lines meet at a known join point. The second is discontinuous, where a gap between the regression lines is available and the join point is unknown (Von Eye and Schuster, 1998; Fomby et al., 1988).

To identify the join points a spline fit analysis will be used, which can visually show an approximate cut-off in the nonlinear relationship between the dependent variable and the independent variables without considering the other control variables. This approximation is assumed to be sufficient to define the join points used for switching regression strategy.

However, some samples may show extreme behaviour that may result pseudo cut-off points. Therefore, this analysis should consider the kernel density of the variable to reduce the effort by defining cut-off points.

Chapter 6

Analyses and Findings

6.1 Introduction

The purpose of the study is to investigate the relationship between innovation and productivity. Chapter 2 reviewed the empirical literature and addressed different points of views regarding the issues discussed. Furthermore, a set of determinants was defined that may affect the relationship between innovation and productivity. Chapter 3 presented the methodology and research design. Chapter 4 described the MIP data used in this work for unbalanced German manufacturing for the time interval between 2003 and 2013, addressed the data's characteristic and weaknesses, and defined steps for data preparation and analysis. Chapter 5 proposed an econometric model and the estimation approach applied in this work.

In this chapter, section 6.2 summarises the aim of the data analyses and states the research hypotheses. Section 6.3 shows the statistical characteristics of the model variables and the main determinants of the model. Section 6.4 presents the major results of the inferential statistics such as general observation and analysis of kernel density and correlation results. Section 6.5 investigates the relationship between innovation and productivity in the pooled cross sectional data. Finally, section 6.6 discusses the estimation results of the structural model.

6.2 Objective

This chapter has two targets: the first is to provide detailed description of each research hypothesis that will be tested using pooled and panel data, present the estimation results, and interpreting them. Additionally, determinants that are assumed to impact on the relationship between innovation and productivity will to be confirmed or rejected by analyses. The second target is to check that the results are consistent and robust across specifications with respect to the potential bias due to the selection effect, endogeneity, and other econometric issues discussed in chapter 5. As stated in section 2.6, the research hypotheses that will be tested in this work are:

- HP1: Labour productivity positively affects the firm's decision to engage in innovation.
- HP2: Labour productivity positively affects the firm's level of innovation expenditure.
- HP3: The level of innovation expenditure positively affects the generation of different types of innovations.
- HP4: Innovation positively affects a firm's labour productivity.

6.3 Descriptive Statistics

6.3.1 Model Variables

Table 6.1 shows the descriptive statistics of the main model variables available in the dataset. Additionally, table 6.2 represents the descriptive statistics of the main determinants of the model available in the dataset. Table 6.3 presents innovation constraints, table 6.4 protection measures, table 6.5 market characteristics, table 6.6 cooperation partnership, and table 6.7 the sources of information.

Table 6.1 Descriptive statistics of model variables

	count	mean	sd	min	max
D	19207	.5898891	.4918664	0	1
INE	19207	.0389312	.0708088	0	.35
INEcensored	343	.35	0	.35	.35
INEX	19207	.0178581	.1324389	0	1
P	21240	.2828535	.1634313	0	.6
Pcensored	2298	.6	0	.6	.6
PX	21240	.1081921	.3106302	0	1
PRD_IMPR	16484	2.471002	2.687998	0	8
PRD_NCHG	13073	7.097223	2.087209	0	8
PRD_MNOV	16480	.8081311	1.643057	0	8
PRC_COST	13367	.7028503	1.329286	0	8
PRC_QUAL	12703	.5958435	1.309975	0	8
ORG_TIME	2213	1.930411	1.020868	0	3
ORG_QUAL	2211	1.940299	.991849	0	3
ORG_COST	2207	1.557771	.9854757	0	3
N	21308				

Table 6.2 Descriptive statistics of model determinants

	count	mean	sd	min	max
SIZE	21315	227.0376	1626.766	0	92784.4
BRANCH	21316	.3430287	.4747321	0	1
INVS	19400	.0571678	.1038631	0	1
INVScensored	66	1	0	1	1
INVSX	19400	.0034021	.0582294	0	1
EMPL_UNI	19516	2.798627	2.12587	0	8
SUBS_PUB	7766	.3305434	.4704391	0	1
N	21316				

Table 6.3 Descriptive statistics of innovation constraints

	count	mean	sd	min	max
H_ECO_RISK	5729	.6032466	.9036297	0	3
H_HIG_COST	5729	.6591028	.9600275	0	3
H_INT_FUND	5724	.4956324	.8543084	0	3
H_EXT_FUND	5720	.4344406	.8234005	0	3
H_ORG_PROB	5714	.3319916	.6401596	0	3
H_INT_RESI	5721	.2363223	.5286366	0	3
H_NQA_EMPL	5726	.3719874	.6925261	0	3
H_TEC_INFO	5719	.2596608	.5486802	0	3
H_MKT_INFO	5720	.2984266	.6181927	0	3
H_ACC_CUST	5725	.3931878	.737099	0	3
H_LEG_INDS	5721	.3308862	.7126029	0	3
H_ADM_PROC	5719	.322434	.729318	0	3
N	5750				

Table 6.4 Descriptive statistics of protection measures

	count	mean	sd	min	max
P_PATNT	7324	.3024304	.459342	0	1
P_REGDS	7089	.1998871	.3999435	0	1
P_TRMKT	7127	.2543847	.4355454	0	1
P_CPYRT	6869	.1154462	.3195829	0	1
N	7477				

Table 6.5 Descriptive statistics of market characteristics

	count	mean	sd	min	max
M_POS_THRE	7997	1.436914	.8486978	0	3
M_CMP_UNPR	7959	1.60774	.8044051	0	3
M_OUT_DATE	7926	.9348978	.8199721	0	3
M_PRO_SUBS	7983	1.708881	.900558	0	3
M_DEM_UNFS	7960	1.705276	.8272421	0	3
M_FOR_PRES	7994	1.630723	.9230289	0	3
M_EXS	18834	.2330838	.2720711	0	.85
N	19838				

Table 6.6 Descriptive statistics of cooperation partnership

	count	mean	sd	min	max
CD_GROUP	5762	.0652551	.2469969	0	1
CA_GROUP	5751	.0528604	.2237741	0	1
CD_CSTMR	5834	.1127871	.3163595	0	1
CA_CSTMR	5788	.0488943	.2156656	0	1
CD_SUPLR	5704	.0568022	.2314847	0	1
CA_SUPLR	5686	.0205769	.1419753	0	1
CD_COMPT	5708	.0667484	.2496077	0	1
CA_COMPT	5654	.0104351	.1016269	0	1
CD_CNSLT	5974	.1606964	.3672814	0	1
CA_CNSLT	5845	.0265184	.1606847	0	1
CD_UNIVR	5796	.1152519	.3193532	0	1
CA_UNIVR	5733	.0151753	.1222605	0	1
N	6170				

Table 6.7 Descriptive statistics of source of information

	count	mean	sd	min	max
I_GROUP	5682	1.596269	1.315811	0	3
I_CSTMR	5700	1.43614	1.275779	0	3
I_SUPLR	5672	1.043724	1.04249	0	3
I_COMPT	5659	1.059021	1.049692	0	3
I_CNSLT	5650	.4709735	.7504928	0	3
I_UNIVR	5659	.6578901	.9221821	0	3
I_RDINS	5626	.4162815	.7412984	0	3
N	5765				

6.4 Inferential Statistics

6.4.1 General Observations

This section describes the distribution of productivity and innovation outputs within different groups of firms e.g. whether a firm is an innovator or non-innovator, or if it relates to a research intensive industry or not.

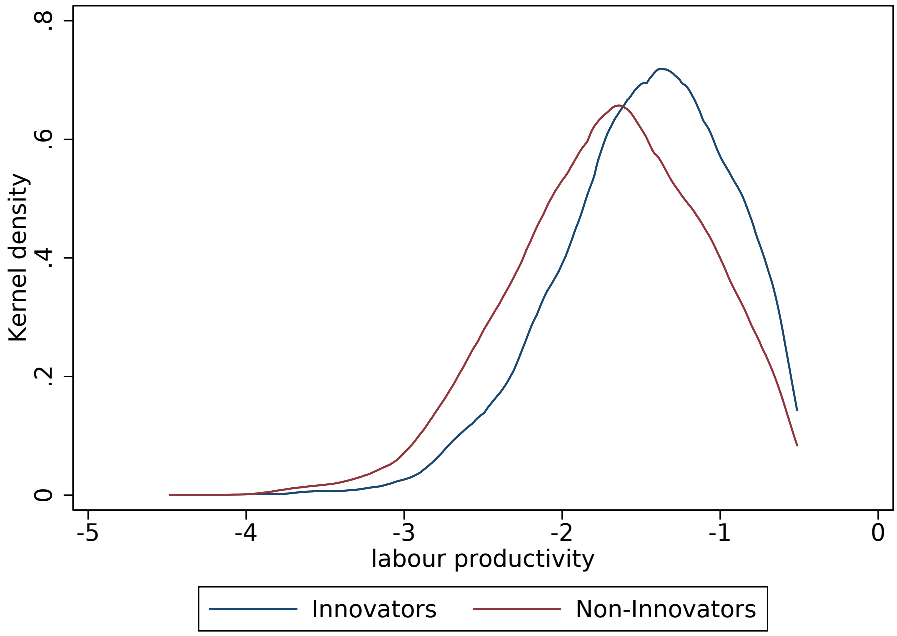


Fig. 6.1 Kernel density of labour productivity among innovators

Using the classification of available firms into the categories of innovative and non-innovative, figure 6.1 shows the distribution of labour productivity for all firms in the dataset. The blue line is associated with innovative firms whereas the red line belongs to non-innovators. Compared to the red distribution, the blue distribution possesses more probability mass at higher values for labour productivity. This implies that innovative firms in general are able to achieve higher labour productivity.

Figure 6.2 shows the kernel density estimates for labour productivity in research intensive industries (blue) and other industries (red). The blue distribution tends to have more probability mass at higher values of labour productivity. This indicates slightly higher labour productivity for companies operating in research-intensive sectors.

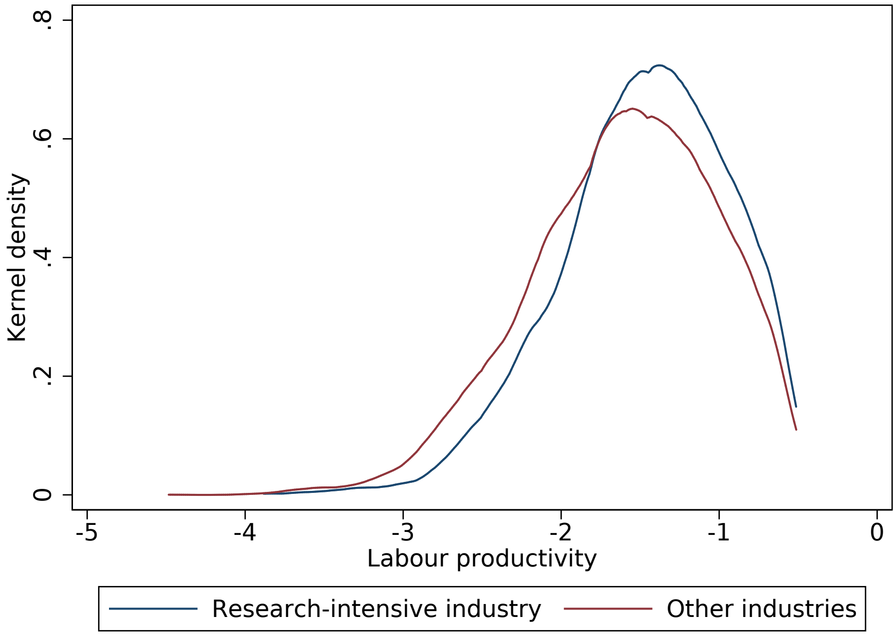


Fig. 6.2 Kernel density of labour productivity among industrial sectors

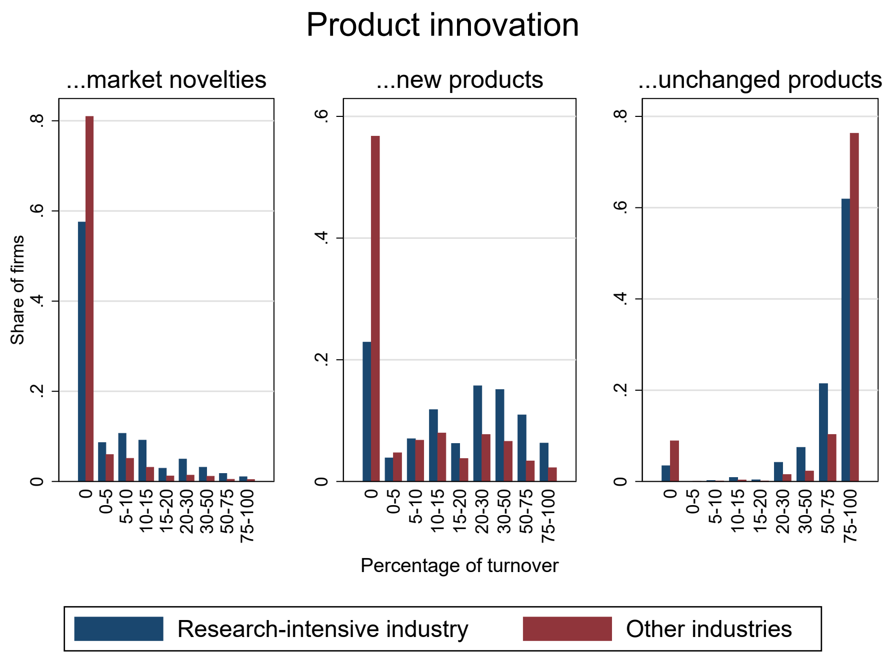


Fig. 6.3 Distribution of product innovation for industry sectors

Figure 6.3 shows the distribution of the proportion of turnover generated by three different classification of product innovations: market novelties, new or clearly improved products, and unchanged products. Research-intensive industries and other industries are considered separately. On one hand, for market novelties and new products to the firm, the bars on the very left side (no turnover generated by these products) are lower for research-intensive industry (blue) compared to other industries (red). At the same time, more firms in the research-intensive industry generate a substantial proportion of their turnover from market novelties (or new products) than firms in other industries. On the other hand, firms from other industries mainly generate their turnover from unchanged or slightly changed products. Hence, it is clear that research intensive industry generates more innovation in the form of new products and market novelties.

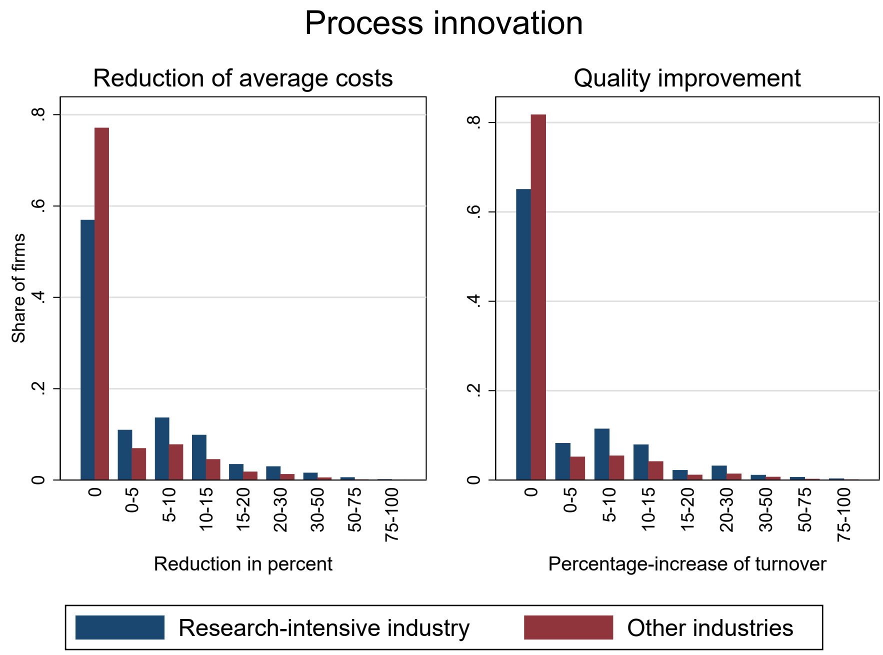


Fig. 6.4 Distribution of process innovation for industry sectors

Figure 6.4 shows the effect of process innovation on reducing average costs or increasing turnover due to quality improvement for research-intensive industries and other industries. The majority of firms in both sectors do not show an improvement in their performance as a result of process innovation. However, relatively more firms from research-intensive industries are able to reduce their average costs and improve their turnover due to quality improvement by means of process innovation. Consequently, research-intensive industries generates more process innovation in the form of cost reduction or quality improvements than other industries.

Figure 6.5 shows the effect of organisational innovation on reducing reaction time to customers, achieving higher quality, or cost reduction, for both research-intensive industries

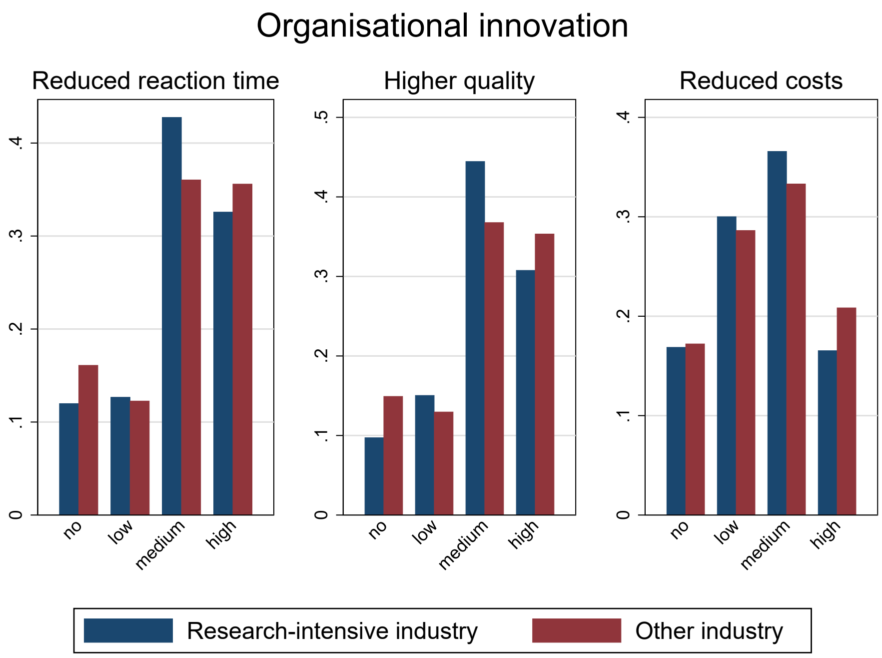


Fig. 6.5 Distribution of organisational innovation for industry sectors

and other industries. Firms of both categories seem to perform very similarly and there is no clear tendency for firms from research-intensive industry to generate more organisational innovation than those related to other industries.

6.4.2 Firm Size

As discussed in chapter 2, firm size plays a key role in the relationship between innovation and productivity. Furthermore, firm size was a core topic in the Schumpeterian innovation theory. Considering the impact of firm size on the innovation behaviour of firm enables greater understanding of many relevant aspects of this work.

Figure 6.6 shows the kernel density estimates of firm size separately for innovative (blue) and non-innovative (red) firms. The blue line is shifted to the right compared to the red line, hence probability mass for innovative firms is moved to the right which means that innovative firms tend to be larger than non-innovative ones. The wiggly behaviour of both density curves is a result of the logarithmic transformation: relatively sparse probability mass for ranges of high values is condensed far more than probability mass in ranges of small values. Additionally, when regarding firm size based on number of employees, we have to take into account that observations recorded in the dataset do not reflect real data because the

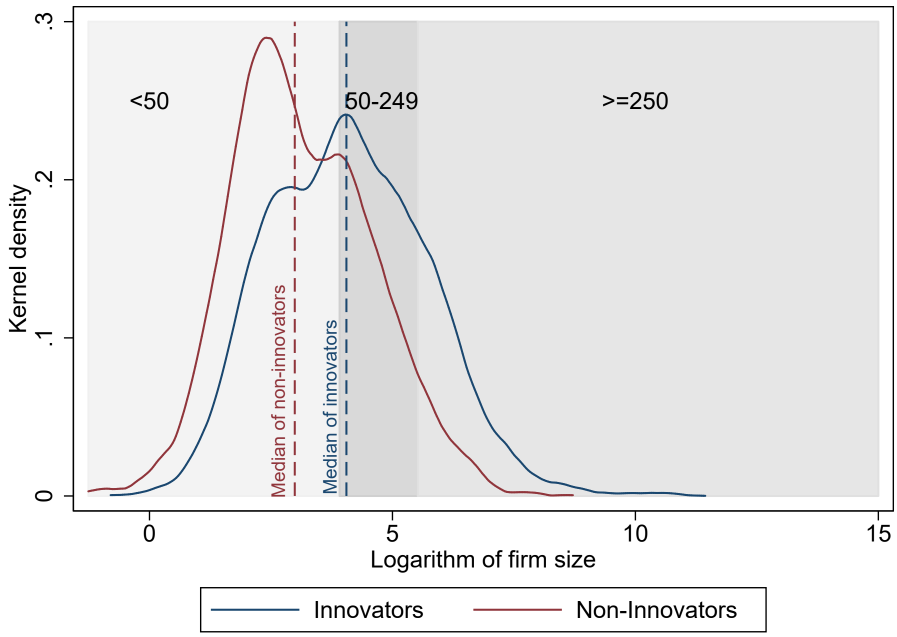


Fig. 6.6 Kernel density of firm size for innovators and non-innovators

data provider has pre-multiplied these values by Independent Identically Distributed (IID) random numbers for anonymity reasons. However, in this case it is not such an issue since aggregated data is analysed where errors cancel themselves out.

Figure 6.7 shows the proportion of firms using public subsides in terms of their size, divided into three different categories. Larger firms have easier access to public subsides since more than half of the large firms use public subsides whereas among small firms less than a quarter have access to public subsides.

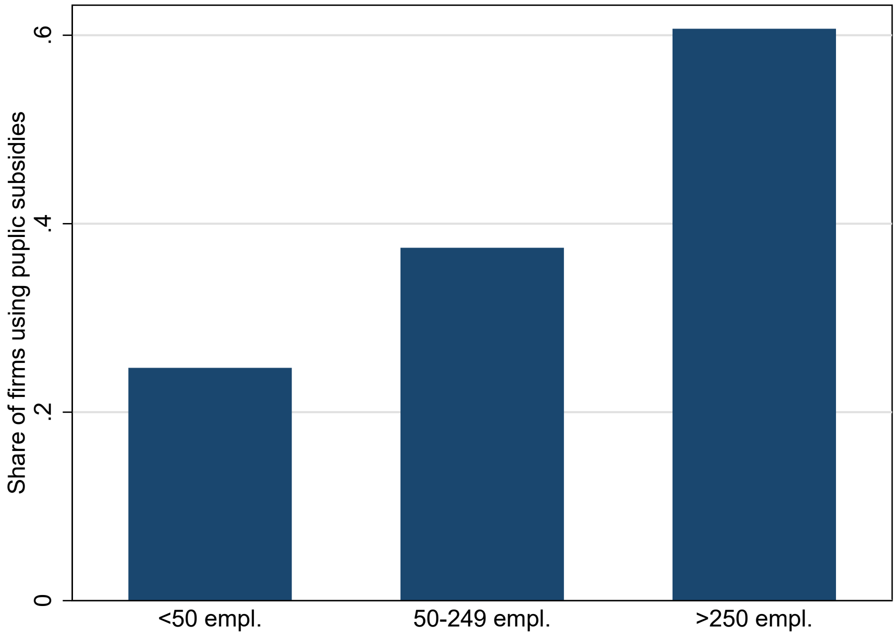


Fig. 6.7 Proportion of firms receiving subsides among firm size categories

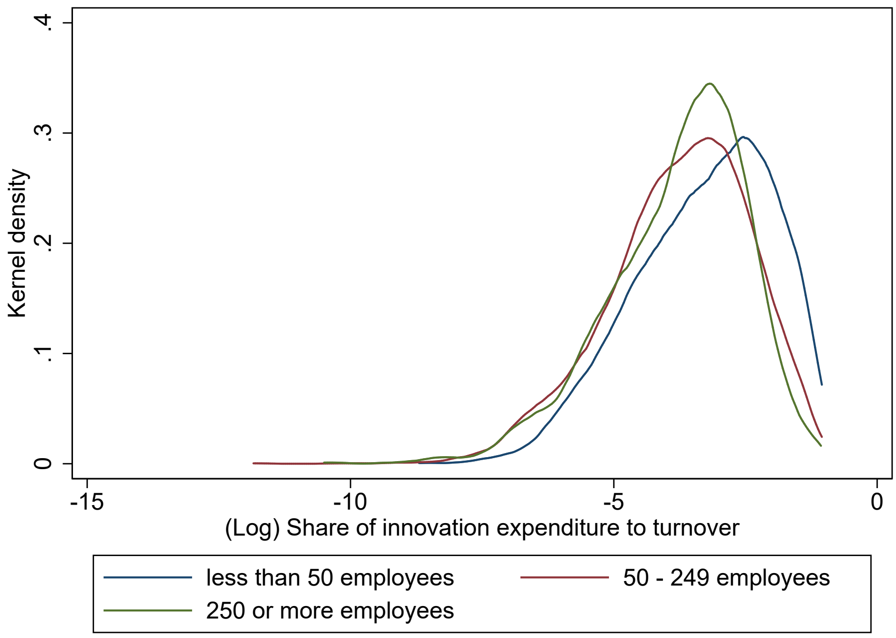


Fig. 6.8 Kernel density estimates for share of innovation expenditure relative to turnover

As shown in figure 6.8, the estimated distribution of the amount that is spent on innovation activities (relative to the firm's turnover) does not differ substantially over different size classes.

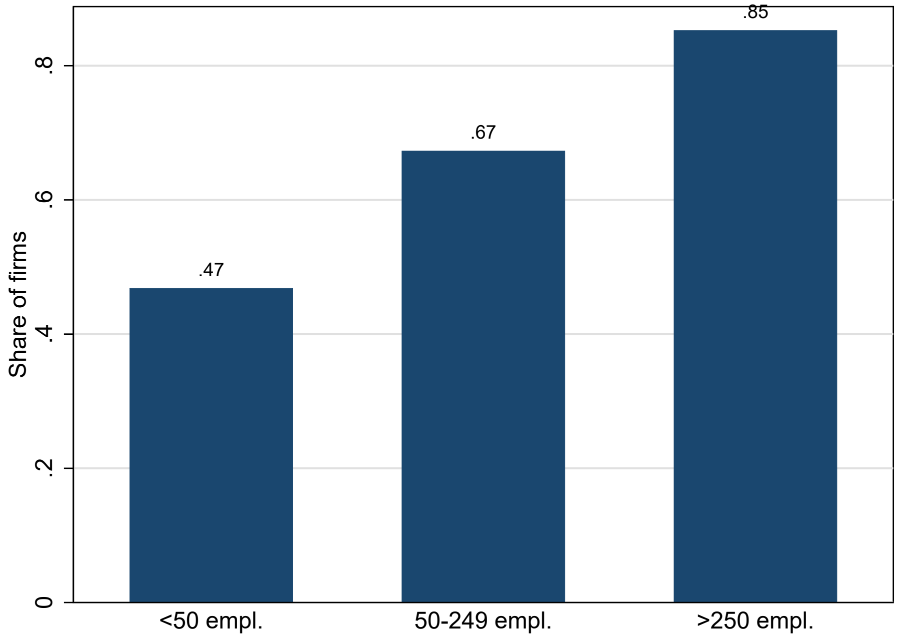


Fig. 6.9 Proportion of firms with positive innovation expenditures

Figure 6.9 shows the proportion of firm expenditure on innovation activities categorised according to their size. Large firms spend the highest proportion and small firms spend the smallest proportion. Almost half of the small firms decide to spend money on innovation activities, whereas two thirds of medium sized firms and almost 90% of large firms decide to do so.

At first glance, the firm size seems to have a positive effect on labour productivity. Figure 6.10 shows the kernel density estimates of labour productivity among three categories of firm size. The blue line (small firms with less than 50 employees) is at the very left and the density has a different shape (skewed, higher peak) compared to the remaining ones which seem to be quite symmetrical. The red line (medium-size firms with between 50 and 249 employees) is located between the blue and the green line (large firms with 250 employees or more). It can be concluded that scale effects also play a role when it comes to explaining labour productivity. The larger a firm is, the higher its labour productivity.

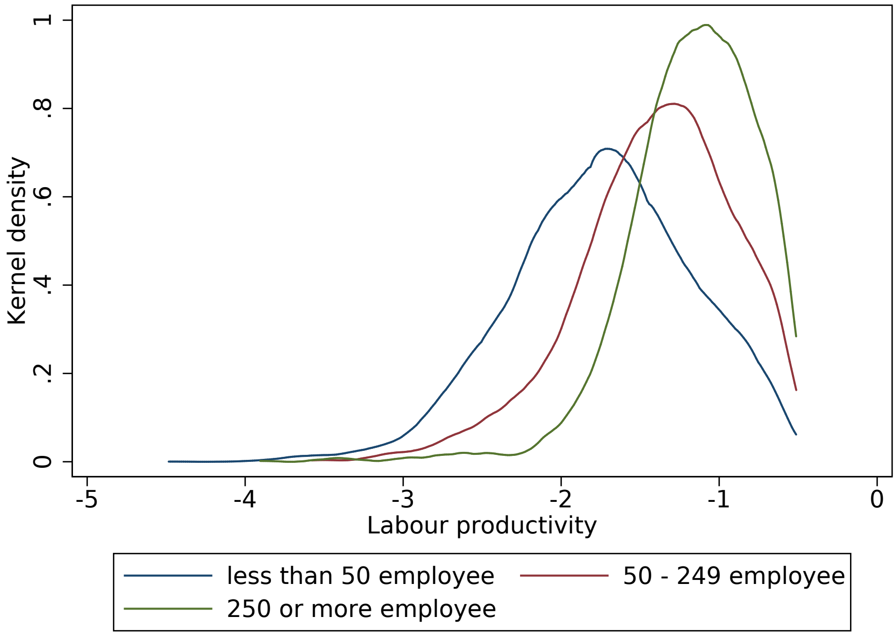


Fig. 6.10 Distribution of labour productivity among firm size categories

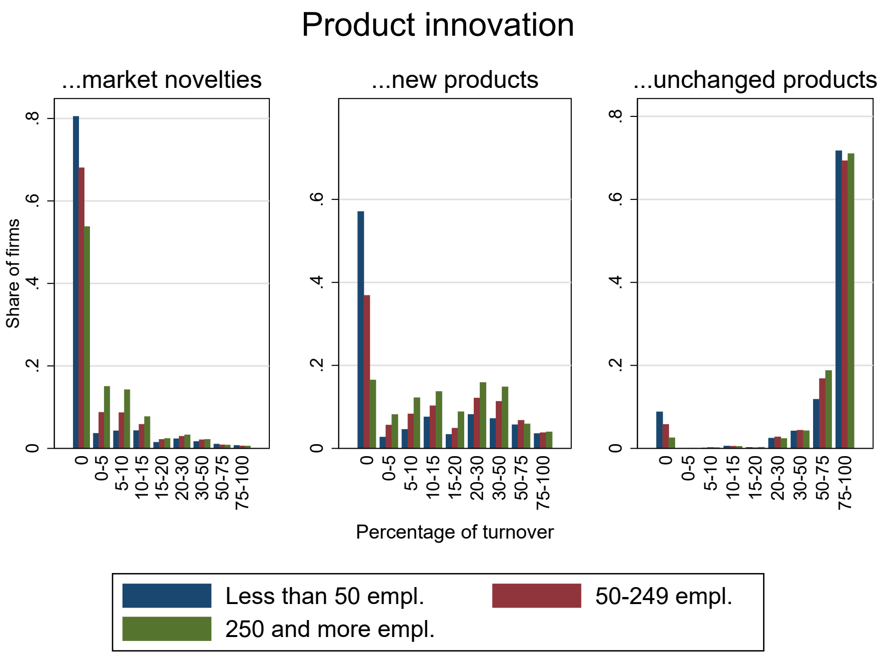


Fig. 6.11 Distribution of product innovation among firm size categories

Figure 6.11 shows the distribution of the proportion of turnover generated by three different classifications of products: market novelties, newly introduced to the market as clearly improved products, and unchanged products. In this case the distinction is made with respect to the three categories of firm size (less than 50 employees, 50 to 249 employees, and 250 or more employees). Apparently, the majority of firms generate very few market novelties. However, it seems to be more likely that larger firms generate a substantial part of their turnover from market novelties. Qualitatively, the same image arises in the case of new products, where the difference lies only in the magnitude: the role of new products is higher for all firm sizes and again there is a clear tendency for larger firms to be more likely to generate a higher proportion of their turnover from new products. The right panel (turnover generated from unchanged products) seems to contradict these findings since the difference between firm sizes found in the previous panels does not appear here. Taking into account that the classes for the proportion of turnover are not uniformly sized since the classes for higher proportions are larger and as such, more information is more aggregated. Consequently, larger firms are relatively more innovative because of the fact that they produce relatively higher turnover from new products or market novelties.

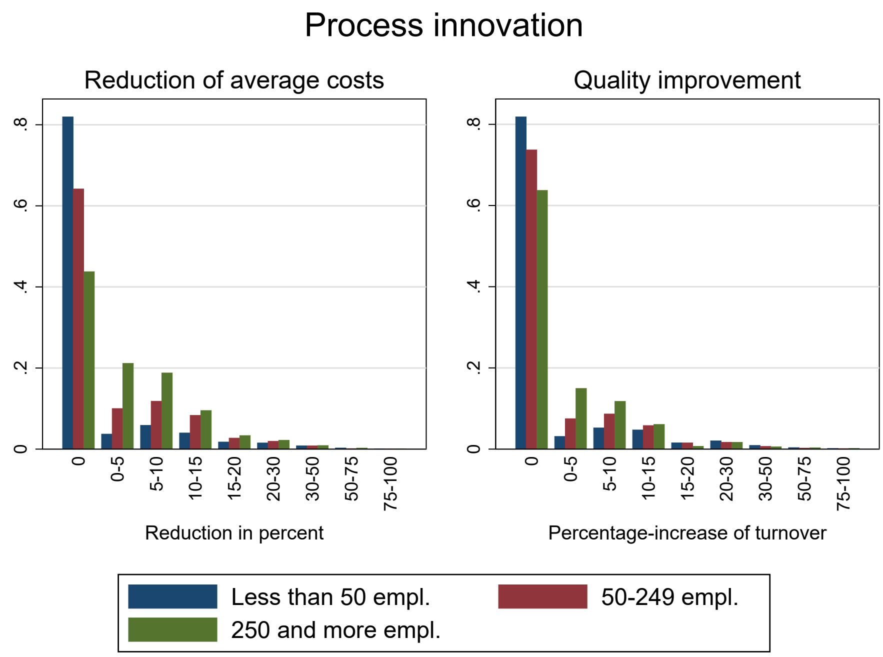


Fig. 6.12 Distribution of process innovation among firm size categories

For process innovation, it also seems that larger firms can exploit scale effects as is revealed by figure 6.12. The left panel depicts how cost reduction is distributed across firms whereas the right panel illustrates the distribution of turnover increases as a result of quality improvements. The vast majority (more than 80%) of small firms (less than 50 employees) are neither able to reduce average cost nor able to increase turnover by means of quality improvements. Medium-sized firms show advances in both cases and these findings are even more relevant for large firms. Hence, larger firms seem to find it much easier to generate process innovation in terms of cost reduction and increase in turnover caused by quality improvement.

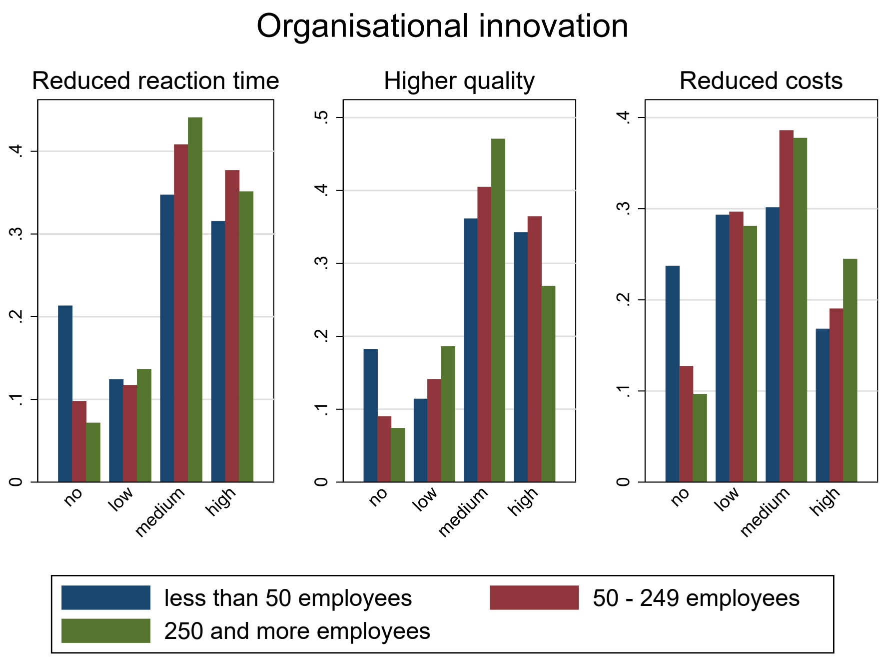


Fig. 6.13 Distribution of organisational innovation among firm size categories

Figure 6.13 shows the distribution of generation of organisational innovation in relation to firm size. Organisational innovation is measured by three different aspects: reducing reaction time, improving quality, and reducing cost. A clear tendency in the relationship is not present. Considering the values ’no’ and ’medium’, a tendency towards more organisational innovation among large firms is apparent (among mid-sized firms and large firms a smaller proportion has this characteristic compared to small firms) but this tendency vanishes for other categories of firm size. Hence, large firms produce more organisational innovation in the form of reduced reaction time, higher quality and lower costs but this conclusion is not that strong.

6.4.3 Innovation Constraints

As shown in figure 6.14, the most frequent factors hampering innovation for all firm size categories are related to financial issues, such as high costs of innovation, the high economic risk, and the insufficient funding of innovation activities.

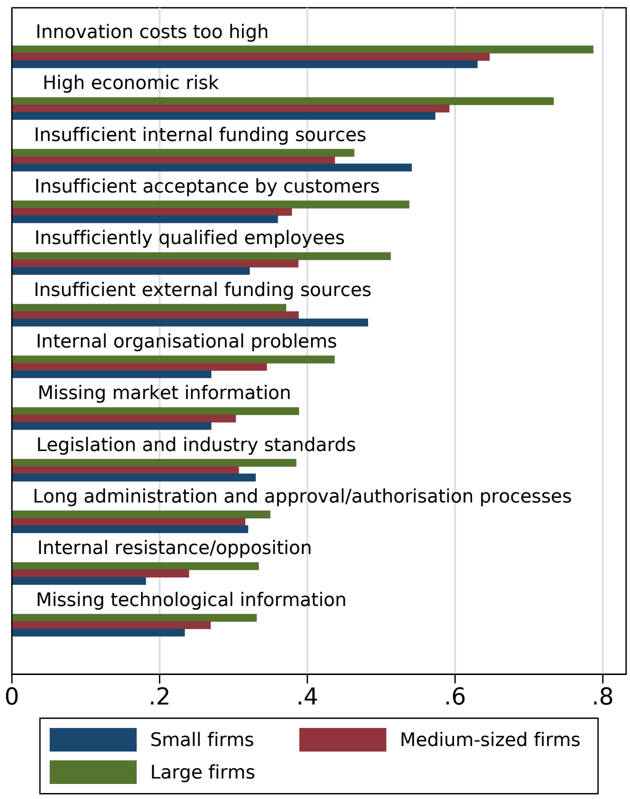


Fig. 6.14 Innovation constraints for each firm size category

6.4.4 Firms Innovation Behaviour

Fosse et al. (2013) argue that if a firm is engaged in one type of innovation, it is not possible to fully separate the effects of the individual innovation from each other.

Table 6.8 shows the correlation behaviour between the different innovation types, which are independent variables in the structural model. Taking a closer look at pairwise correlations between different types of innovation output shows that virtually all variable pairs exhibit a strongly significant correlation. However, the magnitude of dependency varies substantially for different variable pairs. This magnitude is very small for all pairs including PRD_NCHG (share of turnover generated by unchanged or slightly changed products). Furthermore, it seems that the sign of the relationship is not the same in all cases. A relatively high correlation can be observed between variable pairs of product innovation and process innovation. Relatively high correlation is also observed within the group of variables for organisational innovation. There is a correlation between organisational innovation on the one hand and product or process innovation on the other. Hence, firms that produce new

Table 6.8 Correlation between different innovation types

	PRD_IMPR	PRD_NCHG	PRD_MNOV	PRC_COST	PRC_QUAL	ORG_TIME	ORG_QUAL	ORG_COST
PRD_IMPR	1							
PRD_NCHG	-0.0943***	1						
PRD_MNOV	0.568***	-0.0965***	1					
PRC_COST	0.587***	0.0344**	0.416***	1				
PRC_QUAL	0.594***	-0.00702	0.419***	0.458***	1			
ORG_TIME	0.239***	-0.0780**	0.121***	0.243***	0.211***	1		
ORG_QUAL	0.277***	-0.110***	0.196***	0.185***	0.289***	0.599***	1	
ORG_COST	0.156***	-0.0116	0.0836***	0.316***	0.162***	0.494***	0.518***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

or improved products also tend to improve their processes. Organisational innovation, however, seems to be more independent because firms that invent new products or improve their processes do not necessarily tend to carry out organisational innovation.

Table 6.9 Persistence of innovation activities

	Innovation expenditure (share of turnover)
Lagged innovation expenditure (year-1)	0.635***
Lagged innovation expenditure (year-2)	0.573***
Lagged innovation expenditure (year-3)	0.550***
Lagged innovation expenditure (year-4)	0.514***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.9 shows that firms with high innovation expenditure tend to have high spending on innovation activities in the subsequent period as well.

Table 6.10 Persistence of firm productivity

	Turnover / No. of employees (labour productivity)
Lagged labour productivity (year-1)	0.868***
Lagged labour productivity (year-2)	0.827***
Lagged labour productivity (year-3)	0.795***
Lagged labour productivity (year-4)	0.788***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.10 shows that regressing the current value of labour productivity with those from the past show a strong persistence of productivity behaviour from firms.

6.5 Empirical Results of Pooled Data Model

This section presents the pre-estimation results using pooled cross sectional data for the three stages of the structural econometric model.

6.5.1 The Decision and Expenditure Stage

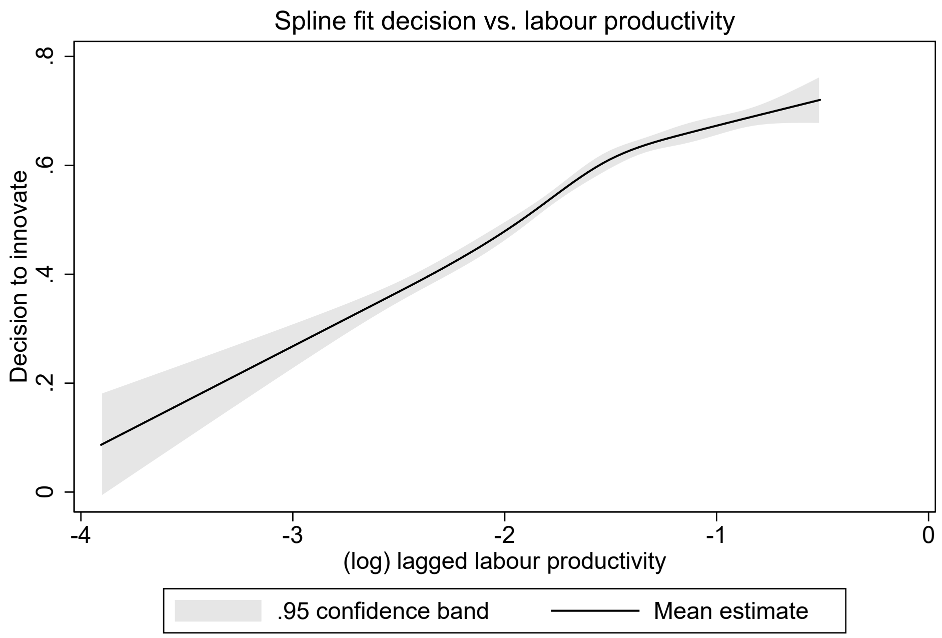


Fig. 6.15 Labour productivity and innovation decisions (pooled)

Figure 6.15 depicts the local estimate of the proportion of the firm's decision to innovative dependent on the lagged labour productivity. The relationship in the pooled sample shows that higher lagged labour productivity encourages firms to take the decision to engage in innovation activities with quite a narrow confidence band. However, the relationship flattens with lagged labour productivity higher than 20% (of turnover / number of employees).

Figure B.1 shows that the innovation decision in relation to the logarithm of labour productivity seems to be linear. The second rise in the density function is caused by the censored samples.

Similarly, figure 6.16 shows a clear positive relationship between lagged labour productivity and innovation expenditure with quite a narrow confidence band. This indicates

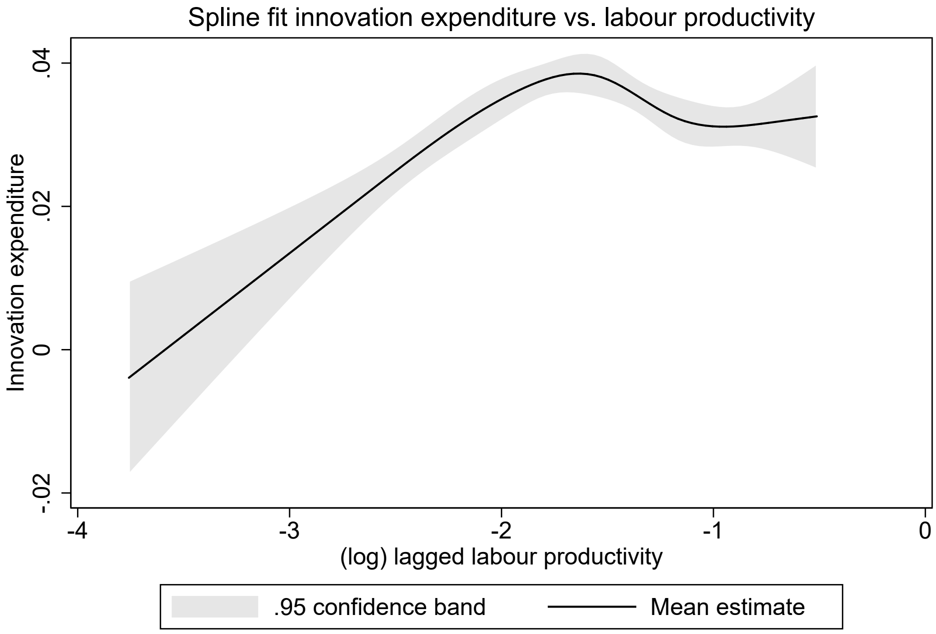


Fig. 6.16 Labour productivity and innovation expenditure (pooled)

that firms with higher lagged labour productivity encourage expenditure on innovation up to a specific level. For firms with lagged labour productivity levels higher than 20%, the relationship turns out to be negative. However, this relationship has no causal interpretation but reflects a relationship that can be observed in the pooled sample and shows a more correlation-like dependency. The mechanism behind this has to be addressed in a framework allowing identification of causal effects.

Figure B.2 shows that the innovation expenditure in relation to the logarithm of labour productivity seems to be linear at the range where the kernel density function indicates a relatively high population concentration. The second rise in the density function is caused by the censored samples.

Table 6.11 shows the regression results for the Heckman model based on a pooled data model in two columns. The left column contains the Odds Ratio (OR) of the participation equation estimated with the Logit model and the corresponding statistics. The right column contains the coefficients of the equation that expresses the expenditure on innovation activities.

Table 6.11 Estimation results of 1st stage (pooled)

	(1)		(2)	
	D		l_INE	
	OR	z	coef.	t
L.l_P	1.295***	(6.85)	-0.232***	(-9.70)
l_SIZE	1.509***	(26.72)	-0.0521***	(-3.68)
BRANCH	3.947***	(28.60)	0.707***	(15.00)
IMR			0.221**	(3.02)
N	11280		6486	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In such a cross-sectional setup, it is not possible to control for individual heterogeneity but the analysis provides the first evidence for the relationships of interest. In the left column of the table, the OR values and corresponding t-statistics of the participation equation are obtained from a Logit model, in which the binary dependent variable is the innovation decision. For all included covariates OR is higher than one and therefore is positively related to the probability of a positive innovation decision and statistically significant at the 0.05% level. These results were then used to predict values for the latent propensity of the decision to innovate, which are then used to compute the IMR which accounts for potential selection bias.

The right column of table 6.11 shows estimates and corresponding t-statistics obtained from a Tobit regression of the equation of innovation expenditure, related to different firm characteristics. All coefficients are highly significant at the 0.05% level but lagged labour productivity seems to be negatively related to the level of innovation expenditure. Hence, this coefficient provides evidence of a relative decrease in innovation expenditure for larger firms. One possible interpretation is that economies of scale are also relevant in the production of innovation. Research-intensive industry seems to be relevant for the extent of innovation expenditure and the significant coefficient of the IMR indicates that a selection bias is clearly relevant for this equation.

6.5.2 The Knowledge Production Stage

Figure 6.17 depicts a regularised spline fit of the scatter plot for product innovation against the logarithm of lagged expenditure on innovation activities. The results propose that a higher amount of expenditure on innovation activities is associated also with an increase of innovation output in case of product improvement and market novelties. Since a log transformation was used for innovation expenditure in this figure, the relationship between the

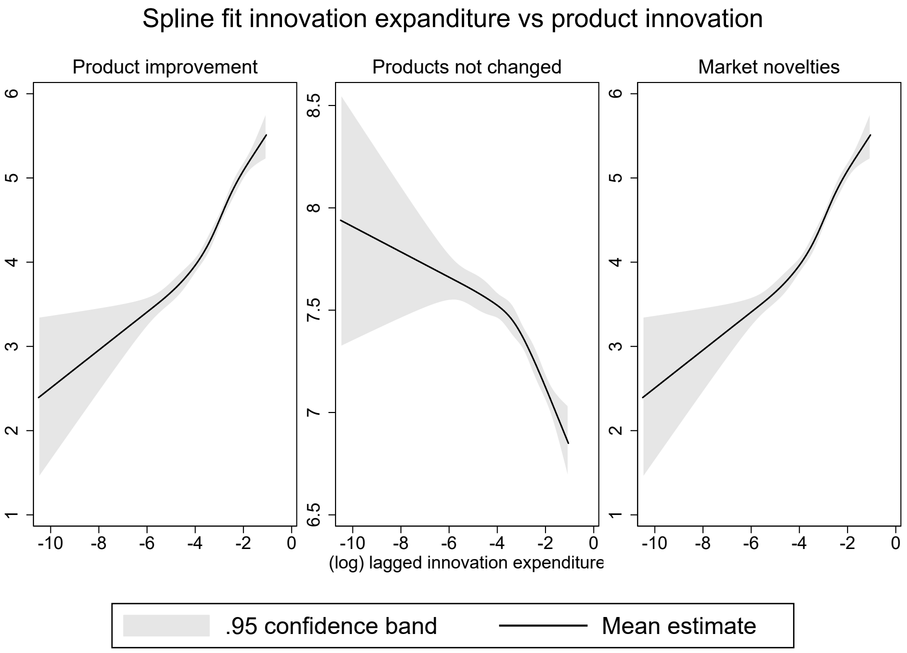


Fig. 6.17 Innovation expenditure and product innovation (pooled)

two non-transformed variables is apparently a non-linear one and the effect of an increase in innovation expenditures seems to become less important the higher their value. However, logarithms in this case seems to provide an appropriate transformation for the relationship in order to be approximated by a linear parametrisation. This relationship cannot be interpreted as causal because it simply reflects the pattern observed for the pooled sample. A similar pattern can also be observed for the cases of process innovation and organisational innovation. However, they differ in the way that, for instance, quality improvement seems to be more reactive to changes in innovation expenditure.

Table 6.12 Estimation results of 2nd stage product innovation (pooled)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.295^{***}	(3.65)	-0.0885	(-0.85)	-0.0307	(-0.34)
l_SIZE	-0.0429	(-1.68)	0.0878^{**}	(2.76)	-0.0412	(-1.49)
EMPL_UNI	0.0774^{***}	(7.21)	-0.0791^{***}	(-5.91)	0.102^{***}	(8.69)
IMR	-0.298^{***}	(-3.59)	0.380^{***}	(3.60)	-0.380^{***}	(-4.06)
N	2590		2279		2564	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results of the second stage are shown in tables 6.12 for product innovation, table 6.13 for process innovation, and table 6.14 for organisational innovation. Different dependent variables are relevant for the second stage (8 in total) and the name of each variable is indicated by the column name. All these variables are only measured on a ordinal scale, hence, an ordered probit model is used to estimate the structural coefficients.

As expected, there is a significant positive relationship between the share of turnover generated by improved products and the level of innovation expenditure. The coefficients corresponding to the other types of product innovation appear statistically insignificant. However, the negative sign indicates that a potential nonlinear relationship might be relevant in the sense that the increase in innovation outcome (in terms of product innovation) is higher when the level of innovation expenditures is below the given threshold.

Another finding is that firms with a higher share of turnover generated by improved products are also associated with a higher number of employees with a university degree. Very similar results are obtained for market novelties. Turnover resulting from unchanged products constitute the remaining part and one could expect coefficients with the opposite sign compared to the previous estimates. However, only the share of employees with a university degree seems to meet this expectation. All other covariates do not exhibit significant coefficients.

A similar pattern can also be observed for the cases of process innovation and organisational innovation. However, they differ in the way that, for instance, quality improvement seems to be more reactive to changes in innovation expenditure. Figure 6.18 depicts a regularised spline fit of the scatter plot for process innovation against the lagged expenditure on innovation activities. The relationship between innovation expenditure and cost reduction (left panel) seems to be a positive and approximately linear. Such a linear relationship, however, can not be reported for the relationship between innovation expenditures and quality improvements. Here it seems, that expanding innovation expenditure when it is already at a comparably high level leads to a by far higher innovation output driven by quality improvements.

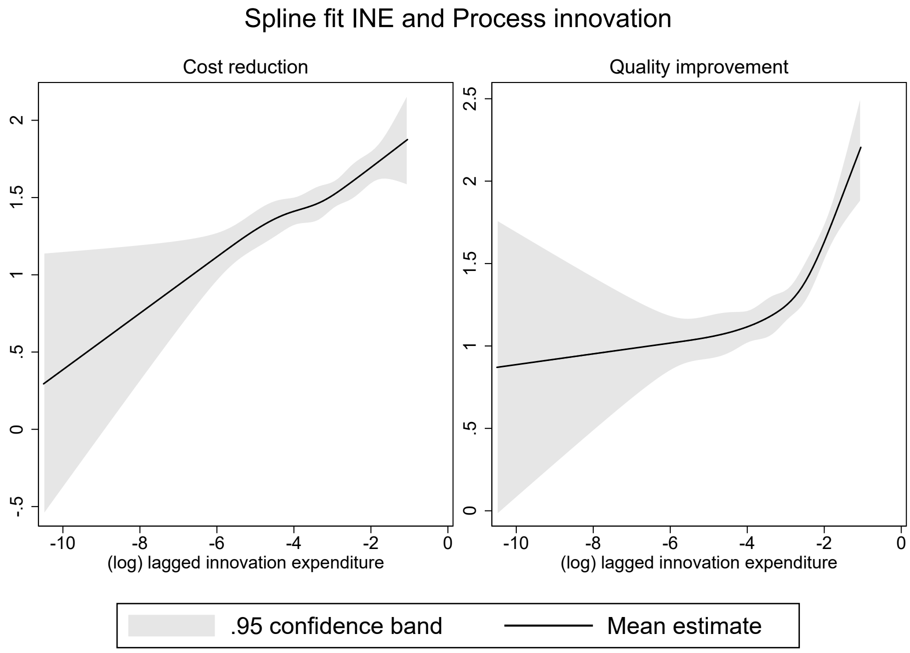


Fig. 6.18 Innovation expenditure and process innovation (pooled)

Table 6.13 Estimation results of 2nd stage process innovation (pooled)

	(1)		(2)	
	PRC_COST		PRC_QUAL	
L.l_INEp	0.0129	(0.12)	0.275*	(2.38)
l_SIZE	-0.0415	(-1.17)	-0.0354	(-0.91)
EMPL_UNI	0.00396	(0.27)	0.0272	(1.75)
IMR	-0.304**	(-2.63)	0.223	(1.82)
N	1667		1514	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.13 shows the results of pooled estimates for process innovation. Investigated variables seem to be insignificant for presenting process innovation.

The relationship between innovation expenditure and organisational innovation seems to turn into a negative one when innovation expenditure reaches a level of 0.02. Figure 6.19 depicts a regularised spline fit of the scatter plot for organisational innovation against the lagged expenditure on innovation activities. In all three panels one can observe a certain peak at approximately 0.02. However, regarding the corresponding confidence bands, one may conclude that these peaks might not be that strong from a statistical point of view. Again note that in general, such a relationship does not permit a causal interpretation because it sim-

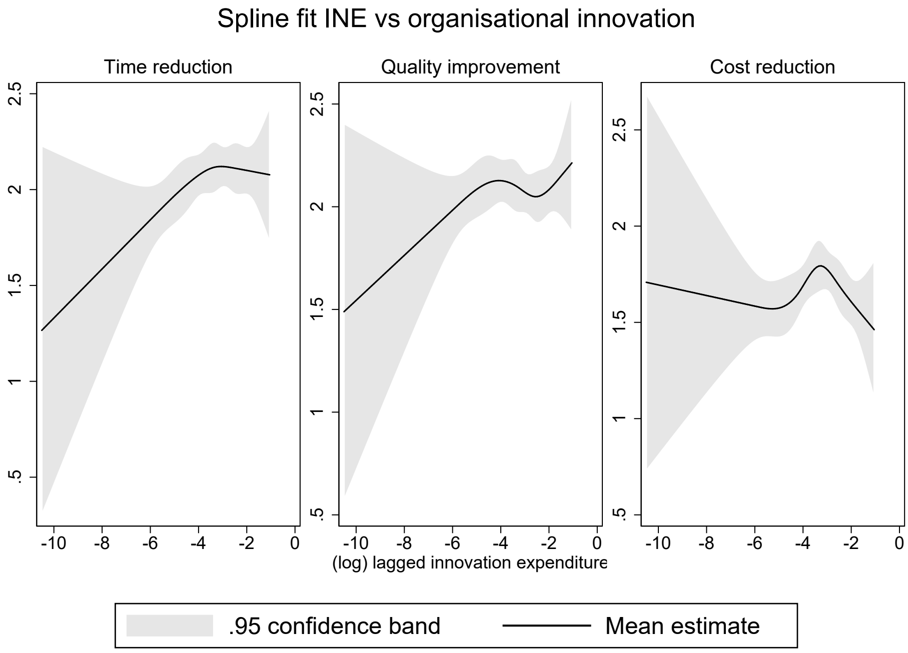


Fig. 6.19 Innovation expenditure and organisational innovation (pooled)

ply reflects the pattern observed for the pooled sample.

Table 6.14 Estimation results of 2nd stage organisational innovation (pooled)

	(1)		(2)		(3)	
	ORG_TIME		ORG_QUAL		ORG_COST	
L.l_INEp	-0.159	(-0.74)	-0.343	(-1.62)	-0.00395	(-0.02)
l_SIZE	0.0711	(1.04)	-0.00865	(-0.13)	0.149*	(2.25)
EMPL_UNI	0.0231	(0.93)	0.0217	(0.87)	-0.00218	(-0.09)
IMR	0.0862	(0.39)	0.137	(0.63)	0.432*	(2.02)
N	511		512		509	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.14 show the regression results of the organisational innovation of the second stage using pooled cross-sectional data. No significant coefficients can be reported, which may be the result of a smaller number of observations by far.

As seen in figure B.3, figure B.4, and figure B.5 for product innovation, figure B.6, and figure B.7 for process innovation, figure B.8, figure B.9, and figure B.10 for organizational innovation, functions to the predicted innovation expenditure show linear behaviour at the range where the kernel density function indicates a relatively high population concentration.

6.5.3 The Production Stage

The local mean of labour productivity via a spline fit conditional on the different measures of product innovation is depicted in figure 6.20. As above, the left panel refers to improved products and the local mean suggests that starting with the improvement of goods comes at the cost of being less productive in the production factor labour. This negative relationship turns into a positive one above a certain threshold of 10-15% and is negative again when the share of turnover generated by improved products is already significantly higher than 30%. However, the overall gradient seems to be negative. Market novelties seem to reduce labour productivity only above the level of 30% of the total turnover. The spline fit of corresponding to turnover resulting from unchanged products supports this pattern, the relationship is moderate positive for a turnover below 30%, and then the slope rises to reach its peak at a turnover between 50-70%.

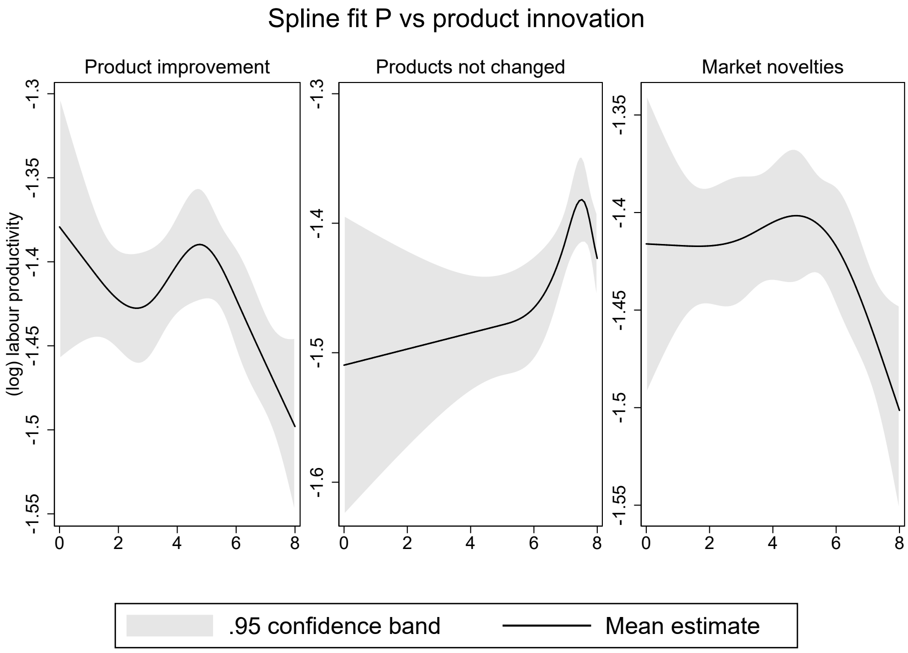


Fig. 6.20 Product innovation and labour productivity (pooled)

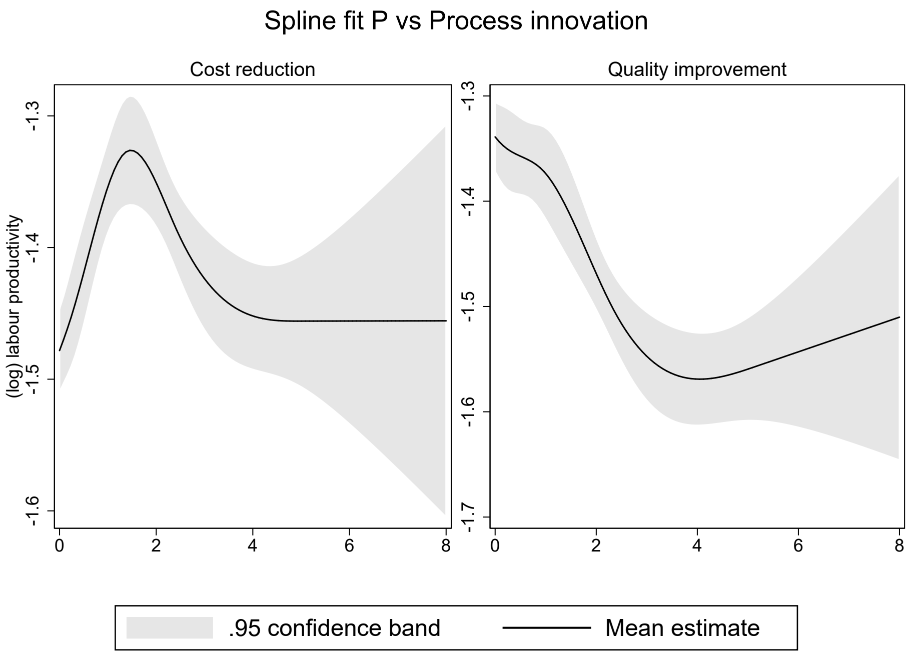


Fig. 6.21 Process innovation and labour productivity (pooled)

The local conditional mean of labour productivity conditional on process innovation is illustrated in figure 6.21. Reducing costs up to a certain specific number of percentage points is associated with increased labour productivity. Reducing costs (left panel) to a greater extent does not seem to be related to higher labour productivity. Improving quality (right panel) through process innovation even seems to lead to lower labour productivity.

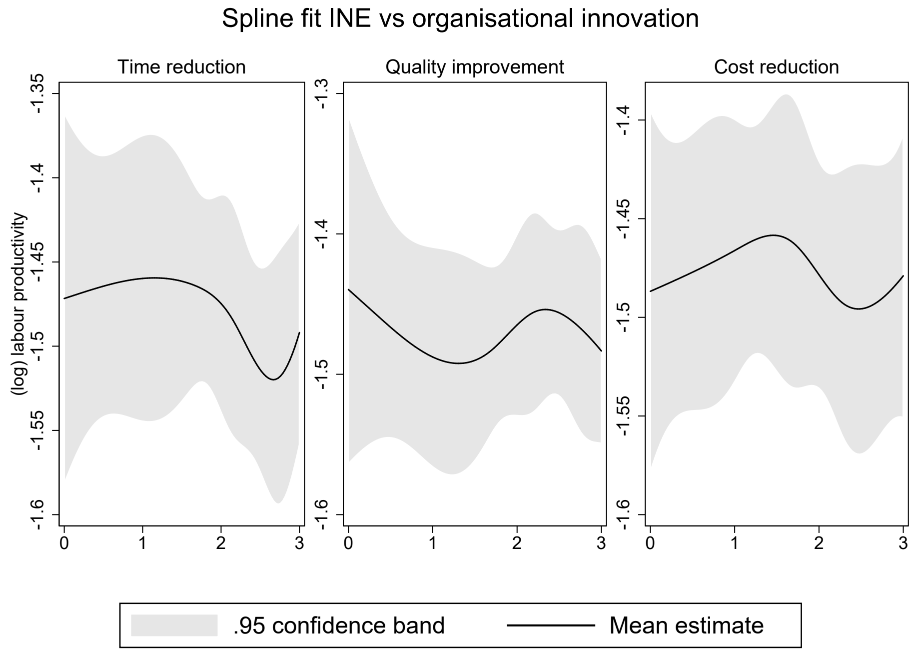


Fig. 6.22 Organisational innovation and labour productivity (pooled)

Figure 6.22 shows that no clear tendency can be identified in the relationship between organisational innovation and labour productivity in pooled data. In addition, the relatively wide confidence bands (point-wise) indicate that the structure of the fitted curves does not contain any significant information.

Table B.27 presents the correlation matrix between the explanatory variables of the third stage. Most of these variables are highly correlated to each other, which indicates a high potential to be composed of less components.

Table B.28 shows that the eigenvalues of the first three components have a value higher than one. The row *Proportion* shows the variation explained by each component and the row *Cumulative* shows the cumulative explanation of data by these components. The first three components explain about 92% of the data variation, which is the threshold chosen. Based on the PCA presented in table B.28, the first three components are considered as representative of the eight original variables from different types of innovations.

Table 6.15 PCA results after choosing the significant components (pooled)

	Comp1	Comp2	Comp3	Unexplained
PRD_IMPRp	0.4563	0.0243	0.2239	0.02715
PRD_NCHGp	-0.4521	-0.0700	-0.2953	8.916e-06
PRD_MNOVp	0.4014	0.2985	0.2854	0.001165
PRC_COSTp	0.4151	0.2184	-0.3455	0.01323
PRC_QUALp	0.0382	-0.5623	0.5049	0.02208
ORG_TIMEp	-0.0766	0.6370	0.1327	0.01143
ORG_QUALp	-0.3059	0.2328	0.6204	0.1041
ORG_COSTp	-0.3911	0.2856	0.0668	0.1257
Variance	3.65664	2.0789	1.95963	-
Difference	1.57775	0.1193	.	-
Proportion	0.4571	0.2599	0.2450	-
Cumulative	0.4571	0.7169	0.9619	-

As shown in table 6.15, for the first component (accounting for more than 45% of the variation), improved products and market novelties (product innovation) and cost reduction (process innovation) are loaded positively, but unchanged product (product innovation), quality improvement (organisational innovation) are loaded negatively. In the second component, market novelties (product innovation), cost reduction (process innovation), and quality improvement (organisational innovation) are loaded positively, whereas quality improvements (process innovation) are loaded negatively. Quality improvements (from both process and organisational innovation) are loaded into the third component. Because these three components do not contain all information in the data, the column 'Unexplained' includes the unaccounted variances in the variables, which equal the sums of squares of the loadings in the dropped components, weighted by their eigenvalues.

Table 6.16 Estimation results of 3rd stage using PCA components (pooled)

	(1) l_P	
L.pc1	0.0122	(0.88)
L.pc2	0.711***	(25.58)
L.pc3	-0.260***	(-10.72)
l_SIZE	-0.549***	(-23.83)
l_INVS	-0.0544***	(-4.73)
IMR	-0.450***	(-5.14)
N	2094	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The estimation results of the pooled model's third stage, where the functional chain is closed by explaining variation of labour productivity are given in table 6.16. The covariates in this Tobit model relate to innovation outcome proxied by the components resulting from PCA as explanatory variables. The transformation of the regression coefficients resulting from loadings into the components and the rotation is given in table 6.17.

Table 6.17 Transformed estimation coefficients after components regression (pooled)

	PCR Coefficients
PRD_IMPRp	-0.0353
PRD_NCHGp	0.0214
PRD_MNOVp	0.1431
PRC_COSTp	0.2501
PRC_QUALp	-0.5306
ORG_TIMEp	0.4177
ORG_QUALp	0.0007
ORG_COSTp	0.1810

The results of the Tobit regression indicate a highly significant relationship between labour productivity and second component (positive) and third component (negative), whereas the first component is not significantly related. The re-transformed coefficients indicate the positive impact of organisational innovation in general, the turnover resulting from market novelties by product innovation, the reduction of average costs by process innovation, and the organisational innovation targeting at reduction of reaction time. However, the re-transformed coefficients indicate a negative relationship for increasing turnover as result of quality improvement by process innovation. The remaining coefficients have a relatively low magnitude. Beyond that, both firm size and intensity of physical capital are negatively related to labour productivity. Additionally, the IMR appears to be significantly related. One could conclude from this, that selectivity issues are also relevant at this stage of the model.

Finally, table B.26 shows the estimation results of the first stage for dummy variables which present different combination of product innovation with process or organisational innovation. No significant combination can be observed.

6.6 Empirical Results of Panel Data Model

This section presents the estimation results using the panel data for the three stages of the structural econometric model.

6.6.1 The Decision and Expenditure Stage

In the first stage, the Heckman model is employed to mitigate potential selection bias, using two equations. The first equation is the participation equation, which estimates the propensity of a firm to be involved in activities that may lead to innovations. Additionally, the IMR is calculated. The second equation is the expenditure equation, which estimates the firm's expenditure on innovation activities.

Estimation Results

Table 6.18 shows the estimation results for the Heckman model in two columns. The left column contains the OR of the participation equation estimated with the Logit model and the corresponding statistics. The right column contains the coefficients of the equation that expresses expenditure on innovation activities.

Table 6.18 Estimation results of 1st stage (panel)

	(1)		(2)	
	D		l_INE	
	OR	z	coef.	t
L.l_P	1.538***	(4.17)	-0.162***	(-6.04)
l_SIZE	2.511***	(18.69)	-0.0763***	(-5.33)
BRANCH	10.72***	(17.55)	0.481***	(13.00)
IMR			0.0435	(1.25)
N	11280		6486	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The estimation results of the first equation using the Logit model presented in the left column of table 6.18 show that OR is larger than one, which means that lagged labour productivity has a positive significant impact on the decision of a firm to innovate. A similar impact can be found for firm size, which positively affects the decision to conduct innovation. Furthermore, the research intensive industry clearly tends to decide to involve themselves in innovation activities more than other industries.

The estimation results of the second equation using the Tobit model presented in the the right column of table 6.18 show a high significant negative impact for lagged labour productivity on the amount of innovation expenditure, which means that productive firms do not tend to expend on innovation. Firm size shows a similarly negative impact, which means that large firms spend less on innovation. Furthermore, the influence of the IMR seems not to be significant for this stage. Hence, the null hypothesis stating no selection bias can not be rejected.

The estimation results of the additional determinants identified for this stage are given in table C.1, table C.3, table C.4, table C.5, table C.6, and table C.2.

The left column of the table C.1 shows the impact of the captured innovation constraints on the firm's decision to innovate or not. It seems that firms which reported internal organisational problems tend to involve in innovation activities. Because only innovators would report problems about innovation, which does not conclude necessarily a causal relationship. The right column of the table C.1 proposes that the firms which reported long administration and approval processes have a higher expenditure on innovation, which could be an reaction to compensate the long administrative time without have a positive meaning in the relationship.

Table C.2 shows that firms which receive public subsidies from the federal government or the EU tend to involve in innovation activities, however, this is a recursive relationship because starting an innovation project may be a condition to apply for public subsidies. However, as seen in the right column, receiving public subsidies does not necessarily lead to an increase in innovation expenditure because even the relationship seems positive but very weak.

The left column of the table C.3 shows that firms use of protection measure such as patents, registered design, or trademarks relating to the decision have a higher likelihood of deciding to innovate. The right column of table C.3 proposes a positive significant relationship between the use of patents and the level of expenditure.

The left column of the table C.4 shows the impact of market characteristics on the firm's decision to innovate. It seems that firms reported that products and services may quickly go out-of-date and firms which are present on foreign markets were mostly firms which decide to innovate. The right column of table C.4 proposes that the same market characteristic is a significant driver of the amount that a firm expends on innovation.

Table C.5 proposes in general that cooperation partnerships are important for the firm's decision to innovate or not. Results presented in the left column of the table suggest that firms which cooperate with customers, suppliers, consultants or universities inside Germany have a higher propensity to innovate. The right column of table C.5 shows that the cooperation with foreign consultants is associated with higher expenditure on innovation. However, other cooperation partnerships do not seem playing a significant role in shaping the level of expenditure on innovation.

Table C.6 proposes in general that the source of information that a firm uses for innovation is important for the innovation decision. The left column of the table shows that firms which rely on sources of information needed for innovation tend to decide to get involved in innovation activities. The right column of table C.6 proposes that firms tend to expend more on innovation if they rely on their own group, universities, or research institutes as a source for information needed for innovation.

Test Results

For the participation equation, the results of the Hausman test are shown in table C.7 for the participation equation, which suggests fixed effects as an appropriate estimator for the first equation. However, this suggestion cannot be followed for the following reasons: Firstly, as the FE estimator makes use of *within-group* variation and the participation equation incorporates a binary dependent variable, all firms which decide to conduct innovation activities in all observed time points are dropped as a consequence of the lack of within-variation. Secondly, in the participation equation the persistency of the innovation decision is considered to be an explanatory variable. Due to the unbalanced nature of the data and the restriction of within-group variation of the dependent variable, the lagged decision is always contrary to the actual one. This will lead to a negative coefficient for the persistence of innovation and inconsistent estimates (for T=const and $N \rightarrow \infty$). For these reasons, the RE model is used.

As shown in table C.8, AIC and BIC indicate results from the Likelihood Ratio (LR) test of the Logit function between the pooled model and the RE panel model. The LR test fails in testing the pooled model against the RE panel set-up and the χ^2 statistic appears to be negative. However, both AIC and BIC indicate that the panel data model is appropriate. Moreover, the Hausman test between pooled and fixed effects model suggests the same results.

As shown in table C.9, the censoring indicator of the innovation expenditure INEX is highly significant, which means that the censored values do affect the estimation results. In this case, a Tobit model is appropriate to solve the issue of censoring dependent variable. The Tobit model can only be estimated with RE in the panel model setup because FE estimation is affected from incidental parameter problem. Hence, a Hausman test is not possible for the second equation.

For the expenditure equation, table C.10 shows the LR test between the Tobit pooled model and the Tobit panel model, which indicates that the panel model is appropriate. However, table C.11 shows that both models propose the same conclusion.

The results of testing multicollinearity are shown in table C.12 for the set of explanatory variables used in this stage. No VIF value higher than 10 is found, therefore, the relevant data can be seen to be free from multicollinearity.

6.6.2 The Knowledge Production Stage

In the second stage, the knowledge production function expresses the relationship between innovation expenditure as input to the innovation process on the right hand side, and product, process, and organisational innovation as output of the innovation process on the left hand side. In this stage, the dependent variables are ordinal indicators. Therefore, an ordered probit model is used.

Estimation Results

The estimation results are reported in table 6.19, table 6.20 and table 6.21 sequentially. To account for potential endogeneity bias, the logarithm of innovation expenditure l_INE has been instrumented using the predicted value estimated in the first stage. The IMR controls for selection bias.

Table 6.19 shows that the lagged logarithm of predicted innovation expenditure is highly significant for the share of turnover resulting from new or clearly improved product PRD_IMPR. However, it seems insignificant for turnover resulting from unchanged or slightly changed products PRD_NCHG and for turnover resulting from market novelties PRD_MNOV.

Table 6.19 Estimation results of 2nd stage product innovation (panel)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.888***	(5.09)	-0.474*	(-2.04)	0.414*	(2.08)
l_SIZE	-0.0375	(-0.89)	0.139*	(2.45)	0.0408	(0.85)
EMPL_UNI	0.0828***	(4.64)	-0.0997***	(-4.09)	0.134***	(6.46)
IMR	-0.256**	(-3.11)	0.393***	(3.44)	-0.232*	(-2.36)
N	2590		2279		2564	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Furthermore, firm size does not seem important in this stage. However, the intensity of employees with university degree has a clearly positive impact on the turnover resulting from product improvements and market novelties, but has a complimentary negative impact on turnover resulting from unchanged products. The estimation results of IMR propose that the selection bias is significant in this equation.

Table C.14 shows that using patents as protection measure is associated with a higher turnover resulting from market novelties. However, this would be an expected behaviour that firms protect their novelty products via patenting them.

Table C.15 shows that firms which are threaten that their products gets quickly out-of-date tend to generate all types of product innovations. However, firms which products can be easily substituted by competitor's products tend to make their turnover from unchanged or slightly changed products and keep back from product innovations.

Table C.16 shows no significant relationship between the cooperation partnership and generating product innovation. Table C.17 shows that firms which rely on competitors to obtain information for innovation produce markedly fewer market novelties. Finally, table C.18 shows that firms which receive public subsidies from the federal government or the EU are more able to generate more product innovation than those do not.

Table 6.20 shows that the lagged logarithm of predicted innovation expenditure and the other control variable do not appear to affect the process innovation, which means that the study was not able to support identifying drivers for process innovation.

Similarly, table 6.21 shows that the lagged logarithm of predicted innovation expenditure and the other control variable has no clear impact on generating organisational innovation,

Table 6.20 Estimation results of 2nd stage process innovation (panel)

	(1) PRC_COST		(2) PRC_QUAL	
L.l_INEp	0.340	(1.49)	0.384	(1.69)
l_SIZE	-0.00417	(-0.07)	-0.102	(-1.73)
EMPL_UNI	0.0152	(0.61)	0.00389	(0.16)
IMR	-0.268*	(-2.25)	0.104	(0.92)
N	1667		1514	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.21 Estimation results of 2nd stage organisational innovation (panel)

	(1) ORG_TIME		(2) ORG_QUAL		(3) ORG_COST	
L.l_INEp	-0.252	(-0.91)	-0.605*	(-2.38)	-0.180	(-0.76)
l_SIZE	0.0770	(1.13)	-0.0598	(-1.01)	0.113	(1.96)
EMPL_UNI	0.0272	(0.96)	0.0202	(0.80)	-0.00518	(-0.22)
IMR	0.0685	(0.50)	-0.0123	(-0.10)	0.234*	(1.99)
N	511		512		509	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

which means that study was not able to support identifying drivers for organisational innovation.

Test Results

In the second stage, the dependent variables are ordinal indicators. Therefore, ordered probit models are the most used ordinal regression techniques. However, these models do not support the fixed effects estimation, which made executing the Hausman test impossible.

The results of testing multicollinearity are shown in table C.19 for the set of explanatory variables used in this stage. No VIF value higher than 10 is found, therefore, the relevant data can be seen to be free from multicollinearity.

Dependent variables are ordinal one and not censored, therefore, no additional test is needed for this stage.

6.6.3 The Production Stage

Principal Component Analysis

Table C.26 presents the correlation matrix between the explanatory variables of the third stage. Most of these variables are highly correlated to each other, which indicates a high potential to be composed of less components. This has been confirmed by the multicollinearity test in table C.27.

Table C.29 shows that the eigenvalues of the first two components are higher than one. The row *Proportion* shows the variation explained by each component and the row *Cumulative* shows the commutative explanation of data by these components. The first two components explain about 94% of the data variation. Furthermore, table C.29 shows the PCA results and the last column shows the portion of the data that is still unexplained by these components. Based on the PCA, the first two components are considered as representative for the eight original variables from different types of innovations as seen in table 6.22, which explain finally 88% of the data variation.

The test results in table C.30 propose that the explanatory variables after PCA can be seen to be free from multicollinearity.

Table 6.22 PCA results after choosing the significant components (panel)

	Comp1	Comp2	Unexplained
PRD_IMPRp	0.4618	-0.0130	0.03525
PRD_NCHGp	-0.4537	-0.0430	0.06458
PRD_MNOVp	0.3860	0.3269	0.05947
PRC_COSTp	0.4052	0.2653	0.08208
PRC_QUALp	0.0241	-0.5901	0.1286
ORG_TIMEp	-0.1288	0.5965	0.03738
ORG_QUALp	-0.2743	0.2874	0.4537
ORG_COSTp	-0.4187	0.1846	0.122
Variance	4.45296	2.56399	-
Difference	1.88898	.	-
Proportion	0.5566	0.3205	-
Cumulative	0.5566	0.8771	-

As shown in table 6.22 in the first component, product improvement, market novelties (product innovation), and cost reduction (process innovation) are loaded positively, but unchanged products (product innovation) and all types of organisational innovation are loaded negatively. In the second component, market novelties (product innovation), cost reduction (process innovation), and all types of organisational innovation are positively loaded, whereas quality improvements (process innovation) is negatively loaded. Because these two components do not contain all information in the data, the column 'Unexplained' includes the unaccounted variances in the variables, which equal the sums of squares of the loadings in the dropped components, weighted by their eigenvalues.

Estimation Results

Table 6.23 Estimation results of 3rd stage using PCA components (panel)

	(1) l_P	
L.pc1	-0.0160	(-1.93)
L.pc2	0.177***	(11.13)
l_SIZE	-0.0325	(-1.59)
l_INVS	-0.0365**	(-3.28)
IMR	-0.157***	(-4.51)
N	2094	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6.23 shows the estimation results of the third stage using a Tobit model. The covariates related to innovation outcome are proxied by the components resulting from PCA as explanatory variables. The transformation of the regression coefficients resulting from loadings into the components and the rotation is given in table 6.24.

Table 6.24 Transformed estimation coefficients after components regression (panel)

	PCR Coefficients
PRD_IMPRp	-0.0097
PRD_NCHGp	-0.0003
PRD_MNOVp	0.0516
PRC_COSTp	0.0404
PRC_QUALp	-0.1046
ORG_TIMEp	0.1074
ORG_QUALp	0.0552
ORG_COSTp	0.0393

The results of the Tobit regression indicate a highly significant positive relationship between the second component and labour productivity, whereas the first component does not seen to be significantly related. Regarding the re-transformed coefficients from the loadings of the original variables in the components, it appears that the turnover resulting from market novelties by product innovation, the reduction of average costs by process innovation and organisational innovation in general impact the labour productivity positively, whereas the increase of turnover resulting from quality improvement by process innovation appear to negatively influence labour productivity.

Beyond that, the intensity of physical capital appears negatively related. The IMR appears to be significantly related, which indicates that the selectivity issue is relevant at this stage of the model. The results propose that the firm's size is not significant for labour productivity.

Table C.21 shows that using protection measure is insignificant for labour productivity. Table C.22 shows that firms which are present on foreign markets are more productive than those are not. Table C.23 shows that cooperation with competitor impact negatively firm's labour productivity. Finally, table C.24 shows that receiving public subsidies from the government or the EU impact negatively firm's labour productivity.

Table C.25 shows the estimation results of the first stage for dummy variables which present different combinations of product innovation with process or organisational innovation. Firms which conduct product innovation together with organisational innovations targeted at reduction of costs appear to have a higher labour productivity.

Test Results

As discussed above, the results of the multicollinearity test presented in table C.27 shows high VIF values for the explanatory variables used to proxy different types of innovation. However, after using the PCA approach to generate components that are highly correlated to the original explanatory variables but orthogonal to each other. The results of testing the multicollinearity of these components presented in table C.30 show that the approach has solved the issue of multicollinearity and no VIF value higher than 10 is available except for firm size, which is slightly above the threshold and can be tolerated.

Table C.32 shows that the dummy indicated for censoring is significant for the estimation, therefore a Tobit model is appropriate for handling the censored samples.

Chapter 7

Conclusion

7.1 Overview of the Study

This work has investigated the relationship between innovation expenditure, innovation outputs, and firm productivity using German manufacturing data. It explores the impact of different factors which may affect these relationships.

In particular, chapter 2 includes definitions of productivity and innovation, describes the innovation process, and highlights the different types of innovations and their impact on productivity. The chapter reviews the chronological evolution of the milestones in studying the relationship between innovation and productivity, starting from the knowledge production function (Pakes and Griliches, 1984), the advanced approach of the CDM framework (Crepon et al., 1998), and the diverse extensions and improvements available in the previous literature. It also summarises the search for factors that drive innovation and productivity, e.g. firm size, human resources, or physical capital. The chapter concludes with a conceptual framework, which represents the innovation process in three stages and determines the main drivers which may affect each stage of the process.

Chapter 3 explains the methodology followed in this work, describes the philosophical position, the research approach, and the rationale for a quantitative research approach. It also presents the research design and the subsequent research process followed in this work.

Chapter 4 describes the main characteristics of the MIP data used to estimate the model, which contains unbalanced data for German firms collected between 2003 and 2013. In this study, unbalanced data is used to keep the number of samples as large as possible in order to gain statistical power. Using balanced data with less samples would carry the risk of bias caused by selecting only frequently-responding firms. However, to improve data balance, the samples which are present in the data for fewer than four years have been dropped from the analysed dataset.

Chapter 5 justifies the methods applied in this research to account for different econometric issues such as heterogeneity, multicollinearity, endogeneity, and selection bias. It also explains how to deal with the data issues addressed in chapter 4, such as censoring in the relevant variables. Furthermore, it describes in detail how the understanding gained from the conceptual framework is implemented in an econometric model using the CDM approach. Finally, this chapter presents the estimation strategy and provides the rationale for the employed methods.

Chapter 6 contains an assessment of the proposed model and empirical results, and outcomes of the regression analysis. The analysis was done using a panel data model and the robustness of the findings was checked by pre-estimating the same econometric model using a pooled data model. The results of tests such as the Hausman test, LR test, non-linearity test, and multicollinearity test are evaluated. Finally, the results of regression analysis of both panel and pooled data models are evaluated and explained.

This chapter tries to answer the research questions by reflecting on the results obtained from chapter 6 to support the research hypotheses with empirical evidence. It also shows how the research aim and the research objectives have been met. Section 7.1 presents an overview of previous chapters. Section 7.2 draws insights from the analysis results and explains each research hypothesis in the context of the results to answer the research questions. Section 7.4 summarises the contribution of this research to knowledge and how this research has closed the gaps addressed in the literature. Section 7.5 presents the limitations of this study. Section 7.3 presents the main implications of the study at micro and macro economic level. Section 7.6 provides some suggestions for future research in this area.

As stated in section 1.5, the research questions are:

1. What is the relationship between innovation and productivity?
2. What are the key determinants affecting innovation and productivity?

The research hypotheses formulated in section 2.6 and the regression analyses conducted in chapter 6 answer these research questions. The research hypotheses are:

- HP1: Labour productivity positively affects the firm's decision to engage in innovation.
- HP2: Labour productivity positively affects the firm's level of innovation expenditure.
- HP3: The level of innovation expenditure positively affects the generation of different types of innovations.
- HP4: Innovation positively affects a firm's labour productivity.

The econometric model developed to test these hypotheses consists mainly of three stages and considers the reciprocal link between productivity and innovation by taking previous labour productivity as an input to the firm's decision to innovate and how much a firm might spend on innovation, as proposed by Baum et al. (2015) and Raymond et al. (2013). It also emphasises the dynamic link between dependent and independent variables in each stage of the innovation process by taking the lagged value of the independent variables in order to account for the time dimension that the innovation process needs between the input and the output, as proposed by Raymond et al. (2013) and Peters (2007). The model considers the extension of Peters (2007) and Parisi et al. (2006) for process innovation, and the extension of Polder et al. (2009) for organisational innovation on the interface between knowledge production function and the production function. Because the various types of innovations as inputs to the production function are highly correlated, a PCA approach is employed to mitigate multicollinearity.

The first stage explains the firm's decision to participate in innovation activities and the level of innovation expenditure in relation to lagged labour productivity. The Heckman model, which consists of two equations, controls for potential selection bias. The first is the participation equation estimated using the Logit model, which assesses the impact of lagged labour productivity on the firm's decision to participate in innovation activities or not. Additionally, the IMR is calculated to control for selection bias in each stage. The second equation evaluates the impact of lagged labour productivity on the level of innovation expenditure for firms which carry out innovation. To estimate this equation, a Tobit model is employed because the dependent variable (innovation expenditure) is right-censored. The analysis shows that controlling for selectivity bias is important for model specifications, but not in all stages.

The second stage examines the knowledge production function, in which innovation expenditure generates economically valuable knowledge in the form of different types of innovations. Each type of innovation is estimated in a separate equation using an ordered probit model because the dependent variables (different innovation outputs) are ordinal.

The third stage is the production function, which assesses the impact of different types of innovations on labour productivity. Due to the high number of explanatory variables in this stage (eight different types of innovations) and to account for multicollinearity, a PCA is conducted. This reduces the number of explanatory variables by generating components, which are correlated to the original explanatory variables but uncorrelated to each other, so that all types of innovations might be estimated in this stage. A Tobit model is employed to estimate this equation because the dependent variable (labour productivity) is censored.

Furthermore, this research accounts for various data issues addressed in chapter 4, and the relevant econometric issues addressed in chapter 5. A single-equation approach is employed to estimate the structural model, which improves robustness against misspecification and allows the estimation of variables with different natures, such as censored, ordinal, or dummy variables. This work accounts for endogeneity in each stage by using the predicted variables resulting from the previous stage, and also accounts for selection bias by using the Heckman model in the first stage.

7.2 Main Findings

This section aims at answering the research questions, linking the empirical evidence obtained to the research hypotheses, summarising the supporting argumentation, and comparing the results with previous research.

In the first equation of the Heckman model, the effect of lagged productivity on the firm's decision to participate in innovation activities is tested using a Logit model. Examining the data shows that lagged labour productivity positively affects the firm's decision to take part in innovation activities. These findings support our first research hypothesis that labour productivity positively affects the firm's decision to engage in innovation and in agreement with Peters et al. (2013). However, the findings are not in line with Baum et al. (2015), who did not find a significant effect for Swedish manufacturing firms but only for service firms.

In the second equation of the Heckman model, the effect of lagged labour productivity on the level of expenditure is tested using a Tobit model to account for censoring of the innovation expenditure as a dependent variable. Data evaluation shows that firms with higher lagged labour productivity expend less on innovation. This suggests the second research hypothesis that labour productivity positively affect the firm's level of innovation expenditure, should be rejected. These findings also differ from the results found by Baum et al. (2015). In their set up, lagged labour productivity only has a significant effect on innovation expenditure for 'other services', and in contrast to our findings, they report a positive relationship. However, Raymond et al. (2013) and Peters et al. (2013) do not report evidence for the impact of lagged productivity on R&D expenditure.

In the knowledge production function, the effect of expenditure level on generating different types of innovation is tested using an ordered probit model. Evaluation of the data shows that the level of expenditure on innovation has a positive impact on the generation of product innovation in the form of new or improved products, but this relationship is not significant for market novelties. This may be due to the fact that market novelties are radical innovations or inventions which are not related to traditional product development. This finding is in agreement with the main body of previous literature investigating the impact of R&D or innovation expenditure on product innovation, such as Crepon et al. (1998), Van Leeuwen and Klomp (2006), Janz et al. (2004), Peters et al. (2013), Roberts and Vuong (2013) and in accordance with Lööf et al. (2001) regarding market novelties for Swedish data. Furthermore, data inspection found no relationship between innovation expenditure and the generation of process or organisational innovation. This results are in line with Parisi et al. (2006) for process innovation in Italian data, but not in line with Peters (2007) for German data, who find that innovation expenditure has a weak positive impact on process innovation. For organisational innovation, the results are in line with Polder et al. (2009), who reported that ICT had a positive impact on organisational innovation, but that R&D expenditure does not. Therefore, the results support the third research hypothesis that innovation expenditure positively affects generation of new-to-the-firm or significantly improved products but not process and organisational innovations.

In the production function, the effect of different innovation outputs on labour productivity is tested using a Tobit model to account for censoring of the labour productivity as a dependent variable. Because the various types of innovation are highly correlated, and to allow testing of the impact of the different types of innovation in the production equation, a PCA approach is employed. This results in a set of components that are correlated with the different types of innovation but uncorrelated with each other. It appears that the impact of product innovation on firm's labour productivity depends on the novelty of the innovation. Firms which make their turnover from products that are new to the firm, or from unchanged products have slightly lower labour productivity. However, those which make turnover from market novelties on average report higher labour productivity.

Process innovation targeted at the reduction of average costs has a positive impact on labour productivity, which is in line with the study by Parisi et al. (2006), Peters (2007), Peters et al. (2013), and Roberts and Vuong (2013). However, it seems that process innovation targeted at increasing turnover as a result of quality improvement impacts labour productivity negatively, at least in the short run. This might be the case because quality improvement is a long-term process and it takes some time until efforts bear fruit. In contrast, all types of organisational innovation positively affect labour productivity. These findings are in line with Polder et al. (2009) and support the fourth research hypothesis that innovation positively affects a firm's labour productivity.

Having tested the impact of different combinations of types of innovation in the production function, it seems that firms which conduct product innovation in association with organisational innovations targeted at cost reduction have higher labour productivity. This finding is in line with Polder et al. (2009).

The analysis shows that large firms tend to be involved in innovation activities. However, of the firms which participate in innovation, large firms expend less on innovation in relation to their turnover than small firms. This result is consistent with the rejection of the second hypothesis and with the inferential statistics presented in section 6.4. Similar results came out by Nguyen and Martin (2010) for Luxembourger firms. Furthermore, firms which are classified as being in the research-intensive sector take part in more innovation activities and expend more on innovation than those from other sectors. In the pooled data model, small firms seem to be more productive than large firms, however, this result could not be confirmed in the panel data model or the inferential statistics because they show that large firms are more productive.

From the traditional production function, a significant negative impact of the intensity of physical capital on labour productivity is available. Baum et al. (2015) had the similar finding for Swedish firms. This stresses that knowledge capital proxied by innovation, rather than the intensity of physical capital, is a source for labour productivity.

Another finding is that receiving public subsidies from the federal government or the EU encourages firms to innovate but it seems not to be significant for the level of expenditure in innovative firms. Those firms have a higher turnover resulting from more new innovation for the firm without evidence that they generate market novelties. Zemplinerova and Hromadkova (2012) had the similar finding for Czech firms. In the end, those firms appear less productive.

It appears that firms which are present in foreign markets tend to participate in and spend more on innovation activities. They also have higher labour productivity, which is in accordance with the previous literature such as Kimura and Kiyota (2006), Hansen (2010), and Wagner (2012).

Data analysis shows that firms which face the threat of their product quickly becoming out-of-date tend to participate in and expend on innovation activities, and they also carry out more product innovation in the form of market novelties. This might indicate a clear positive effect of the Schumpeterian technology push on innovation.

Qualified personnel proxied by employees with an university degree seems to be very important for product innovation are the main driver of generating market novelties. Furthermore, relying on competitors to obtain information for innovation obviously leads to fewer market novelties. However, in the knowledge production function, no significant drivers for process or organisational innovation could be found.

It appears that firms which are members of a firm group and cooperate internationally achieve higher labour productivity. However, this may be due to the fact that spillover effect and the fact that the creation of know-how and the necessary expenses takes place in another country while the outcome is captured in the German data.

Another finding is that the level of expenditure on innovation increases because of the long administration and approval process, which does not necessarily lead to more innovation but compensates for process constraints. Moreover, firms which rely on their group, universities, or research institutes to obtain information for innovation tend to expend more on innovation. What motivates firms to participate in innovation activities is relying on their group companies, suppliers, customers, or universities as a source for the information needed for innovation. Additionally, firms which use protection mechanisms such as patents or trademarks have a higher propensity to innovate.

7.3 Implications

On the **micro-economic** level, this study promotes product innovation in the form of market novelties as an important source of labour productivity. This might be due to the high risk associated with their high potential for profit in the case of success (Aschhoff, 2013). Interestingly, market novelties appear not to be driven by innovation expenditure but mainly by the intensity of qualified personnel and the technology push. The generation of new-to-the-firm or significantly improved products has a weak negative impact on labour productivity, however, this appears to be a necessary initial step towards generating market novelties. This study demonstrates that process innovation targeted at the reduction of average costs, as well as organisational innovation in general, leads to improvement in labour productivity. Furthermore, enhancing international cooperation between firms which are members of a firm group should improve labour productivity driven by the exchanges of ideas and innovations. Furthermore, being present in foreign markets positively affects labour productivity.

On the **macro-economic** level, firms which receive public subsidies from the federal government or the EU tend to be involved in innovation activities and generate product innovation in the form of product improvements. However, those firms also appear to be less productive than those which do not receive such subsidies.

7.4 Contribution to Knowledge

This research attempts to close the gap in knowledge addressed in section 1.5 and resulting from the critical review conducted of the previous empirical studies. The contribution of this research to knowledge can be summarised as follows:

Firstly, this research improves the coverage of the existing research which tests the relationship between innovation and productivity using German CIS data. Most of the available research in this area uses cross-sectional data, while this research incorporates a panel-data approach which allows us to also control for unobserved heterogeneity and dynamic relationships. It uses data for German manufacturing firms between 2003-2013, whereas the latest research of Peters et al. (2013) and Roberts and Vuong (2013) uses panel data up to 2009.

Secondly, the reciprocal link between productivity and innovation, in which the firm's previous labour productivity might affect the decision to innovate, as proposed by Baum et al. (2015), and the level of expenditure, as proposed by Raymond et al. (2013) and Baum et al. (2015), is considered in this research. The research on this specific reciprocal link is scarce for German data. Peters et al. (2013) and Roberts and Vuong (2013) model a mutual dynamic dependency between productivity and R&D activities, which at its core incorporates the firm's decision to engage in R&D activities as a dynamic programming problem. Our approach, however, is based on the CDM framework and extends this set up to also account for this reciprocal link. Furthermore, it allows for more refined analysis as our model includes not only the innovation decision but also the extent of innovation expenditure.

Thirdly, the main CDM framework (Crepon et al., 1998) and most of the previous empirical work on this topic use product innovation as an input to the production function. Parisi et al. (2006) and Peters (2007) expanded the framework to include process innovation, and Polder et al. (2009) and Hall et al. (2012) expanded it to encompass organisational innovation. The impact of organisational innovation on productivity has not been tested for German data before.

Fourthly, due to the high correlation among various types of innovation, it was impossible to consider them together as inputs in the production function. Therefore, Hall et al. (2012) estimates a firm's predicted probability of innovation and uses it as an input in the production function to proxy innovation. Peters et al. (2013), Roberts and Vuong (2013), and Moreno and Huergo (2010), also take binary/dummy variables to proxy innovation outcomes. However, this level of abstraction does not allow refined understanding of which type of innovation is most relevant for productivity. This research provides a novel methodology to take different types of innovations as input to the production function of the CDM framework. It employs a PCA approach to generate a number of innovation components which are correlated to the different types of innovations but uncorrelated to each other. These components were used to estimate the impact of various types of innovation as inputs to the production function. Therefore, investigating the impact of different types of innovation, especially product innovation in the form of market novelties or organisational innovation, has not been done before with German data.

Table 7.1 presents an overview of the latest relevant studies to have carried out extensions for the CDM framework and highlights the methodological contribution of this research.

Table 7.1 Methodological comparison between this research and latest key studies

	Baum et al. (2015)	Raymond et al. (2013)	Peters et al. (2013)	Polder et al. (2009)	This research
Decision					
Productivity as input	Yes	No	Yes	No	Yes
Time lag	1 year	No	1 year	No	1 year
Expenditure					
Productivity as input	Yes	Yes	No	No	Yes
Time lag	1 year	1 year	No	No	1 year
Knowledge production function					
Innovation input	R&D expenditure	R&D expenditure	R&D expenditure	R&D expenditure	Innovation expenditure
ICT as input	No	No	No	Yes	Not possible*
Time lag	1 year	1 year	1 year	No	1 year
Cobb-Douglas production function					
Product innovation	Dummy (0/1)	Dummy (0/1)	Dummy (0/1)	Dummy (0/1)	Share of sales; market novelties
Process innovation	No	No	Dummy (0/1)	Dummy (0/1)	Cost reduction; increasing turnover
Org. innovation	No	No	No	Dummy (0/1)	Three effect indicators **
Time lag	No	1 year	1 year	No	1 year
Data					
Source	Sweden	The Netherlands and France	Germany	The Netherlands	Germany
Year	2006-2012	1994-2004	2003-2009	2002-2006	2003-2013

* See limitations in section 7.5
** See table 5.3

7.5 Limitations

This section discusses the weaknesses and limitations of this study. In general, the quality of its results is related to the quality of the investigated data. As discussed in chapter 4, the dataset used suffers from different issues such as being highly unbalanced, the fact that most of the variables of interest were censored, and in addition that the ordinal representation of variables causes a loss of information. Moreover, using some variables of interest led to a dramatic reduction in the number of observations and thereby decreases the results' statistical power. Even though the econometric methods described were applied in order to reduce the impact of data quality on the analysis; these should still be considered as limitations.

Firstly, the early intention of this study was to investigate the impact of both innovation expenditure and investment in ICT on innovation within German firms. It was motivated by the approach proposed by Polder et al. (2009) and Hall et al. (2012), which states that the former may affect product and process innovation, while the latter may affect organisational innovation. The lack of data integration between MIP and ICT survey data made it impossible to conduct this investigation. Each survey uses different firm identifiers, thus it was impossible to merge both data sources in order to test this relationship. This study tried to find other proxies for ICT such as the turnover of physical capital, however, no empirical evidence could be found in the dataset to support this.

Secondly, this research was unable to find empirical evidence within the investigated factors for determinants that may drive process or organisational innovations.

Thirdly, an important area to be intentionally investigated is the impact of individual and organisational knowledge accumulation on innovation and productivity measured by firm age and employee age. It was not possible to investigate this for the knowledge production function due the absence of these indicators in the latest surveys.

Finally, investigating the impact of firm's regional area location on innovation and productivity was not possible due to the lack of this information in the dataset. The only available indicator is whether a firm operates in East or West Germany, which is very generic.

7.6 Suggestions for Future Research

In this study, the relationships between productivity and innovation inputs, innovation inputs and innovation outputs, and innovation outputs and productivity have been investigated. In this section, some proposals for future research will be suggested.

The MIP data has a high potential for use in different research topics for two reasons: Firstly, because the large number of indicators captured in the data offers attractive opportunities to inspect different aspects of innovation activities. Secondly, the data is becoming more mature and is expanding quantitatively and qualitatively over time, which may enable understanding of the innovation process and the relationships between the different types of innovations and their key determinants. Furthermore, the survey questionnaire is expanding, leading to new data about micro innovation behaviour and backgrounds. Future work might thus lead in the following directions:

Firstly, more indicators are captured in the MIP survey, such as the indicator for financing innovation activities (available from 2007), and indicators for the motivation of foreign innovation activities (available from 2009). Additionally, new indicators are available which cover modern topics and extensions to the innovation economy such as environmental innovations or open innovations.

Secondly, the German data for organisational and marketing innovation have not been sufficiently investigated yet because the focus of most previous studies was on production and process innovation. Therefore, this area is an important potential area for gaining better understanding about these important types of innovations.

Thirdly, the impact of ICT on both innovation and productivity, as proposed by Polder et al. (2009), has not been investigated yet for German data. As mentioned in section 7.5, this topic has not been sufficiently examined for German data. Unifying the identification number of the sample firms within the datasets of MIP and ICT by ZEW would offer the opportunity to investigate this important topic empirically for German data. This research area could be pursued in two directions: Firstly, the use of ICT is expected to affect the effectiveness and efficiency of the innovation process for both manufacturing and service sectors; this can be named 'productivity of innovation process'. Secondly, investigation of the direct link between the use of ICT and firm performance or labour productivity in comparison to the link through organisational innovation.

Fourthly, the innovation process in the knowledge-intensive service sector might be a subject of future research. As discussed in chapter 1, the innovation economy shifts the firm's business spectrum from the manufacturing sector to the knowledge-intensive sector, which is growing rapidly in Germany. Therefore, investigating this sector, using a similar model to that used in this work, is an open topic for future research.

Finally, the productivity of R&D activities should be studied in more detail by analysing different indicators to obtain understanding of innovation as a white box system and to find out how the efficiency and effectiveness of the innovation process can be optimised.

References

Abdih, Y. and Joutz, F. (2005). Relating the knowledge production. function to total factor productivity: An endogenous growth puzzle. Working paper, International Monetary Fund.

Abowd, J. M., Haltiwanger, J., Jarmin, R., Lane, J., Lengermann, P., McCue, K., McKinney, K., and Sandusky, K. (2002). The relation among human capital, productivity and market value: Building up from micro evidence. Technical report, NBER/CRIW conference on "Measuring Capital for the New Economy".

Acs, Z. J. and Audretsch, D. B. (1988). Innovation in large and small firms: An empirical analysis. *American Economic Review*, 78:678–690.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and innovation: An inverted u-relationship. *The Quarterly Journal of Economics*, pages 701–728.

Aghion, P. and Howitt, P. (1998). *Endogenous Growth Theory*. MIT Press.

Alavifar, A., Karimimalayer, M., and Anuar, M. K. (2012). Structural equation modeling vs multiple regression. *IRACST – Engineering Science and Technology: An International Journal (ESTIJ)*, 2(2).

Almus, M. and Czarnitzki, D. (2002). The effects of public r&d subsidies on firms' innovation activities: The case of eastern germany. Discussion Paper 01-10, Centre for European Economic Research (ZEW).

Andrew, J., Haanaes, K., Michael, D., Sirkin, H., and Taylor, A. (2009). Measuring innovation 2009. the need for action. Senior management survey, The Boston Consulting Group (BCG).

Araujo, N. and Costa, L. (2012). The productivity of innovation in portugal. Technical Report 05, Catholic University of Portugal at Porto.

Armbruster, H., Bikfalvi, A., Kinkel, S., and Lay, G. (2008). Organizational innovation: The challenge of measuring non-technical innovation in large-scale surveys. *Technovation*, 28:644–657.

Arrow, K. (1962). *The Rate and Direction of Inventive Activity: Economic and Social Factors*, chapter Economic Welfare and the Allocation of Resources for Invention, pages 609–626. UMI.

Arun, M. (2013). *Driving the Economy through Innovation and Entrepreneurship. Emerging Agenda for Technology Management*, chapter Innovations in an Emerging Software Cluster, pages 499–509. Springer.

Arundel, A., O'Brien, K., and Torugsa, A. (2013). *Handbook of Innovation Indicators and Measurement*, volume 4, chapter How Firm managers understand innovation: implication for the design of innovation surveys, pages 88–108. Edward Elgar Publishing.

Aschhoff, B. (2013). Innovation in germany - results of the german cis 2006 to 2010. Background Report 13-01, Center for European Economic Research (ZEW).

Atkinson, R. and Wial, R. (2008). Boosting productivity, innovation, and growth through a national innovation foundation. Technical report, The Information Technology and Innovation Foundation (ITIF).

Audretsch, D. B. and Lehmann, E. E. (2015). *The Seven Secrets of Germany. Economic Resilience in an Era of Global Turbulance*. Oxford University.

Balasubramanian, N. and Lee, J. (2007). Firm age and innovation. *Industrial and Corporate Change*, 17(5):1019–1047.

Baltagi, B. (2005). *Econometric Analysis of Panel Data*. J. Wiley & Sons, 3 edition.

Baltagi, B. (2011). *EconomerEco, Fifth Edition*. Sprinter, 5 edition.

Bartelsman, E. J., Haltiwanger, J. C., and Scarpetta, C. (2009). Cross-country differences in productivity: The role of allocation and selection. NBER Working Paper 15490, The National Bureau of Economic Research.

Baum, C., Lööf, H., Nabavi, P., and Stephan, A. (2015). A new approach to estimation of the r&d-innovation-productivity relationship. Boston College Working Papers in Economics 876, Boston College Department of Economics.

Belderbos, R., Carree, M., and Lokshin, B. (2004). Cooperative r&d and firm performance. Technical report, ELSEVIER.

Bennett, J. and McGuinness, S. (2009). Assessing the impact of skill shortages on the productivity performance of high-tech firms in northern ireland. *Applied Economics*, 41(6):727–737.

Benoit, K. (2011). Linear regression models with logarithmic transformations. resreport, London School of Economics.

Bertschek, I. and Meyer, J. (2008). Do older workers lower it-enabled productivity? firm-level evidence from germany. Zew discussion paper, Centre for European Economic Research (ZEW).

Bloom, N., Fischer, G., Rasul, I., Rodriguez-Clare, A., Suri, T., Udry, C., Verhoogen, E., Woodruff, C., and Zane, G. (2013). Firm capabilities and economic growth. Igc evidence paper, International Growth Center.

Bond, S., Elston, J., Mairesse, J., and Mulkay, B. (1997). Financial factors and investment in belgium, france, germany, and the uk: a comparison using company panel data. Working Paper 5900, National Bureau of Economic Research.

Bond, S., Harhoff, D., and Van Reenen, J. (2003). Corporate r&d and productivity in germany and the united kingdom. Technical report, Centre for Economic Performance, London School of Economics and Political Science.

Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, 117:339–376.

Bryan, F., Manly, J. A., and Navarro, A. (2016). *Multivariate Statistical Methods: A Primer, Fourth Edition*. Chapman and Hall/CRC, 4 edition.

Bryman, A. (2012). *Social Research Methods, 4th Edition*. Oxford University Press.

Brynjolfsson, E. and Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, 14(4):23–48.

Brynjolfsson, E., McAfee, A., Sorell, M., and Zhu, F. (2006). Scale without mass: Business process replication and industry dynamics. Technical Report 07-016, HBS Working Paper.

Caldwell, B. J. (2003). *Beyond Positivism Economic Methodology in the Twentieth Century*. Taylor & Francis e-Library, London and New York.

Carson, D., Gilmore, A., Perry, C., and Gronhaug, K. (2001). *Qualitative Marketing Research*. Sage.

Cassiman, B. and Martínez-Ros, E. (2007). Product innovation and exports. evidence from spanish manufacturing. Wroking paper, IESE Business School, K.U. Leuven and CEPR.

Cassiman, B. and Veugelers, R. (2001). R&d cooperation and spillover: Some empirical evidence from belgium. *American Economic Review*.

Chiswick, B. R. (2005). High skilled immigration in the international arena. Discussion Paper 1782, The Institute for the Study of Labor (IZA), Bonn.

Christensen, C. M. (1997). *The Innovator's Dilemma. When New Technologies Cause Great Firms to Fail*. Harvard Business School Press.

Christensen, C. M., Raynor, M. E., and McDonald, R. (2015). What is disruptive innovation? *Harvard*.

Ciriaci, D., Moncada-Paterno-Castello, and Voigt, P. (2012). Does size or age of innovative firms affect their growth persistence? evidence from a panel of innovative spanish firms. Working paper, Institute for Prospective Technological Studies.

Coad, A., Segarra, A., and Teruel, M. (2012). Innovation and firm growth: Does firm age play a role? Technical report, Universitat Rovira i Virgili.

Cobb, C. W.and Douglas, P. H. (1928). A theory of production. *The American Economic Review*, 18:139–165.

Coe, D. T., Helpman, E., and Hoffmaister, A. W. (2008). International r&d spillovers and institutions. Working paper 08/104, International Monetary Fund.

Cohen, W. (2010). *Handbook of the Economics of Innovation*, volume 04, chapter Fifty years of empirical studies of Innovative activity and performance, pages 129–213. Elsevier B. V.

Cohen, W. and Levinthal, D. (1989). Innovation and learning: The two faces of r & d. *The Economic Journal*, 99(397):569–596.

Cohen, W. M. and Klepper, S. (1996). Firm size and the nature of innovation within industries: The case of process and product r&d. *The Review of Economics and Statistics*, 78(2):232–43.

Cohen, W. M. and Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1):128–152.

Cooper, D. R. and Schnindler, P. S. (2013). *Business Research Methods, Twelfth Edition*. McGraw-Hill's, 12 edition.

Corrado, C., Hulten, C., and Sichel, D. (2005). Intangible capital and economic growth. measuring capital and technology: an expanded frameworks. *Studies in Income and Wealth*, 65.

Cortright, J. (2001). New growth theory, technology and learning: A practitioner's guide. Reviews of Economic Development Literature and Practice 4, U.S. Economic Development Administration.

Cotis, J. P. (2004). *Understanding Economic Growth. Macro level, Industry level, Firm level*. Organisation for Economic Co-operation and Development (OECD).

Crass, D., Licht, G., and Peters, B. (2015). *Intangible Assets and Investments at the Sector Level: Empirical Evidence for Germany*, chapter 4, pages 57–111. Springer.

Crepon, B., Duguet, E., and Mairesse, J. (1998). Research, innovation, and productivity, an economic analysis at the firm level. Working Paper 6696, NBER.

Crespi, G., Criscuolo, C., and Haskel, J. (2006). Information technology, organisational change and productivity growth: Evidence and from uk firms. *CEP Discussion Papers*, 558.

Crespi, G. and Zuniga, P. (2012). Innovation and productivity: Evidence from six latin american countries. *World Development*, 40(2):273–290.

Creswell, J. W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Method Approaches. Fourth Edition*. SAGE Publications, 4 edition.

Criscuolo, C. and Haskel, J. (2003). Innovations and productivity growth in the uk: evidence from cis2 and cis3. Discussion paper EBPF03-3, CeRiBA.

Croitoru, A. (2012). Schumpeter, j.a., 1934 (2008), the theory of economic development: An inquiry into profits, capital, credit, interest and the business cycle, translated from the german by redvers opie, new brunswick (u.s.a) and london (u.k.): Transaction publishers. *Journal of Comparative Research in Anthropology and Sociology*, 3(2):137–148.

De Loecker, J. and Goldberg, P. K. (2013). Firm performance in a global market. *The Annual Review of Economics*, 23.

D'Este, P., Iammarino, S., Savona, M., and Tunzelmann, N. v. (2012). What hampers innovation? revealed barriers versus deterring barriers. *Research Policy*, 41(2):482–488.

Donselaar, P., Erken, H., and Klomp, L. (2004). *Fostering Productivity*, volume 1, chapter R&D and Innovation: Drivers of Productivity Growth, pages 75–92. ELSEVIER.

Drucker, P. (1986). *The Frontiers of Management*. Routledge.

Drucker, P. (1999). *Management Challenges for the 21st Century*. Perfect Bound.

Dubner, S. J. (2008). How can we measure innovation? Technical report, Freakonomics. Url: http://freakonomics.com/2008/04/25/how-can-we-measure-innovation-a-freakonomics-quorum/, Url-date: 12.01.2017.

Duguet, E. and Lelarge, C. (2012). Does patenting increase the private incentives to innovate? a microeconometric analysis. *ANNALS OF ECONOMICS AND STATISTICS*, 107/108.

Ethridge, D. E. (2004). *Research Methodology in Applied Economics*. Blackwell.

Eurostat (2008). Information society: Ict impact assessment by linking data from different sources. Technical report, Eurostat.

Fallahi, F., Sojoodi, S., and Aslaninia, N. M. (2010). Determinants of labor productivity in manufacturing firms of iran: Emphasizing on labor education and training. Technical report, University of Tabriz, Department of Economics.

Fernandes, A. M. and Paunov, C. (2012). The risks of innovation. Policy Research Working Paper 6103, The World Banl.

Filippetti, A., Frenz, M., and Archibugi, D. (2009). The effects of the economic downturn on innovation: Creative destruction versus creative accumulation. Conference paper, 4th Annual Conference of the GARNET Network.

Fomby, T. B., Hill, R. C., and Johnson, S. R. (1988). *Advanced Econometric Methods*. Springer Science+Business Media, LLC.

Forth, J. and Mason, G. (2004). The impact of high-level skill shortages on firmlevel performance: Evidence from the uk technical graduates employers survey. Technical Report 235, National Institute of Economic and Social Research, London, England.

Fosse, H., Jacobsen, J., and Sorensen, A. (2013). Ict, innovation and productivity growth. Technical report, Center for Economic and Business Research, Copenhagen Business School.

Franco, C., Marzucchi, A., and Montresor, S. (2012). Absorptive capacity, innovation cooperation and human-capital. evidence from 3 european countries. Technical Report 25637, Institute for Prospective Technological Studies, Joint Research Centre, European Commission.

Frees, E. W. (2004). *Longitudinal and Panel Data: Analysis and Applications for the Social Sciences*. Cambridge University Press.

Frietsch, R., Schmoch, U., van Looy, B., Walsh, J. P., Devroede, R., Du Plessis, M., Jung, T., Meng, Y., Neuhäusler, P., Peters, B., and Schubert, T. (2010). The value and indicator function of patents. Technical report, Fraunhofer Institute for Systems and Innovation Research.

Galbraith, J. K. (1994). *A Journey Through Economic Time: A Firsthand View*. Houghton Mifflin.

Gault, F. (2013). *Handbook of Innovation Indicators and Measurement*, volume 19, chapter Innovation indicators and measurement: challanges, pages 441–464. Edward Elgar Publishing.

Gonzalez, X., Jaumandreu, J., and Pazo, C. (2005). Barriers to innovation and subsidy effectiveness. *The RAND Journal of Economic*, 36(4).

Gordon, R. A. (2015). *Regression Analysis for the Social Sciences, Second Edition*. Routledge.

Gottschalk, S. (2013). *The Scientific-Use-Files of the Mannheim Innovation Panel. Guide to the survey's datasets for external users, 1993 to 2011 surveys*. Centre for European Economic Research.

Grant, K., Hackney, R., and Edgar, D. (2009). *Strategic Information Systems Management*. Cengage Learning EMEA.

Gretton, P., Gali, J., and Parham, D. (2003). The effect of icts and complementary innovations on australian productivity growth. Technical report, Productivity Commission Canberra, Australia.

Griffith, R., Huergo, E., Mairesse, J., and Peters, B. (2006). Innovation and productivity across four european countries. *NBER Working Papers*, 12722.

Griliches, Z. (1979). R&d and productivity: The econometric evidence. *Bell Journal of Economics*, 10(1):92–116.

Griliches, Z. (1994). Productivity, r&d, and data constraints. *American Economic Review*, 84:1–23.

Griliches, Z. (1998). *R&D and Productivity: The Unfinished Business*, chapter R&D and Productivity: The Econometric Evidence, pages 269–283. University of Chicago Press.

Grossman, G. M. and Helpman, E. (1991). Quality ladders and product cycles. *Quarterly Journal of Economics*, 106:557–586.

Guellec, D. and van Pottelsberghe, B. d. l. P. (2004). From r&d to productivity growth: Do the institutional settings and the source of funds of r&d matter? Technical report, Center Emile Bernheim, Research INstitute in Management Sciences.

Guo, S. and Fraser, M. W. (2009). *Propensity Score Analysis: Statistical Methods and Applications*. Advanced Quantitative Techniques in the Social Sciences. SAGE Publications, Inc, 1 edition.

Hagen, H. O., Glanz, J., and Nilsson, M. (2008). *Yearbook on Productivity 2008*, chapter ICT investment, ICT use and productivity, pages 39–72. Statistic Sweden.

Hall, B., Lotti, F., and Mairesse, J. (2009). Innovation and productivity in smes. and empirical evidence for italy. working papers 718, Banca D' Italia.

Hall, B. H. (2011a). Innovation and productivity. Working Paper 17178, National Bureau of Economic Research.

Hall, B. H. (2011b). Using productivity growth as an innovation indicator. Report for the high level panel on measuring innovation, dg research, european commission, University of Maastricht and UC Berkeley.

Hall, B. H. (2013). Using patent data as indicators. Power point presentation, EMAEE.

Hall, B. H., Lotti, F., and Mairesse, J. (2008). Employment, innovation, and productivity: evidence from italian microdata. *Industrial and Corporate Change*, 17(4):813–839.

Hall, B. H., Lotti, F., and Mairesse, J. (2012). Evidence on the impact of r&d and ict investments on innovation and productivity in italian firms. *Economics of Innovation and New Technology*, 22(3):300–328.

Hall, B. H. and Mairesse, J. (2006). Empirical studies of innovation in the knowledge driven economy. Working paper 12320, National Bureau of Economic Research.

Hamilton, D. S. and Quinlan, J. P. (2008). *Germany and Globalization*. Center for Transatlantic Relations.

Hansen, T. (2010). Exports and productivity: An empirical analysis of german and austrian firmlevel performance. Discussion Paper 317, University of Munich.

Harhoff, D. (1990). R&d spillovers, technological proximity, and productivity growth – evidence from german panel data. *Schmalenbach Business Review*, 52:238–260.

Hashi, I. and Stojcic, N. (2010). The impact of innovation activities on firm performance using a multi-stage model: Evidence from the community innovation survey 4. *CASE Network Studies & Analyses*, 410.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer, New York.

Hausman, J. A. (1978). Specification tests in economerics. *Econometrica*, 46:1251–1271.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–166.

Helpman, E. (2004). *The Mystery of Economic Growth*. Harvard University Press.

Helpman, E., Melitz, M., and Yeaple, S. (2004). Export versus fdi with heterogeneous firms. *American Economic Review*, 94:300–316.

Henderson, R. M. (1993). Underinvestment and incompetence as responses to radical innovation: Evidence form the photolithographic alignment equipment industry. *RAND Journal of Economics*, 24(2):248–270.

Hsiao, C. (2014). *Analysis of Panel Data, third edition*. Cambridge Universirty Press, 3 edition.

Huergo, E. and Jaumandreu, J. (2004). Firms' age, process innovation and productivity growth. *International Journal of Industrial Organization*, 22:541–559.

Janz, N., Lööf, H., and Peters, B. (2004). Firm level innovation and productivity - is there a common story across countries? Discussion Paper 03-26, Center for European Economic Research (ZEW).

Javorik, B. S. (2004). Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages. *The American Economic Review*, 94(5):605–627.

Jefferson, G. H., Huamao, B., Xiaojing, G., and Xiaoyun, Y. (2002). R and d performance in chinese industry. *Economics of Innovation and New Technology*.

Jienwatcharamongkhol, V. and Tavassoli, S. (2014). Closing the gap: An empirical evidence on firm's innovation, productivity, and exports. Technical report, Center for Innovation and Technology Research.

Johansson, B. and Lööf, H. (2009). Innovation, r&d and productivity - assessing alternative specifications of cdm-models. Working Paper 159, The Royal Institute of Technology Centre of Excellence for Science and Innovation Studies (CESIS).

Johansson, B. and Lööf, H. (2010). Innovation strategy and firm performance. what is the long-run impact of persistent r&d? Working Paper 240, Centre of Excellence for Science and Innovation Studies (CESIS).

Jolliffe, I. T. (2002). *Principal Component Analysis, Second Edition*. Springer.

Jorgenson, D. W. (1991). *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth*, chapter Productivity and Economic Growth, pages 19–118. University of Chicago Press.

Junge, M., Severgini, B., and Sorensen, A. (2012). Product-marketing innovation, skills, and firm productivity growth. Technical report, Copenhagen Business School.

Kangasniemi, M. and Robinson, C. (2008). The impact of organisational change on productivity. sectoral evidence from the uk, 1995-2005. Working Paper 29, EU KLEMS Project.

Katchova, A. (2013a). Instrumental variables. Technical report, The Econometrics Academy.

Katchova, A. (2013b). Principal component analysis and factor analysis. Technical report, The Econometrics Academy.

Kemp, R., Folkeringa, M., de Jong, J., and Wubben, E. (2002). Innovation and firm performance. Research Report H200207, SCALES SCientific AnaLysis of Entrepreneurship and SMEs.

Kennedy, P. (2008). *A Guide to Econometrics, Sixth Edition*. Blackwell.

Khan, M., Luintel, K. B., and Theodoridis, K. (2010). How robust is the r&d productivity relationship? evidence from oecd countries. Working Paper 1, WIPO Economic Research Working Papers.

Kimura, F. and Kiyota, K. (2006). Exports, fdi and productivity: Dinamic evidence from japanese firms. *Review of World Economics*.

Klomp, L. and van Leeuwen, G. (2001). Linking innovation and firm performance: A new approach. *International Journal of the Economics of Business*, 8(3):343–364.

Knell, M. and Nas, S. O. (2006). What is missing in the analysis of input-output rrelationship of innovation processes? Technical report, Blue Sky II Forum.

Koellinger, P. (2008). The relationship between technology, innovation, and firm performance: Empirical evidence on e-business in europe. Erim report, Eramus Research Institute of Management.

Krugman, P. (1997). *The Age of Diminished Expectations. U.S. Economic Policy in the 1990s*. The MIT Press Cambridge, Massachusetts.

Kwark, N. and Shyn, Y. (2006). International r&d spillovers revisited: Human capital as an absorptive capacity for foreign technology. *International Economic Journal*.

Kwon, D. B. (2009). Human capital and its meaasurement. Conference paper, The 3rd OECD World Forum on "Statistics, Knowledge and Policy".

Levitt, B. and March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14:319–340.

Lööf, H. (2002). *The Knowledge Production Function : Evidence from New Micro Data*. PhD thesis, KTH, Industrial Economics and Management.

Lööf, H. and Heshmati, A. (2002a). Knowledge capital and performance heterogeneity: A firm-level innovation study. *International Journal of Production Economics*, 76(1):61–85.

Lööf, H. and Heshmati, A. (2002b). The link between firm level innovation and aggregate productivity growth. Working paper 15, Insti tutet för studier av utb i ldning och forskning, Stockholm.

Lööf, H. and Heshmati, A. (2002c). On the relationship between innovation and performance: A sensitivity analysis. Working Paper 446, SSE/EFI Working Paper Series in Economics and Finance.

Lööf, H., Heshmati, A., Asplund, R., and Naas, O. (2001). Innovation and performance in manufacturing firms: A comparison of the nordic countries. Working Paper Series in Economics and Finance 457, SSE/EFI.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22:3–42.

Lucas, R. E. (2002). *Lectures on Economic Growth*. Harvard University Press.

Mairesse, J. and Mohnen, P. (2004). The importance of r&d for innovation: A reassessment using french survey data. Technical report, MERIT - Maastricht Economic Research Institute on Innovation and Technology.

Mairesse, J. and Mohnen, P. (2010). *Handbook of the Economics of Innovation*, volume 26, chapter Using innovation surveys for econometric analysis., pages 1129–1155. Elsevier.

Mairesse, J., Mohnen, P., Zhao, Y., and Zhen, F. (2012). Globalization, innovation and productivity in manufacturing firms: A study of four sectors of china. ERIA Discussion Paper ERIA-DP-2012-10, Europe and China Co-ordination Research Fund (ERIA).

Mairesse, J. and Robin, S. (2009). Innovation and productivity: a firm-level analysis for french manufacturing and services using cis3 and cis4 data. Paper proposal, Thematic Meeting of the French Economic Association 'Firms, Markets and Innovation', Sophia Antipolis.

Mankiw, N. G. (2007). *Principles of Economics. The third edition*. Cengag.

Marx, K. (1867). *Capital: Critique of Political Economy*. Verlag von Otto Meisner.

McCraw, T. K. (2009). *Prophet of Innovation: Joseph Schumpeter and Creative Destruction*. Harvard University Press.

Meyer, J. (2008). The adoption of new technologies and the age structure of the workforce. Discussion Paper 08-045, Center for European Economic Research.

Milbergs, E. (2007). Innovation metrics: Measurement to insight. White paper, Center for International Science and Technology Policy.

Mohnen, P., Mairesse, J., and Dagenais, M. (2006). Innovativity: A comparison across seven european countries. Working Paper 12280, National Bureau of Economic Research.

Mohnen, P. and Röller, L. H. (2003). Complementarities in innovation policy. Report 022, Maastricht Economic Research Institute on Innovation and Technology.

Mohr, J., Sengupta, S., and Slater, S. (2010). *Marketing of High-technology Products and Innovations*. Prentice Hall, Pearson.

Monjon, S. and Waelbroeck, P. (2003). Assessing spillovers from universities to firms: Evidence from french firm-level data. *International Journal of Industrial Organization*, 21(9):1255–1270.

Moreno, L. and Huergo, E. (2010). Does history matter for the relationship between r&d, innovation and productivity? In *2nd Conference on corporate R&D*.

Morris, L. (2008). Innovation metrics, the innovation process and how to measure it. White paper, InnovationLabs LLC.

Mukherjee, C., White, H., and Wuyts, M. (1998). *Econometrics and Data Analysis for Developing Countries*. Routledge.

Nagaoka, S., Motohashi, K., and Akira, G. (2010). *Handbook of the Economics of Innovation*, volume 25, chapter Patent Statistics as an Innovation Indicator, pages 1083–1127. Elsevier B. V.

Nguyen, T. U. and Martin, L. (2010). Impact of r&d and ict on innovation and productivity. empirical evidence from micro data. Technical report, CEPS/INSTEAD.

OECD (1996). The knowledge-based economy. Technical report, OECD.

OECD (2001a). *Measuring Productivity. Measurement of afggrigate and industry-level productivity growth*. Organisation for Economic Co-Operation and Development (OECD).

OECD (2001b). R&d and productivity growth: panel data analysis of 16 oecd countries. OECD Economic Studies 33, OECD.

OECD (2002). *The Measurement of Scientific and Technological Activities. Frascati Manual. Proposed Standard Practice for Surveys on Research and Experimental Development*. Organisation for Economic Co-operation and Development.

OECD (2005). *Oslo Manual: Proposed Guidelines for Collecting and Interpreting Technological Innovation Data, Third Edition*. Organisation for Economic Co-operation and Development, Eurostat.

OECD (2009). *Innovation in Firms: A Microeconomic Perspective*. Organisation for Economic Co-operation and Development.

OECD (2010a). *Measuring Innovation: A New Perspective*. Organisation for Economic Co-operation and Development. Original PDF is not available.

OECD (2010b). *The OECD innovation Strategy*. Organisation for Economic Co-operation and Development, Eurostat.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6).

Pakes, A. and Griliches, Z. (1984). *R&D, Patents, and Productivity*, chapter Patents and R&D at the Firm level: A First Look, pages 55–72. University of Chicago Press.

Parisi, M., Schiantarelli, F., and Sembenelli, A. (2006). Productivity, innovation and r&d: micro-evidence from italy. *European Economic Review*, 8(50):2037–2061.

Pavitt, K. (1988). *Handbook of Quantitative Studies of Science and Technology*, chapter Uses and abuses of patent statistics, pages 509–536. Elsevier Science Pub.

Peeters, C. and Pottelsberghe, de la Potterie, B. v. (2004). Innovation capabilities and firm labour productivity. Technical report, Workgroup CEB.

Peters, B. (2005). The relationship between product and process innovations and firm performance: Microeconometric evidence. *ZEW*.

Peters, B. (2006). Persistence of innovation: Stylised facts and panel data evidence. Working Paper 06-30, Danish Research Unit for Industrial Dynamics.

Peters, B. (2007). *Inovation and Firm Performance: An Empirical Investigation for German Firms*, volume 38. ZEW Economic Studies.

Peters, B., Roberts, M. J., Vuong, V. A., and Fryges, H. (2013). Firm r&d, innovation, and productivity in german industry. Technical report, ZEW, NBER, and EWI.

Polder, M., van Leeuwen, G., Mohnen, P., and Raymond, W. (2009). Productivity effects of innovation modes. Discussion Paper 09033, Statistics Netherlands.

Polder, M., van Leeuwen, G., Mohnen, P., and Raymond, W. (2010). Product, process and organizational innovation: drivers, complementarily and productivity effects. Working Paper 2010-035, United Nations University.

Raes, S., Gelauff, G., Klomp, L., and Roelandt, T. (2004). *Fostering Productivity*, chapter A Policy Agenda for Higher Productivity Growth, pages 305–321. Emerald Group Publishing Limited.

Rammer, C. and Peters, B. (2013). *Handbook of Innovation Indicators and Measurement*, chapter Innovation Panel Surveys in Germany, pages 135–177. Edward Elgar Publishing.

Rammer, C. and Schubert, T. (2016). Concentration on the few? r&d and innovation in german firms 2001 to 2013. Technical report, Centre for European Economic Research (ZEW).

Rao, S., Ahmad, A., Horsman, W., and Kaptein-Russell, P. (2001). The importance of innovation for productivity. Technical report, Industry Canada.

Raymond, W., Mairesse, J., Mohnen, P., and Palm, F. (2013). Dynamic models of r&d, innovation and productivity: Panel data evidence for dutch and french manufacturing. NBER Working Paper 19074, National Bureau of Economic Research (NBER).

Roberts, M. J. and Vuong, V. A. (2013). Empirical modeling of r&d demand in a dynamic framework. Technical report, The Pennsylvania State University and University of Cologne and Institute for Energy Economics.

Robin, S. and Schubert, T. (2010). Cooperation with public research institutions and success in innovation: Evidence from france and germany. Discussion Paper 24, Fraunhofer Institute for Systems and Innovation Research ISI.

Rogerson, P. A. (2001). *Statistical Methods for Geography*. Sage.

Romer, D. (2012). *Advanced Marcroeconomics. Forth Edition*. McGraw-Hill.

Romer, P. (1992). *Proceedings of the World Bank Annual Conference on Development Economics*, chapter Two Strategies for Economic Development: Using Ideas and Producing Ideas, pages 63–116. The World Bank.

Romer, P. (1993). Implementing a national technology strategy with self-organizing industry investment boards. Technical report, National Bureau of Economic Research (NBER).

Romer, P. M. (1990). Endogenous technological change. *The Journal of Political Economy*, 98(5).

Rose, S., Shipp, S., Lal, B., and Stone, A. (2009). Frameworks for measuring innovation: Initial approaches. Working Paper 06, Science and Technology Policy Institute.

Saunders, M., Lewis, P., and Thornhill, A. (2009). *Research methods for business students, 5th Edition*. 5th Edition. Prentice Hall, 5 edition.

Scherer, F. (1967). Market structure and the employment of scientists and engineers. *American Economic*, 57:524–531.

Scherer, F. and Ross, D. (1990). *Industrial market Structure and Economic Performance*. Houghton Mifflin.

Schmidt, T. (2005). Knowledge flows and r&d co-operation: Firm-level evidence from germany. *ZEW Discussion Papers*, 05-22.

Schneider, H. (2017). *Creative Destruction and the Sharing Economy: Uber As Disruptive Innovation (New Thinking in Political Economy series)*. Elgar.

Schneider, L. (2007). Alterung und technologisches innovationspotential. eine linked-employer-employee-analyse. Technical Report 2, IWH.

Schneider, R. (2013). Low productivity growth in germany. Working paper 166, Allianz.

Schramm, C. (2008). Innovation measurement. tracking the state of innovation in the american economy. Report, The Advisory Committee on Measuring Innovation in the 21st Century Economy.

Schubert, T. (2010). Market and organizsational innovation in entrepreneurial innovation processes and their relation to market structure and firm characteristics. *Review of Industrial Organization*, 32:189–212.

Schubert, T. and Andersson, M. (2013). Old is gold? the effects of employee age on innovation and the moderating effects of employment turnover. Technical report, Center for Innovation, Research and Competence in the Learning Economy (CIRCLE).

Schumpeter, J. (1934). *The Theory of Economic Development. An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, volume 46 of *Harvard Economic Studies*. Harvard University Press.

Schumpeter, J. (1942). *Capitalism, Socialism and Democracy*. Harper, New York/London.

Schumpeter, J. (1943). *Postwar economic problems*, chapter Capitalism in the postwar world. McGraw-Hill Book Company.

Seddighi, H. (2012). *Introductory Econometrics: A Practical Approach*. Routledge.

Sekaran, U. (2003). *Research Methods For Business, A Skill-Building Approach. Fourth Edition*. John Wiley & Sons, Inc., 4 edition.

Sengupta, J. (2011). *Understanding Economic Growth. Modern Theory and Experience*. Springer.

Sengupta, J. (2014). *Theory of Innovation. A New Paradigm of Growth*. Springer.

Shanks, S. and Zheng, S. (2006). Econometric modemodel of r&d and australia's productivity. Technical report, Australian Government, Productivity Commission.

Sinn, H. W. (2004). The dilemma of globalisation: A german perspective. *Economie internationale*, 4(100):111–120.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. William Strahan, Thomas Cadell.

Solow, R. (1987). We'd better watch out. *New York Times Book Review, New York Times*, 36.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 39:65–94.

Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics ans Statistics*, 39:312–320.

Soskice, D. and Hall, P. A. (2001). *Varieties of Capitalism*, chapter An Introduction to Varieties of Capitalism, pages 1–68. Oxford University Press.

Spiegel (2004). Globalization's toll: Goodbye 'made in germany'? Technical report, Spiegel-Online. Url: http://www.spiegel.de/international/globalization-s-toll-goodbye-made-in-germany-a-324921-2.html, Url-date: 12.01.2017.

Spiezia, V. (2011). Are ict users more innovative? an analysis of ict-enabled innovation in oecd firms. *OECD Journal: Economic Studies*.

Squicciarini, M. and Torsti, L. (2008). Going global: The challenges for knowledge-based economies. MPRA Paper 9663, Finnish Ministry of Employment and the Economy.

Stone, A., Rose, S., Lal, B., and Shipp, S. (2008). Measuring innovation and intangibles: A business perspective. Document D-3704, IDA, Science and Technology Policy Institute.

Stoneman, P. (1995). *Handbook of the Economics of Innovation and Technological Change*. Blackwell, Cambridge MA.

Supan, A. B. (2008). Capital's contribution to productivity and the nature of competition. Brookings papers, University of Mannheim, Center for Economic Policy Research.

Suryanarayana, T. and Mistry, P. (2016). *Principal Component Regression for Crop Yield Estimation*. Springer.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–365.

Thietart, R. A. e. a. (2001). *Doing Management Research*. SAGE Publications.

Van Ark, B., Jaeger, K., Manole, V., and Metz, A. (2009). Productivity, performance, and progress. germany in the international comparative perspective. Working paper, The Conference Board, Friedrich-Ebert-Stiftung.

Van Beveren, I. and Vandenbussche, H. (2009). Product and process innovation and the decision to export: Firm-level evidence for belgium. LICOS Discussion Papers 24709, LICOS - Centre for Institutions and Economic Performance, KU Leuven.

Van Leeuwen, G. and Farooqui, S. (2008). *Yearbook on Productivity 2008*, chapter ICT investment, ICT use and productivity, pages 9–38. Statistic Sweden.

Van Leeuwen, G. and Klomp, L. (2006). On the contribution of innovation to multi-factor productivity growth. *Economics of Innovation and New Technology*, 15(4-5):367–390.

Van Leeuwen, G. and Van der Wiel, H. (2003). Spillover effect of ict. Technical report, CPB Netherlands Bureau for Economic Policy Analysis.

Van Winden, P. and Reitsma, A. (2004). *Fostering Productivity*, volume 1, chapter Human capital and productivity, pages 115–140. ELSEVIER.

Vickery, G. and Wurzburg, G. (1998). The challenge of measuring and evaluating organisational change in enterprises. Technical report, Organisation For Economic Co-operation And Development (OECD).

Vieira, E., Vazquez-Rozas, E., and Niera, I. (2008). The innovation factor: An econometric model of productivity in european regions. *Regional and Sector Economic Studies*, 8(1).

Vogt, W. P. (2005). *Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sciences. Third Edition*. Sage Publications Binding.

Von Eye, A. and Schuster, C. (1998). *Regression Analysis for Social Sciences*. Academic Press.

Wagner, J. (2012). Exports, r&d and productivity: A test of the bustos-model with german enterprise data. Working Paper Series in Economics 244, University of Lüneburg.

Washington, S. P., Karlaftis, M. G., and Mannering, F. L. (2011). *Statistical and Econometric Methods for Transportation Data Analysis*. Chapman & Hall/CRC.

Weidmann, J. (2014). Demographic challenges in germany. Technical report, Deutsche Bundesbank. Speech, Url: https://www.bundesbank.de/Redaktion/EN/Reden/2014/2014_11_27_weidmann.html, Url-date: 12.01.2017.

Wilson, J. (2014). *Essentials of Business Research. A Guide to Doing Your Research Project. Second Edition*. SAGE Publications, 2 edition.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press Cambridge, MassachusThe , London, England.

Wooldridge, J. M. (2013). *Introductory EEconometric. A Modern Approach. Fifth Edition*. Cengange Learning.

Zemplinerova, A. and Hromadkova, E. (2012). Determinants of firm's innovation. *Prague Economic Papers*, 4:487–503.

ZEW (2009). Questionnaire sheet of community innovation survey 2009. Technical report, Center for European Economic Research (ZEW).

ZEW (2011). Questionnaire sheet of community innovation survey 2011. Technical report, Center for European Economic Research (ZEW).

Zikmund, W. G., Babin, J., Carr, J., and Griffin, M. (2009). *Business Research Methods, Eighth Edition*. South-Western College Pub, 8 edition.

Appendix A

Empirical Studies Based on CDM Framework

The following table A.1 includes an overview of the known empirical studies that used the CDM model (Crepon et al., 1998) as basis for investigating the relationship between innovation and productivity. As shown, the table contains information about proxies used for innovation input, innovation output, whether ICT has been considered in the model or not, the origin countries and the industry of used data, time period of the data, and the employed estimation approach.

Table A.1 Overview of empirical studies based on CDM model

Study	**Variables**	**Data**	**Methods**	**IN →OUT**	**OUT →P**
Baum et al. (2015)	IN: R&D expenditure; OUT: PRD sales; P: labour productivity	Sweden, manufacturing 2006-2012 panel	ML	positive, weak	positive

To be continued on next page.

Table A.1 - *Continued from previous page*

Study	**Variables**	**Data**	**Method**	**IN →OUT**	**OUT →P**
Crepon et al. (1998)	IN: R&D stock ; OUT: PRD sales and number of patents; P: labour productivity	France, manufacturing 1986-1990 cross section	ALS	positive	positive
Criscuolo and Haskel (2003)	IN: Innov. expenditure; OUT: PRD sales (new and market novelties); P: TFP growth	UK, manufacturing 1994-2000 cross section	OLS	positive	positive, weak
Griffith et al. (2006)	IN: R&D expenditure; OUT: PRD and PRC dummy; P: labour productivity	France, Germany, Spain, and UK, manufacturing 1998-2000 cross section	FIML, 2SLS	positive for PRD and PRC	positive PRD except Germany; positive PRC only France
Hall et al. (2012)	IN: ICT, R&D expenditure; OUT: PRD, PRC, and ORG dummy; P: labour productivity	Italy, manufacturing 1992-2003 panel	ML	R&D expenditure positive for PRD and PRD; ICT positive for ORG	PRD and PRC positive in combination w. ORG
Janz et al. (2004)	IN: R&D expenditure; OUT: PRD sales per employee, PRC; P: labour productivity	Germany and Sweden, manufacturing 1998-2000 cross section	FIML, 2SLS	positive for PRD and PRC	positive

To be continued on next page.

Table A.1 - *Continued from previous page*

Study	Variables	Data	Method	IN →OUT	OUT →P
Jefferson et al. (2002)	IN: Innov. expenditure; OUT: PRD sales; P: labour productivity	China, L and M-sized manufacturing 1997-1999 cross section	OLS, IV	positive	positive
Klomp and van Leeuwen (2001)	IN: R&D expenditure; OUT: PRD sales ; P: growth of sales	Dutch, manufacturing 1994-1996 cross section	OLS, FIML	positive, weak	positive
Lööf and Heshmati (2002c)	IN: R&D expenditure; OUT: PRD sales (new and market novelties); P: labour productivity	Sweden, manufacturing and service 1996-1998 cross section	FIML and 2SLS	positive (not for market novelties)	positive
Mairesse and Robin (2009)	IN: R&D intensity; OUT: PRD sales and PRC; P: labour productivity	France, manufacturing 1998-2002 cross section	ML, FIML	positive	PRD positive, PRC no impact
Mairesse et al. (2012)	IN: R&D expenditure; OUT: PRD sales per employee ; P: labour productivity	China, manufacturing 2005-2006 pooled	n.a.	positive	positive
Moreno and Huergo (2010)	IN: R&D intensity; OUT: PRD and PRC dummy; P: labour productivity	Spain, 1990-2005 panel	ML, IV	positive	positive

To be continued on next page.

Table A.1 - *Continued from previous page*

Study	Variables	Data	Method	IN →OUT	OUT →P
Nguyen and Martin (2010)	IN: ICT, R&D expenditure; OUT: PRD, PRC, and ORG; P: labour productivity	Luxembourg manufacturing and service 2004-2006 cross section	ML	Both inputs are positive for PRD, PRC and ORG	PRD, PRC, ORG positive
Parisi et al. (2006)	IN: R&D expenditure; OUT: PRD and PRC dummy; P: labour productivity growth	Italy, manufacturing 1992-1997 panel	IV	positive for PRD and PRC	PRD no impact, PRC positive
Peters (2007)	IN: Innov. expenditure; OUT: PRD and PRC dummy; P: labour productivity	German firms 2000-2003 pooled	ML, LS	positive for PRD, positive weak for PRC	PRD and PRC positive
Polder et al. (2009)	IN: ICT, R&D expenditure; OUT: PRD, PRC, and ORG; P: value added per employee	Dutch, manufacturing 2002-2006 panel	ML	PRD positive	PRD, PRC positive if combined with ORG
Raymond et al. (2013)	IN: R&D expenditure; OUT: PRD sales; P: labour productivity	Dutch and French manufacturing 1994-2004 panel	FIML	positive	positive

To be continued on next page.

Table A.1 - *Continued from previous page*

Study	Variables	Data	Method	IN →OUT	OUT →P
Roberts and Vuong (2013)	IN: R&D expenditure; OUT: PRD and PRC dummy; P: labour productivity	German manufacturing up to 2009 panel	dynamic model	positive for PRD and PRC	PRD and PRC positive
Van Leeuwen and Klomp (2006)	IN: R&D expenditure; OUT: PRD sales; P: value added	Dutch manufacturing 1994-1996 cross section	FIML	positive	no impact
Zemplinerova and Hromadkova (2012)	IN: Innov. expenditure; OUT: PRD sales; P: labour productivity	Czech, 2004-2006 panel	3SLS	PRD positive	PRD positive

IN: innovation input(s); OUT: innovation output(s); P: performance/productivity variable.
PRD: product innovation; PRC: process innovation; ORG: organisational innovation; □= impact.

Appendix B

Pooled Tests and Further Analyses

B.1 The Decision and Expenditure Stage

B.1.1 Additional Determinants

Table B.1 The 1st stage with innovation constraints (pooled)

	(1)		(2)	
	D		l_INE	
L.l_P	1.469***	(4.73)	-0.158***	(-3.53)
l_SIZE	1.461***	(11.59)	-0.0692***	(-3.49)
BRANCH	3.611***	(12.40)	0.738***	(11.49)
H_ECO_RISK	1.254	(1.79)	0.139**	(2.97)
H_HIG_COST	1.246	(1.67)	-0.122*	(-2.56)
H_INT_FUND	0.918	(-0.55)	0.0360	(0.67)
H_EXT_FUND	1.043	(0.28)	0.0419	(0.81)
H_ORG_PROB	1.677***	(3.33)	-0.0533	(-1.07)
H_INT_RESI	0.784	(-1.41)	0.00364	(0.06)
H_NQA_EMPL	1.000	(0.00)	0.0169	(0.39)
H_TEC_INFO	1.145	(0.79)	-0.0921	(-1.60)
H_MKT_INFO	1.352*	(1.97)	0.0840	(1.62)
H_ACC_CUST	0.807	(-1.89)	-0.00597	(-0.14)
H_LEG_INDS	1.164	(0.99)	-0.0489	(-0.94)
H_ADM_PROC	0.892	(-0.81)	0.167***	(3.38)
IMR_H_			0.247*	(2.23)
N	2668		1673	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.2 The 1st stage with public subsidies (pooled)

	(1)		(2)	
	D		l_INE	
L.l_P	1.595***	(5.35)	-0.276***	(-6.54)
l_SIZE	1.276***	(6.84)	-0.164***	(-10.85)
BRANCH	2.237***	(7.28)	0.485***	(10.84)
PUB_SUBS	244.6***	(22.45)	0.331*	(2.24)
IMR_PUB_			-0.112	(-1.38)
N	4058		1780	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.3 The 1st stage with protection measures (pooled)

	(1)		(2)	
	D		l_INE	
L.l_P	1.058	(0.82)	-0.290***	(-7.74)
l_SIZE	1.317***	(9.15)	-0.121***	(-7.78)
BRANCH	3.149***	(12.60)	0.522***	(10.51)
P_PATNT	5.285***	(11.32)	0.313***	(5.41)
P_REGDS	1.931***	(4.08)	-0.0336	(-0.67)
P_TRMKT	2.991***	(7.73)	0.118*	(2.29)
P_CPYRT	1.715**	(2.74)	0.00937	(0.17)
IMR_P_			0.0725	(0.95)
N	3712		2105	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.4 The 1st stage with market characteristics (pooled)

	(1)		(2)	
	D		l_INE	
L.l_P	1.032	(0.44)	-0.294***	(-7.97)
l_SIZE	1.422***	(11.85)	-0.0719***	(-3.88)
BRANCH	3.005***	(12.08)	0.531***	(9.72)
M_POS_THRE	0.850**	(-2.97)	-0.0266	(-0.99)
M_CMP_UNPR	1.091	(1.25)	-0.00148	(-0.05)
M_OUT_DATE	1.489***	(7.82)	0.215***	(7.94)
M_PRO_SUBS	0.974	(-0.54)	-0.109***	(-4.59)
M_DEM_UNFS	1.041	(0.60)	0.0275	(0.89)
M_FOR_PRES	1.068	(1.41)	-0.000597	(-0.03)
M_EXS	4.792***	(8.66)	0.479***	(5.43)
IMR_M_			0.247*	(2.51)
N	3655		2224	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.5 The 1st stage with co-partnership (pooled)

	(1)		(2)	
	D		l_INE	
L.l_P	1.252**	(2.81)	-0.245***	(-4.76)
l_SIZE	1.368***	(9.29)	-0.153***	(-6.76)
BRANCH	2.650***	(9.42)	0.293***	(4.50)
CD_GROUP	1	(.)	0	(.)
CD_CSTMR	19.69***	(4.71)	0.00334	(0.04)
CD_SUPLR	18.90***	(3.89)	-0.0956	(-0.88)
CD_COMPT	1	(.)	0	(.)
CD_CNSLT	110.4***	(4.65)	0.0130	(0.15)
CD_UNIVR	26.42***	(5.44)	0.145	(1.69)
CA_GROUP	16.83**	(2.69)	-0.0298	(-0.22)
CA_CSTMR	2.362	(0.69)	0.118	(0.73)
CA_SUPLR	6.431	(1.53)	0.00914	(0.04)
CA_COMPT	1	(.)	0	(.)
CA_CNSLT	1	(.)	0	(.)
CA_UNIVR	1	(.)	0	(.)
IMR_C_			-0.192**	(-2.66)
N	2794		1077	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.6 The 1st stage with source of information (pooled)

	(1)		(2)	
	D		l_INE	
L.l_P	1.093	(0.56)	-0.302***	(-7.39)
l_SIZE	1.068	(0.95)	-0.108***	(-7.32)
BRANCH	2.188***	(3.76)	0.442***	(10.50)
I_GROUP	3.372***	(12.88)	0.116***	(3.75)
I_CSTMR	2.159***	(7.10)	0.0591*	(2.38)
I_SUPLR	1.833***	(4.95)	-0.0506*	(-2.05)
I_COMPT	1.473**	(2.93)	0.0411	(1.59)
I_CNSLT	1.641**	(2.86)	-0.0260	(-0.94)
I_UNIVR	2.236***	(4.12)	0.100***	(3.48)
I_RDINS	0.782	(-1.07)	0.125***	(4.01)
IMR_I_			0.111	(1.59)
N	3020		1716	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

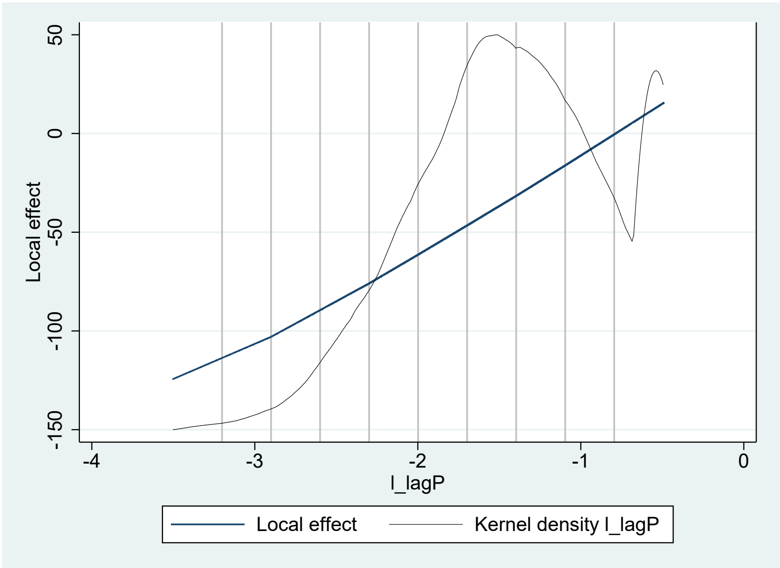


Fig. B.1 Non linearity test for labour productivity and innovation decision

B.1.2 Test Results

PS: The second rise in the kernel density is due to censoring.

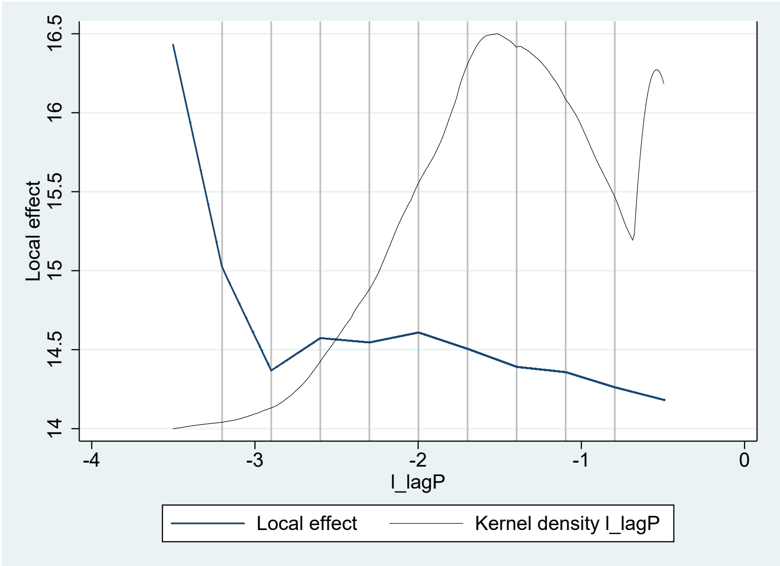


Fig. B.2 Non linearity test for labour productivity and innovation expenditure

PS: The second rise in the kernel density is due to censoring.

B.2 The Knowledge Production Stage

B.2.1 Additional Determinants

Table B.7 The 2nd stage product innovation with protection measures (pooled)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.259*	(2.01)	-0.0459	(-0.29)	0.0245	(0.17)
l_SIZE	-0.125**	(-2.91)	0.155**	(3.00)	-0.127**	(-2.71)
EMPL_UNI	0.0644***	(3.64)	-0.0594**	(-2.83)	0.0773***	(4.00)
IMR	-0.399**	(-3.07)	0.497**	(3.15)	-0.496***	(-3.41)
P_PATNT	0.134	(1.72)	-0.0340	(-0.37)	0.338***	(3.98)
P_REGDS	0.124	(1.50)	-0.148	(-1.52)	0.125	(1.43)
P_TRMKT	-0.0902	(-1.20)	0.169	(1.86)	0.141	(1.73)
P_CPYRT	0.261**	(2.87)	-0.267*	(-2.52)	0.0397	(0.41)
N	972		974		961	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.8 The 2nd stage product innovation with market characteristics (pooled)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.287*	(2.26)	-0.0419	(-0.27)	-0.135	(-0.96)
l_SIZE	-0.0443	(-1.06)	0.102*	(2.05)	-0.0390	(-0.85)
EMPL_UNI	0.0571***	(3.46)	-0.0451*	(-2.32)	0.0798***	(4.35)
IMR	-0.208	(-1.62)	0.277	(1.78)	-0.238	(-1.63)
M_POS_THRE	0.0630	(1.36)	-0.0598	(-1.07)	-0.0553	(-1.08)
M_CMP_UNPR	-0.130*	(-2.26)	0.106	(1.56)	-0.0172	(-0.27)
M_OUT_DATE	0.315***	(7.92)	-0.308***	(-6.61)	0.184***	(4.24)
M_PRO_SUBS	-0.191***	(-4.47)	0.195***	(3.75)	-0.178***	(-3.73)
M_DEM_UNFS	0.0229	(0.41)	-0.0428	(-0.64)	-0.0838	(-1.32)
M_FOR_PRES	0.0333	(0.81)	-0.00453	(-0.09)	0.0422	(0.93)
M_EXS	0.138	(1.08)	-0.0659	(-0.44)	0.356*	(2.56)
N	1088		1089		1068	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.9 The 2nd stage product innovation with co-partnership (pooled)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.491**	(3.06)	-0.254	(-1.26)	-0.234	(-1.32)
l_SIZE	-0.0761	(-1.43)	0.115	(1.77)	-0.120*	(-2.07)
EMPL_UNI	0.0694**	(3.21)	-0.0673*	(-2.54)	0.0822***	(3.44)
IMR	-0.257	(-1.65)	0.360	(1.85)	-0.455**	(-2.58)
CD_GROUP	0.192	(1.52)	-0.154	(-1.01)	0.201	(1.49)
CD_CSTMR	-0.0761	(-0.71)	0.157	(1.21)	-0.0587	(-0.50)
CD_SUPLR	-0.00942	(-0.07)	0.172	(1.08)	-0.242	(-1.74)
CD_COMPT	0.179	(1.52)	-0.228	(-1.62)	0.0338	(0.26)
CD_CNSLT	0.0942	(0.98)	-0.0590	(-0.50)	0.121	(1.17)
CD_UNIVR	0.0787	(0.77)	-0.148	(-1.22)	0.224*	(2.01)
CA_GROUP	-0.0543	(-0.37)	0.179	(0.95)	0.111	(0.70)
CA_CSTMR	0.114	(0.72)	-0.0785	(-0.41)	-0.0487	(-0.29)
CA_SUPLR	-0.374	(-1.65)	0.143	(0.50)	-0.117	(-0.49)
CA_COMPT	0.374	(1.15)	-0.621	(-1.72)	0.384	(1.22)
CA_CNSLT	0.231	(1.14)	-0.142	(-0.60)	0.215	(1.02)
CA_UNIVR	-0.251	(-0.91)	0.634	(1.83)	-0.0819	(-0.29)
N	709		711		698	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.10 The 2nd stage product innovation with source of information (pooled)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.161	(1.05)	0.0112	(0.06)	-0.217	(-1.28)
l_SIZE	-0.108*	(-2.18)	0.156**	(2.58)	-0.0634	(-1.16)
EMPL_UNI	0.0743***	(3.88)	-0.0745***	(-3.29)	0.0959***	(4.47)
IMR	-0.494**	(-3.21)	0.588**	(3.06)	-0.288	(-1.65)
I_GROUP	0.130*	(2.40)	-0.0886	(-1.34)	0.140*	(2.25)
I_CSTMR	0.0197	(0.45)	0.0376	(0.71)	0.103*	(2.06)
I_SUPLR	0.0444	(0.99)	-0.0584	(-1.09)	0.0242	(0.48)
I_COMPT	0.00719	(0.16)	-0.0308	(-0.56)	-0.185***	(-3.56)
I_CNSLT	0.0253	(0.50)	-0.0156	(-0.26)	0.0105	(0.19)
I_UNIVR	0.0229	(0.45)	0.0593	(0.97)	0.123*	(2.16)
I_RDINS	0.0488	(0.91)	-0.101	(-1.59)	0.0626	(1.07)
N	816		817		805	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.11 The 2nd stage product innovation with public subsidies (pooled)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.468**	(3.23)	-0.230	(-1.32)	-0.0296	(-0.19)
l_SIZE	-0.0503	(-1.12)	0.0783	(1.48)	-0.0477	(-0.99)
EMPL_UNI	0.0862***	(4.46)	-0.0972***	(-4.22)	0.0930***	(4.43)
IMR	-0.194	(-1.37)	0.293	(1.72)	-0.332*	(-2.13)
PUB_SUBS	0.268**	(3.08)	-0.263*	(-2.43)	0.250**	(2.60)
N	850		852		833	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.12 The 2nd stage process innovation with protection measures (pooled)

	(1) PRC_COST		(2) PRC_QUAL	
L.l_INEp	0.159	(0.95)	0.326	(1.81)
l_SIZE	0.0126	(0.22)	-0.0295	(-0.47)
EMPL_UNI	-0.0186	(-0.75)	0.0143	(0.53)
IMR	-0.252	(-1.37)	0.273	(1.42)
P_PATNT	0.241*	(2.18)	0.00303	(0.03)
P_REGDS	0.161	(1.46)	0.168	(1.40)
P_TRMKT	-0.0456	(-0.43)	-0.0283	(-0.24)
P_CPYRT	-0.00559	(-0.05)	0.103	(0.77)
N	631		587	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.13 The 2nd stage process innovation with market characteristics (pooled)

	(1) PRC_COST		(2) PRC_QUAL	
L.l_INEp	0.179	(1.06)	0.264	(1.45)
l_SIZE	0.00493	(0.09)	-0.0230	(-0.37)
EMPL_UNI	-0.0281	(-1.22)	0.0136	(0.55)
IMR	-0.209	(-1.12)	0.289	(1.49)
M_POS_THRE	-0.0326	(-0.54)	-0.0389	(-0.58)
M_CMP_UNPR	-0.0695	(-0.97)	-0.0862	(-1.10)
M_OUT_DATE	0.101	(1.95)	0.199***	(3.57)
M_PRO_SUBS	0.0411	(0.74)	-0.137*	(-2.30)
M_DEM_UNFS	-0.0353	(-0.48)	0.0118	(0.15)
M_FOR_PRES	0.0885	(1.65)	-0.0716	(-1.24)
M_EXS	0.0802	(0.48)	-0.199	(-1.08)
N	706		651	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.14 The 2nd stage process innovation with co-partnership (pooled)

	(1)		(2)	
	PRC_COST		PRC_QUAL	
L.l_INEp	0.0582	(0.27)	0.436	(1.84)
l_SIZE	-0.00325	(-0.04)	0.000815	(0.01)
EMPL_UNI	-0.0145	(-0.47)	0.0150	(0.44)
IMR	-0.244	(-1.05)	0.354	(1.44)
CD_GROUP	-0.00676	(-0.04)	0.0880	(0.51)
CD_CSTMR	-0.00964	(-0.07)	0.343*	(2.27)
CD_SUPLR	-0.197	(-1.08)	-0.359	(-1.80)
CD_COMPT	0.0826	(0.54)	0.214	(1.31)
CD_CNSLT	0.0606	(0.46)	0.0462	(0.32)
CD_UNIVR	-0.178	(-1.18)	-0.0105	(-0.06)
CA_GROUP	-0.0238	(-0.13)	-0.109	(-0.51)
CA_CSTMR	0.0214	(0.10)	-0.125	(-0.51)
CA_SUPLR	0.326	(1.15)	-0.243	(-0.73)
CA_COMPT	0.346	(0.96)	0.299	(0.76)
CA_CNSLT	-0.147	(-0.50)	0.104	(0.32)
CA_UNIVR	0.128	(0.39)	0.656	(1.91)
N	452		410	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.15 The 2nd stage process innovation with source of information (pooled)

	(1)		(2)	
	PRC_COST		PRC_QUAL	
L.l_INEp	0.259	(1.33)	0.267	(1.21)
l_SIZE	-0.000885	(-0.01)	-0.0672	(-0.90)
EMPL_UNI	-0.0464	(-1.75)	0.0477	(1.66)
IMR	-0.0847	(-0.40)	0.395	(1.73)
I_GROUP	0.0632	(0.96)	0.0218	(0.31)
I_CSTMR	0.133*	(2.35)	0.114	(1.82)
I_SUPLR	-0.00572	(-0.10)	0.146*	(2.25)
I_COMPT	0.0213	(0.35)	-0.0103	(-0.15)
I_CNSLT	-0.0680	(-1.07)	-0.111	(-1.59)
I_UNIVR	0.0693	(0.99)	-0.0152	(-0.20)
I_RDINS	0.0331	(0.42)	0.160	(1.85)
N	535		486	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.16 The 2nd stage process innovation with public subsidies (pooled)

	(1) PRC_COST		(2) PRC_QUAL	
L.l_INEp	0.0101	(0.05)	0.335	(1.58)
l_SIZE	-0.0292	(-0.47)	-0.0740	(-1.06)
EMPL_UNI	-0.0141	(-0.52)	-0.0194	(-0.67)
IMR	-0.266	(-1.30)	0.347	(1.60)
PUB_SUBS	-0.0737	(-0.66)	0.442***	(3.48)
N	529		474	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.17 The 2nd stage organisational innovation with protection measures (pooled)

	(1) ORG_TIME		(2) ORG_QUAL		(3) ORG_COST	
L.l_INEp	-0.191	(-0.57)	-0.554	(-1.68)	-0.0200	(-0.06)
l_SIZE	0.0658	(0.63)	-0.102	(-1.00)	0.146	(1.48)
EMPL_UNI	-0.0268	(-0.64)	0.0242	(0.58)	0.0109	(0.27)
IMR	0.326	(0.94)	0.154	(0.45)	0.517	(1.55)
P_PATNT	0.388	(1.92)	0.199	(1.00)	-0.146	(-0.76)
P_REGDS	-0.444*	(-2.20)	-0.372	(-1.86)	0.175	(0.91)
P_TRMKT	0.357	(1.79)	0.558**	(2.76)	0.0287	(0.15)
P_CPYRT	-0.0146	(-0.06)	-0.0908	(-0.36)	-0.132	(-0.55)
N	207		208		208	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.18 The 2nd stage organisational innovation with market characteristics (pooled)

	(1) ORG_TIME		(2) ORG_QUAL		(3) ORG_COST	
L.l_INEp	-0.213	(-0.90)	-0.438	(-1.87)	-0.00175	(-0.01)
l_SIZE	0.0791	(1.01)	0.0228	(0.29)	0.143	(1.88)
EMPL_UNI	0.0220	(0.78)	0.0209	(0.75)	0.000738	(0.03)
IMR	0.100	(0.40)	0.140	(0.55)	0.369	(1.50)
M_POS_THRE	0.0606	(0.83)	0.112	(1.53)	0.0404	(0.57)
M_CMP_UNPR	0.0239	(0.31)	-0.0186	(-0.24)	-0.00573	(-0.08)
M_OUT_DATE	0.183**	(2.75)	0.111	(1.67)	0.0652	(1.02)
M_PRO_SUBS	-0.0424	(-0.52)	-0.139	(-1.70)	0.106	(1.33)
M_DEM_UNFS	-0.00460	(-0.05)	0.0754	(0.87)	-0.160	(-1.89)
M_FOR_PRES	-0.0679	(-1.04)	-0.115	(-1.74)	-0.0296	(-0.46)
M_EXS	-0.0470	(-0.22)	-0.352	(-1.63)	-0.0858	(-0.40)
N	455		455		453	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.19 The 2nd stage organisational innovation with co-partnership (pooled)

	(1) ORG_TIME		(2) ORG_QUAL		(3) ORG_COST	
L.l_INEp	-0.319	(-0.87)	-0.543	(-1.50)	-0.149	(-0.42)
l_SIZE	-0.00635	(-0.05)	-0.0318	(-0.28)	0.101	(0.89)
EMPL_UNI	-0.0455	(-1.06)	-0.0394	(-0.93)	-0.0183	(-0.44)
IMR	0.118	(0.31)	0.121	(0.33)	0.502	(1.37)
CD_GROUP	-0.453	(-1.49)	-0.135	(-0.45)	0.0117	(0.04)
CD_CSTMR	-0.421	(-1.85)	-0.0985	(-0.45)	-0.327	(-1.50)
CD_SUPLR	0.269	(0.78)	0.117	(0.35)	-0.0936	(-0.28)
CD_COMPT	0.393	(1.67)	0.0448	(0.20)	0.0842	(0.37)
CD_CNSLT	-0.142	(-0.68)	-0.187	(-0.91)	-0.459^{*}	(-2.23)
CD_UNIVR	0.398	(1.45)	0.370	(1.37)	0.544^{*}	(1.99)
CA_GROUP	0.478	(1.45)	0.190	(0.61)	0.453	(1.41)
CA_CSTMR	0.424	(1.29)	-0.0983	(-0.32)	0.328	(1.03)
CA_SUPLR	0.331	(0.63)	-0.456	(-0.93)	0.142	(0.29)
CA_COMPT	0.355	(0.45)	-0.750	(-0.95)	0.278	(0.36)
CA_CNSLT	-0.203	(-0.51)	0.110	(0.29)	-0.0165	(-0.04)
CA_UNIVR	-0.599	(-0.62)	1.216	(1.25)	-0.223	(-0.24)
N	236		237		233	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.20 The 2nd stage organisational innovation with source of information (pooled)

	(1) ORG_TIME		(2) ORG_QUAL		(3) ORG_COST	
L.l_INEp	-0.0869	(-0.39)	-0.357	(-1.62)	0.0131	(0.06)
l_SIZE	0.0743	(1.03)	-0.0147	(-0.21)	0.113	(1.63)
EMPL_UNI	0.0299	(1.08)	0.0129	(0.47)	-0.00532	(-0.20)
IMR	0.167	(0.72)	0.204	(0.90)	0.421	(1.89)
I_GROUP	0.140	(1.79)	0.126	(1.62)	0.202^{**}	(2.64)
I_CSTMR	0.139^{*}	(1.98)	0.0709	(1.01)	-0.0271	(-0.39)
I_SUPLR	0.236^{***}	(3.57)	0.190^{**}	(2.92)	0.132^{*}	(2.06)
I_COMPT	0.0247	(0.38)	0.0459	(0.71)	0.117	(1.84)
I_CNSLT	0.0456	(0.63)	0.0245	(0.34)	0.107	(1.52)
I_UNIVR	-0.118	(-1.57)	-0.0482	(-0.65)	-0.0330	(-0.45)
I_RDINS	0.0138	(0.17)	0.0588	(0.74)	-0.0345	(-0.44)
N	486		487		484	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.21 The 2nd stage organisational innovation with public subsidies (pooled)

	(1)		(2)		(3)	
	ORG_TIME		ORG_QUAL		ORG_COST	
L.l_INEp	0.244	(0.73)	-0.105	(-0.32)	0.114	(0.35)
l_SIZE	0.262**	(2.60)	0.0524	(0.53)	0.119	(1.23)
EMPL_UNI	0.0372	(0.93)	0.0250	(0.63)	-0.0360	(-0.93)
IMR	0.269	(0.77)	0.249	(0.71)	0.373	(1.10)
PUB_SUBS	-0.320	(-1.14)	-0.0987	(-0.35)	-0.397	(-1.43)
N	217		217		214	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B.2.2 Test Results

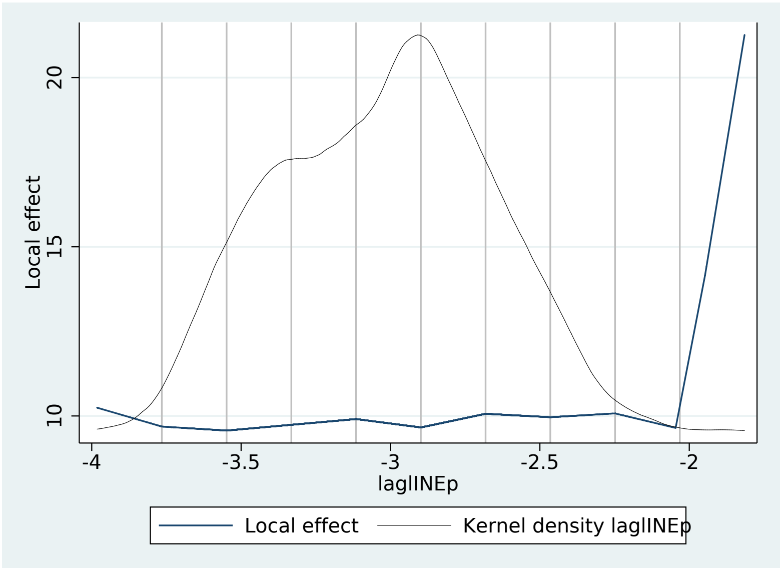


Fig. B.3 Non linearity test for expenditure and product innovation PRD_IMPR

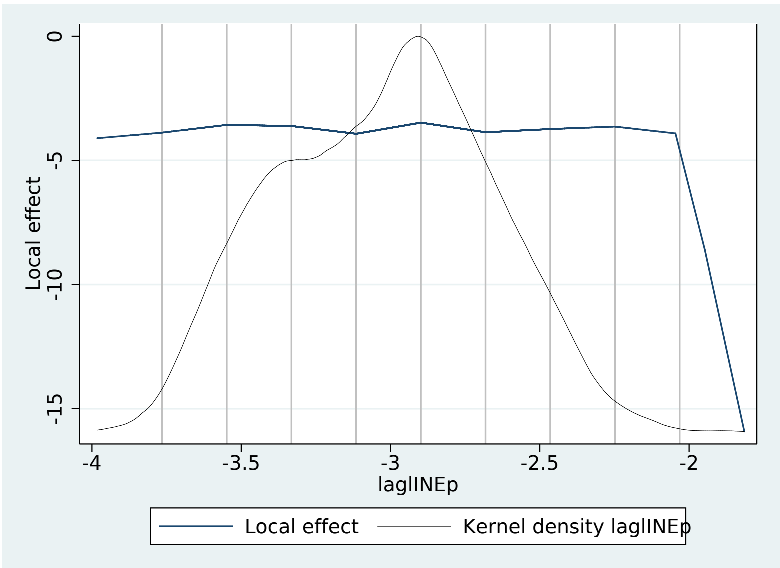


Fig. B.4 Non linearity test for expenditure and product innovation PRD_NCHG

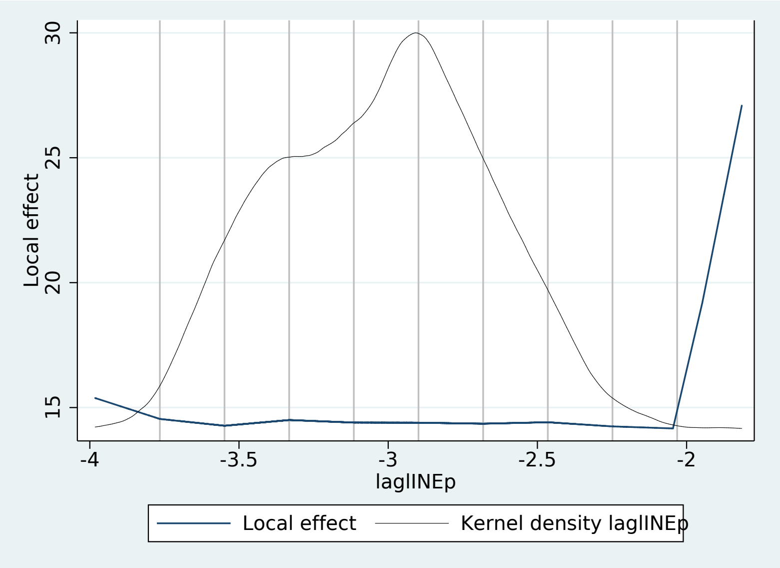


Fig. B.5 Non linearity test for expenditure and product innovation PRD_MNOV

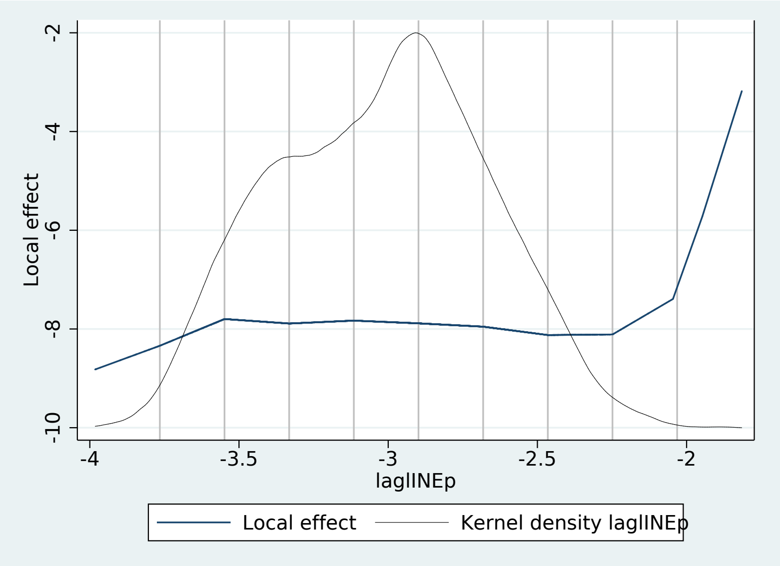


Fig. B.6 Non linearity test for expenditure and process innovation PRC_COST

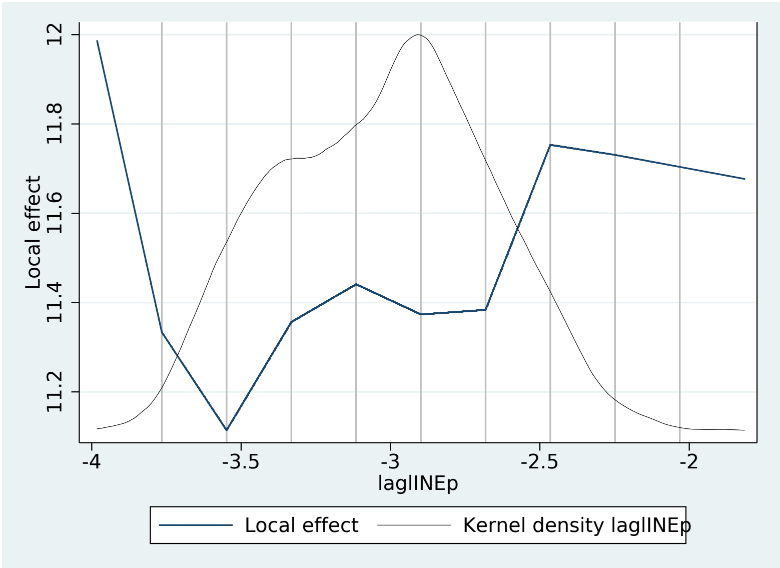


Fig. B.7 Non linearity test for expenditure and process innovation PRC_QUAL

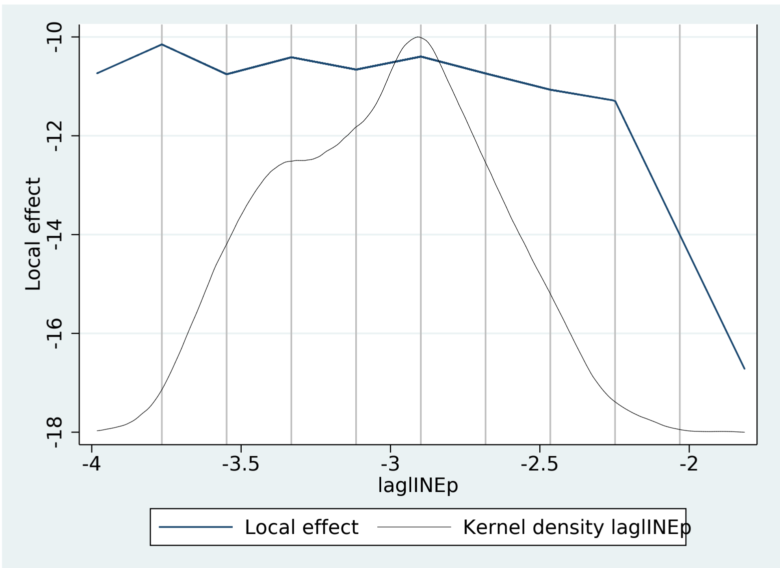


Fig. B.8 Non linearity test for expenditure and organisational innovation ORG_TIME

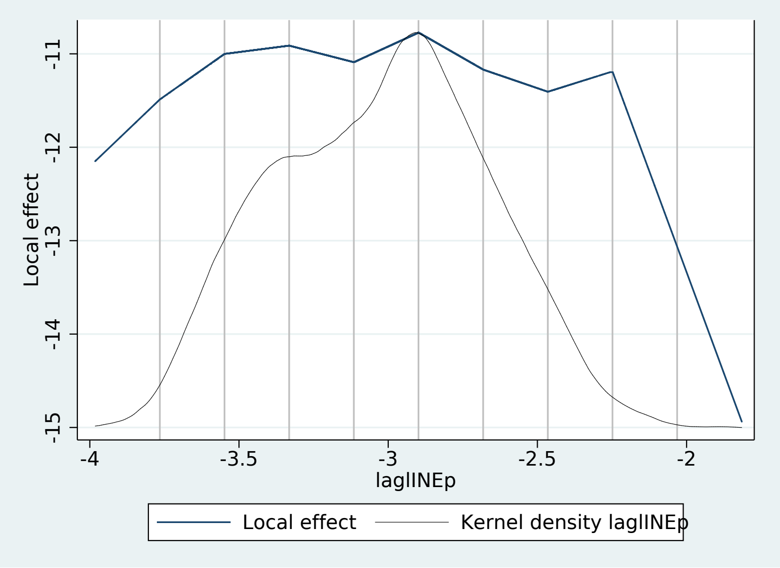


Fig. B.9 Non linearity test for expenditure and organisational innovation ORG_COST

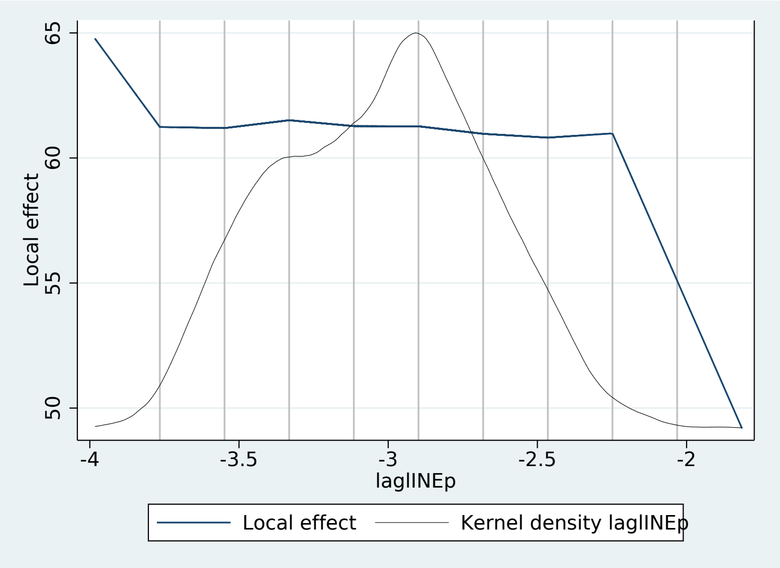


Fig. B.10 Non linearity test for expenditure and organisational innovation ORG_QUAL

B.3 The Production Stage

B.3.1 Additional Determinants

Table B.22 The 3rd stage with protection measures (pooled)

	(1)	
	l_P	
L.pc1	0.00149	(0.08)
L.pc2	0.739***	(16.38)
L.pc3	-0.272***	(-7.32)
l_SIZE	-0.610***	(-15.20)
l_INVS	-0.0301	(-1.67)
IMR	-0.515***	(-4.18)
P_PATNT	0.0566	(1.54)
P_REGDS	-0.0237	(-0.57)
P_TRMKT	0.0394	(1.09)
P_CPYRT	-0.0351	(-0.82)
N	713	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.23 The 3rd stage with market characteristics (pooled)

	(1)	
	l_P	
L.pc1	-0.000939	(-0.05)
L.pc2	0.719***	(16.66)
L.pc3	-0.254***	(-7.13)
l_SIZE	-0.602***	(-15.52)
l_INVS	-0.0438*	(-2.46)
IMR	-0.444***	(-3.82)
M_POS_THRE	-0.00118	(-0.05)
M_CMP_UNPR	0.00148	(0.05)
M_OUT_DATE	-0.0102	(-0.53)
M_PRO_SUBS	0.0114	(0.58)
M_DEM_UNFS	0.0105	(0.33)
M_FOR_PRES	-0.0355	(-1.78)
M_EXS	0.502***	(7.91)
N	719	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.24 The 3rd stage with co-partnership (pooled)

	(1)	
	l_P	
L.pc1	0.0183	(0.77)
L.pc2	0.802***	(14.32)
L.pc3	-0.326***	(-6.87)
l_SIZE	-0.658***	(-13.08)
l_INVS	-0.0580**	(-2.71)
IMR	-0.359*	(-2.40)
CD_GROUP	0.00788	(0.14)
CD_CSTMR	-0.0341	(-0.69)
CD_SUPLR	0.0589	(1.09)
CD_COMPT	-0.135**	(-2.62)
CD_CNSLT	0.0353	(0.83)
CD_UNIVR	-0.0302	(-0.68)
CA_GROUP	0.196**	(2.72)
CA_CSTMR	0.0635	(0.80)
CA_SUPLR	0.00188	(0.02)
CA_COMPT	0.235	(1.51)
CA_CNSLT	0.150	(1.58)
CA_UNIVR	-0.182	(-1.48)
N	473	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.25 The 3rd stage with public subsidies (pooled)

	(1)	
	l_P	
L.pc1	0.0221	(1.04)
L.pc2	0.733***	(14.60)
L.pc3	-0.299***	(-7.08)
l_SIZE	-0.533***	(-11.55)
l_INVS	-0.0417*	(-2.05)
IMR	-0.334*	(-2.47)
PUB_SUBS	-0.115**	(-2.94)
N	560	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B.3.2 Test Results

Table B.26 Estimation results of 3rd stage combination of innovation types (pooled)

	(1)	
	l_P	
L.pc1	0.0613	(0.89)
L.pc2	0.751***	(5.93)
L.pc3	-0.316**	(-2.82)
l_SIZE	-0.556***	(-5.43)
l_INVS	-0.109*	(-2.28)
IMR	-0.291	(-0.68)
L.CPRD_PRC_COST	-0.0994	(-1.34)
L.CPRD_PRC_QUAL	0.103	(1.42)
L.CPRD_ORG_TIME	-0.113	(-0.78)
L.CPRD_ORG_QUAL	-0.135	(-0.68)
L.CPRD_ORG_COST	0.298	(1.73)
N	142	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.27 Correlation among explanatory variables used in 3rd stage (pooled)

	PRD_IMPRp	PRD_NCHGp	PRD_MNOVp	PRC_COSTp	PRC_QUALp	ORG_TIMEp	ORG_QUALp	ORG_COSTp
PRD_IMPRp	1							
PRD_NCHGp	-0.982***	1						
PRD_MNOVp	0.895***	-0.933***	1					
PRC_COSTp	0.764***	-0.776***	0.798***	1				
PRC_QUALp	0.174***	-0.122***	-0.186***	-0.393***	1			
ORG_TIMEp	-0.0744***	0.0137	0.345***	0.129***	-0.771***	1		
ORG_QUALp	-0.533***	0.409***	-0.232***	-0.606***	-0.117***	0.491***	1	
ORG_COSTp	-0.705***	0.720***	-0.469***	-0.638***	-0.357***	0.604***	0.611***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.28 PCA results on explanatory variables of 3rd stage (pooled)

	Comp1	Comp2	Comp3	Comp4
PRD_IMPRp	0.4563	0.0243	0.2239	0.2984
PRD_NCHGp	-0.4521	-0.0700	-0.2953	-0.0054
PRD_MNOVp	0.4014	0.2985	0.2854	0.0618
PRC_COSTp	0.4151	0.2184	-0.3455	-0.2084
PRC_QUALp	0.0382	-0.5623	0.5049	0.2691
ORG_TIMEp	-0.0766	0.6370	0.1327	0.1936
ORG_QUALp	-0.3059	0.2328	0.6204	-0.5843
ORG_COSTp	-0.3911	0.2856	0.0668	0.6421
Eigenvalue	4.4445	2.3319	0.9188	0.3048
Proportion	0.5556	0.2915	0.1149	0.0381
Cumulative	0.5556	0.8470	0.9619	1.0000

Appendix C

Panel Tests and Further Analyses

C.1 The Decision and Expenditure Stage

C.1.1 Additional Determinants

Table C.1 The 1st stage with innovation constraints (panel)

	(1)		(2)	
	D		l_INE	
L.l_P	1.915***	(4.00)	-0.129**	(-3.03)
l_SIZE	2.054***	(8.57)	-0.0722***	(-4.16)
BRANCH	9.882***	(9.01)	0.584***	(11.60)
H_ECO_RISK	1.123	(0.53)	0.0771*	(1.96)
H_HIG_COST	1.609*	(2.09)	-0.0755	(-1.86)
H_INT_FUND	0.940	(-0.23)	0.00130	(0.03)
H_EXT_FUND	0.963	(-0.15)	0.0272	(0.60)
H_ORG_PROB	2.177**	(2.90)	-0.0326	(-0.79)
H_INT_RESI	0.689	(-1.25)	0.00573	(0.12)
H_NQA_EMPL	0.963	(-0.17)	0.0335	(0.88)
H_TEC_INFO	1.221	(0.67)	-0.0317	(-0.64)
H_MKT_INFO	1.744*	(2.09)	0.0176	(0.39)
H_ACC_CUST	0.759	(-1.43)	0.0317	(0.84)
H_LEG_INDS	1.257	(0.87)	-0.0638	(-1.42)
H_ADM_PROC	0.830	(-0.77)	0.149***	(3.54)
IMR_H_			0.0678	(1.24)
N	2668		1628	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.2 The 1st stage with public subsidies (panel)

	(1) D		(2) l_INE	
L.l_P	2.199***	(4.40)	-0.204***	(-4.79)
l_SIZE	1.662***	(6.17)	-0.136***	(-8.16)
BRANCH	4.389***	(6.00)	0.370***	(8.10)
PUB_SUBS	14622.7***	(12.05)	0.249*	(2.43)
IMR_PUB_			-0.0754*	(-2.01)
N	4058		1719	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3 The 1st stage with protection measures (panel)

	(1) D		(2) l_INE	
L.l_P	1.037	(0.25)	-0.218***	(-5.84)
l_SIZE	1.798***	(7.95)	-0.0903***	(-5.69)
BRANCH	9.830***	(9.85)	0.508***	(10.92)
P_PATNT	17.08***	(9.21)	0.210***	(4.51)
P_REGDS	2.727***	(3.43)	-0.00448	(-0.10)
P_TRMKT	6.335***	(6.57)	0.104*	(2.33)
P_CPYRT	2.389*	(2.50)	0.0157	(0.32)
IMR_P_			0.0679	(1.78)
N	3712		2047	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4 The 1st stage with market characteristics (panel)

	(1) D		(2) l_INE	
L.l_P	0.976	(-0.15)	-0.249***	(-6.62)
l_SIZE	2.203***	(9.47)	-0.0768***	(-4.56)
BRANCH	10.63***	(9.58)	0.430***	(9.64)
M_POS_THRE	0.747*	(-2.55)	-0.00267	(-0.11)
M_CMP_UNPR	1.162	(1.07)	-0.00819	(-0.29)
M_OUT_DATE	2.106***	(6.64)	0.133***	(5.80)
M_PRO_SUBS	0.957	(-0.42)	-0.0697**	(-3.14)
M_DEM_UNFS	1.125	(0.89)	0.0277	(1.04)
M_FOR_PRES	1.099	(1.00)	-0.00563	(-0.27)
M_EXS	27.45***	(7.44)	0.330***	(4.10)
IMR_M_			0.0356	(0.86)
N	3655		2178	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.5 The 1st stage with co-partnership (panel)

	(1)		(2)	
	D		l_INE	
L.l_P	1.422*	(2.23)	-0.242***	(-5.39)
l_SIZE	1.883***	(7.21)	-0.130***	(-6.80)
BRANCH	6.184***	(7.11)	0.340***	(6.44)
CD_GROUP	1	(.)	0.0753	(1.08)
CD_CSTMR	104.0***	(4.58)	-0.0491	(-0.83)
CD_SUPLR	66.62***	(3.74)	0.0281	(0.44)
CD_COMPT	1	(.)	0.119	(1.95)
CD_CNSLT	2305.7***	(4.89)	0.00565	(0.11)
CD_UNIVR	222.4***	(5.39)	0.0874	(1.63)
CA_GROUP	49.04*	(2.53)	0.0650	(0.80)
CA_CSTMR	1.073	(0.04)	0.188*	(2.19)
CA_SUPLR	6.440	(1.03)	0.109	(0.96)
CA_COMPT	1	(.)	-0.0377	(-0.23)
CA_CNSLT	1	(.)	0.351**	(3.15)
CA_UNIVR	1	(.)	0.0436	(0.33)
IMR_C_			-0.0977**	(-2.99)
N	2794		1374	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.6 The 1st stage with source of information (panel)

	(1)		(2)	
	D		l_INE	
L.l_P	1.093	(0.56)	-0.274***	(-6.62)
l_SIZE	1.068	(0.95)	-0.0877***	(-5.76)
BRANCH	2.188***	(3.76)	0.425***	(9.89)
I_GROUP	3.372***	(12.88)	0.0871**	(3.03)
I_CSTMR	2.159***	(7.10)	0.0524*	(2.29)
I_SUPLR	1.833***	(4.95)	-0.0202	(-0.86)
I_COMPT	1.473**	(2.93)	0.00678	(0.28)
I_CNSLT	1.641**	(2.86)	-0.0168	(-0.64)
I_UNIVR	2.236***	(4.12)	0.0860**	(3.21)
I_RDINS	0.782	(-1.07)	0.0837**	(2.86)
IMR_I_			0.0970	(1.46)
N	3020		1690	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C.1.2 Test Results

Table C.7 Hausman test for the participation equation

	(1) D
N	3133
h_chi2	124.5
h_p	8.30e-27
h_df	3
h_rank	3
bic	2342.6
aic	2324.5

Table C.8 LR test between pooled and RE Logit model

	D	
	panel RE	pooled
N	11280	11280
Log-Likelihood	-5205.877	-6573.56
AIC	10421.75	13155.12
BIC	10458.41	13184.44
LR-Test (Chi2, df=1)	2735.36	
LR-Test P-Value	0	

Table C.9 Robustness check (censoring) for expenditure equation

	(1) l_INE	
L.l_P	-0.131***	(-4.67)
l_SIZE	-0.0539***	(-3.90)
BRANCH	0.445***	(12.69)
IMR	0.0492	(1.46)
INEX	1.412***	(24.90)
INVSX	0.456	(1.43)
L.PX	-0.0456	(-1.13)
N	6179	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.10 LR test for 2nd equation between pooled and panel Tobit model

	l_INE	
	panel RE	pooled
N	6486	6486
Log-Likelihood	-7363.67	-8378.878
AIC	14741.34	16769.76
BIC	14788.78	16810.42
LR-Test (Chi2, df=1)	2030.4	
LR-Test P-Value	0	

Table C.11 Compare for 2nd equation between pooled and panel Tobit model

	(1)		(2)	
	Tobit(panel)		Tobit(pooled)	
main				
L.l_P	-0.1615***	(0.0267)	-0.2475***	(0.0224)
l_SIZE	-0.0763***	(0.0143)	-0.0691***	(0.0104)
BRANCH	0.4810***	(0.0370)	0.6321***	(0.0288)
IMR	0.0435	(0.0349)	0.0761**	(0.0280)
N	6486		6486	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.12 Multicollinearity test for 1st stage

Variable	VIF	Tolerance	R-Squared
lagLogP	1.27	.79	.21
l_SIZE	3.06	.327	.673
BRANCH	1.77	.565	.435
IMR	4.32	.231	.769

C.2 The Knowledge Production Stage

C.2.1 Additional Determinants

Table C.14 The 2nd stage product innovation with protection measures (panel)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.685**	(2.83)	-0.377	(-1.21)	0.350	(1.39)
l_SIZE	-0.159**	(-2.59)	0.211**	(2.67)	-0.0929	(-1.48)
EMPL_UNI	0.0882***	(3.30)	-0.0847*	(-2.51)	0.116***	(4.02)
IMR	-0.335**	(-2.98)	0.462**	(3.07)	-0.351**	(-2.84)
P_PATNT	0.226*	(2.03)	-0.0651	(-0.46)	0.421***	(3.56)
P_REGDS	0.188	(1.66)	-0.261	(-1.80)	0.197	(1.66)
P_TRMKT	-0.106	(-1.01)	0.214	(1.59)	0.174	(1.56)
P_CPYRT	0.286*	(2.31)	-0.315*	(-2.06)	0.0221	(0.17)
N	972		974		961	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.15 The 2nd stage product innovation with market characteristics (panel)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.843***	(3.57)	-0.387	(-1.32)	-0.0179	(-0.07)
l_SIZE	-0.0118	(-0.20)	0.124	(1.72)	-0.00811	(-0.13)
EMPL_UNI	0.0804**	(3.27)	-0.0707*	(-2.32)	0.108***	(4.09)
IMR	-0.0983	(-0.92)	0.224	(1.65)	-0.156	(-1.33)
M_POS_THRE	0.108	(1.66)	-0.105	(-1.28)	-0.0674	(-0.98)
M_CMP_UNPR	-0.183*	(-2.34)	0.194	(1.95)	-0.0216	(-0.26)
M_OUT_DATE	0.392***	(6.90)	-0.399***	(-5.56)	0.229***	(3.89)
M_PRO_SUBS	-0.235***	(-3.90)	0.243**	(3.16)	-0.230***	(-3.54)
M_DEM_UNFS	0.0405	(0.55)	-0.0762	(-0.82)	-0.0967	(-1.18)
M_FOR_PRES	0.0661	(1.18)	-0.0246	(-0.35)	0.0558	(0.93)
M_EXS	0.222	(1.18)	-0.114	(-0.49)	0.498*	(2.53)
N	1088		1089		1068	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.16 The 2nd stage product innovation with co-partnership (panel)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.994***	(3.75)	-0.700	(-1.90)	-0.128	(-0.39)
l_SIZE	-0.0465	(-0.75)	0.123	(1.41)	-0.0769	(-0.91)
EMPL_UNI	0.0799**	(2.92)	-0.0918*	(-2.33)	0.122**	(3.25)
IMR	-0.193	(-1.71)	0.385*	(2.26)	-0.334*	(-2.09)
CD_GROUP	0.170	(1.09)	-0.126	(-0.59)	0.250	(1.26)
CD_CSTMR	-0.0359	(-0.27)	0.0589	(0.32)	-0.104	(-0.62)
CD_SUPLR	0.0273	(0.18)	0.242	(1.11)	-0.267	(-1.36)
CD_COMPT	0.223	(1.56)	-0.334	(-1.69)	0.0461	(0.26)
CD_CNSLT	0.0878	(0.76)	-0.0727	(-0.45)	0.129	(0.88)
CD_UNIVR	0.0576	(0.46)	-0.129	(-0.75)	0.369*	(2.19)
CA_GROUP	-0.0484	(-0.26)	0.216	(0.79)	0.119	(0.48)
CA_CSTMR	0.172	(0.88)	-0.0714	(-0.27)	-0.105	(-0.42)
CA_SUPLR	-0.392	(-1.42)	0.231	(0.58)	-0.150	(-0.43)
CA_COMPT	0.362	(0.92)	-0.704	(-1.38)	0.461	(1.00)
CA_CNSLT	0.291	(1.18)	-0.215	(-0.65)	0.332	(1.07)
CA_UNIVR	-0.264	(-0.77)	0.733	(1.48)	0.0854	(0.20)
N	709		711		698	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.17 The 2nd stage product innovation with source of information (panel)

	(1) PRD_IMPR		(2) PRD_NCHG		(3) PRD_MNOV	
L.l_INEp	0.528*	(2.40)	-0.258	(-1.01)	-0.198	(-0.78)
l_SIZE	-0.0647	(-1.24)	0.120	(1.92)	-0.0395	(-0.64)
EMPL_UNI	0.0874***	(3.79)	-0.0871**	(-3.22)	0.120***	(4.23)
IMR	-0.297**	(-2.88)	0.365**	(2.82)	-0.183	(-1.46)
I_GROUP	0.147*	(2.31)	-0.0959	(-1.26)	0.161*	(2.08)
I_CSTMR	0.0189	(0.37)	0.0462	(0.76)	0.140*	(2.19)
I_SUPLR	0.0432	(0.82)	-0.0590	(-0.95)	0.0190	(0.30)
I_COMPT	0.00877	(0.16)	-0.0291	(-0.46)	-0.223***	(-3.36)
I_CNSLT	0.0328	(0.56)	-0.0190	(-0.27)	0.0122	(0.17)
I_UNIVR	0.0268	(0.45)	0.0608	(0.86)	0.146*	(2.05)
I_RDINS	0.0472	(0.75)	-0.104	(-1.43)	0.0869	(1.18)
N	816		817		805	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.18 The 2nd stage product innovation with public subsidies (panel)

	(1)		(2)		(3)	
	PRD_IMPR		PRD_NCHG		PRD_MNOV	
L.l_INEp	1.320***	(4.55)	-0.894*	(-2.40)	0.346	(1.21)
l_SIZE	-0.00768	(-0.12)	0.0703	(0.84)	0.0250	(0.38)
EMPL_UNI	0.115***	(3.78)	-0.140***	(-3.49)	0.125***	(4.01)
IMR	-0.0727	(-0.57)	0.210	(1.24)	-0.185	(-1.40)
PUB_SUBS	0.415**	(3.11)	-0.533**	(-2.82)	0.297*	(2.15)
N	850		852		833	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C.2.2 Test Results

Table C.19 Multicollinearity test for 2nd stage

Variable	VIF	Tolerance	R-Squared
lagLogINE	1.88	.532	.468
l_SIZE	4.16	.24	.76
EMPL_UNI	1.14	.875	.125
IMR	3.06	.327	.673

C.3 The Production Stage

C.3.1 Additional Determinants

Table C.21 The 3rd stage with protection measures (panel)

	(1)	
	l_P	
l_P		
L.pc1	-0.0607***	(-4.66)
L.pc2	0.229***	(8.49)
l_SIZE	-0.142***	(-4.41)
l_INVS	-0.0315	(-1.91)
IMR	-0.360***	(-6.80)
P_PATNT	0.0368	(1.06)
P_REGDS	-0.0611	(-1.66)
P_TRMKT	0.0358	(1.19)
P_CPYRT	-0.00215	(-0.06)
N	713	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.22 Estimation of 3rd stage with market characteristics (panel)

	(1)	
	l_P	
l_P		
L.pc1	-0.0644***	(-5.15)
L.pc2	0.252***	(9.73)
l_SIZE	-0.183***	(-5.75)
l_INVS	-0.0372*	(-2.44)
IMR	-0.294***	(-5.84)
M_POS_THRE	-0.0343	(-1.73)
M_CMP_UNPR	0.0135	(0.53)
M_OUT_DATE	0.0134	(0.75)
M_PRO_SUBS	-0.00958	(-0.52)
M_DEM_UNFS	-0.0176	(-0.73)
M_FOR_PRES	-0.00929	(-0.57)
M_EXS	0.538***	(7.28)
N	719	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.23 The 3rd stage with co-partnership (panel)

	(1)	
	l_P	
l_P		
L.pc1	-0.0665***	(-4.15)
L.pc2	0.248***	(7.07)
l_SIZE	-0.165***	(-4.20)
l_INVS	-0.0661**	(-3.22)
IMR	-0.315***	(-4.87)
CD_GROUP	-0.0157	(-0.30)
CD_CSTMR	-0.0134	(-0.31)
CD_SUPLR	0.0701	(1.54)
CD_COMPT	-0.119**	(-2.78)
CD_CNSLT	0.0499	(1.42)
CD_UNIVR	0.0196	(0.48)
CA_GROUP	0.230**	(3.26)
CA_CSTMR	-0.0364	(-0.48)
CA_SUPLR	-0.0818	(-0.88)
CA_COMPT	0.0403	(0.31)
CA_CNSLT	0.0858	(0.93)
CA_UNIVR	-0.0503	(-0.46)
N	473	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.24 The 3rd stage with public subsidies (panel)

	(1)	
	l_P	
l_P		
L.pc1	-0.0396**	(-2.95)
L.pc2	0.205***	(6.89)
l_SIZE	-0.0548	(-1.57)
l_INVS	-0.0436*	(-2.30)
IMR	-0.268***	(-4.75)
PUB_SUBS	-0.108**	(-2.74)
N	560	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.25 Estimation results of 3rd stage combination of innovation types (panel)

	(1)	
	l_P	
l_P		
L.pc1	0.00739	(0.58)
L.pc2	0.280***	(11.13)
l_SIZE	-0.0996***	(-4.12)
l_INVS	-0.125***	(-5.26)
IMR	-0.145**	(-3.14)
L.CPRD_PRC_COST	-0.0409	(-1.29)
L.CPRD_PRC_QUAL	0.0533	(1.67)
L.CPRD_ORG_TIME	-0.203**	(-2.96)
L.CPRD_ORG_QUAL	-0.0686	(-0.86)
L.CPRD_ORG_COST	0.277***	(4.79)
N	142	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C.3.2 Test Results

Table C.26 Correlation among explanatory variables used in 3rd stage (panel)

	PRD_IMPRp	PRD_NCHGp	PRD_MNOVp	PRC_COSTp	PRC_QUALp	ORG_TIMEp	ORG_QUALp	ORG_COSTp
PRD_IMPRp	1							
PRD_NCHGp	-0.970***	1						
PRD_MNOVp	0.841***	-0.867***	1					
PRC_COSTp	0.789***	-0.799***	0.857***	1				
PRC_QUALp	0.129***	-0.0636***	-0.356***	-0.449***	1			
ORG_TIMEp	-0.254***	0.189***	0.304***	0.119***	-0.844***	1		
ORG_QUALp	-0.568***	0.376***	-0.201***	-0.415***	-0.319***	0.561***	1	
ORG_COSTp	-0.837***	0.866***	-0.534***	-0.673***	-0.286***	0.577***	0.495***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.27 Multicollinearity test for 3rd stage before PCA

Variable	VIF	Tolerance	R-Squared
PRD_IMPRp	2.95e+13	0	1
PRD_NCHGp	4.73e+13	0	1
PRD_MNOVp	6.21e+13	0	1
PRC_COSTp	1.24e+14	0	1
PRC_QUALp	7.75e+13	0	1
ORG_TIMEp	1.74e+13	0	1
ORG_QUALp	7.92e+12	0	1
ORG_COSTp	2.54e+13	0	1
l_SIZE	7.52e+13	0	1
l_INVS	1.02	.981	.019
IMR	1.98e+14	0	1

Table C.29 PCA results on explanatory variables of 3rd stage (panel)

	Comp1	Comp2	Comp3	Comp4
PRD_IMPRp	0.4618	-0.0130	0.0800	0.3046
PRD_NCHGp	-0.4537	-0.0430	-0.3101	-0.0822
PRD_MNOVp	0.3860	0.3269	0.1469	0.3684
PRC_COSTp	0.4052	0.2653	-0.2554	-0.3447
PRC_QUALp	0.0241	-0.5901	0.3284	0.4184
ORG_TIMEp	-0.1288	0.5965	0.0087	0.3337
ORG_QUALp	-0.2743	0.2874	0.8131	-0.2751
ORG_COSTp	-0.4187	0.1846	-0.2038	0.5327
Eigenvalue	4.5221	2.4948	0.6479	0.3352
Proportion	0.5653	0.3119	0.0810	0.0419
Cumulative	0.5653	0.8771	0.9581	1.0000

Table C.30 Multicollinearity test for 3rd stage after PCA

Variable	VIF	Tolerance	R-Squared
pc1	3.42	.292	.708
pc2	5.36	.186	.814
l_SIZE	8.31	.12	.88
l_INVS	1.02	.982	.018
IMR	8.04	.124	.876

Table C.32 Robustness check (censoring) 4th equation

	(1) l_P	
pc1	-0.0463^{***}	(-7.13)
pc2	0.111^{***}	(10.79)
l_SIZE	-0.0664^{***}	(-4.86)
l_INVS	-0.0420^{***}	(-5.86)
IMR	-0.256^{***}	(-9.99)
PX	0.454^{***}	(23.62)
N	3556	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix D

Software Tools

The Interface between Software Tools

Figure D.1 illustrates the relationship between software tools used in this work.

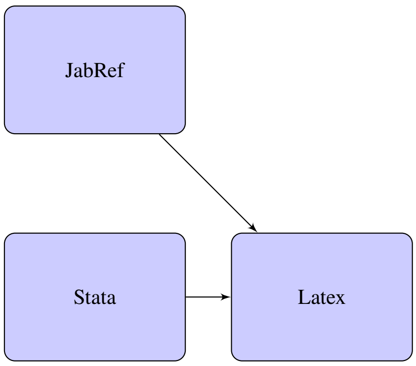


Fig. D.1 The interaction between software tools

Reference and Bibliography Management

To support a systematic review of the existing literature, the the software tool JabRef ® has been employed as a reference and bibliography management. The advantage of JabRef is that it uses 'BibLaTeX' and 'BibTex' as a format, hence, it is easily interfaced to LaTeX editor, which is used in this work. All referenced literature is saved in a database and linked automatically and in traceable manner into the text editor LaTeX, which reduces the potential of human mistakes and improves the quality of referencing.

Data Management and Analyses

To test the research hypotheses using the MIP dataset, a statistical software tool for estimating the econometric model is needed. This study employs Stata ® software tool version 14, which has a reach variety of characteristics:

- Powerful and efficient environment for research in applied economics and analysing statistical results.
- Wide range of functions, features, graphic options, command sets for different kinds of data, which includes a variety of commands to enhance analysis of panel data.
- Comprehensive language reduce systematic errors and acceptable usability if compared with other statistic analyses tools such as SAS, EVIEWS or R.
- Statistical packages available in menus, the data editor, script editor, variable manager, log viewer, and graphs drawing are all functions that enhance tool usability and capability.
- Using the same functions by either command-line or menus.
- Moderate price if compared with other commercial tools of the same purpose.

From dataset point of view, the size of used data in this project optimises Stata ® usage. Dataset is available in Stata format, which facilitate importing data to the work environment.

Stata ® has an interface to LaTeX(the used script editor in this work) that enhances integrating analyses outputs such as lists of variables, summary tables, graphs, estimation results, which minimise errors.

Abbreviations

2SLS Two-stage least square. 149, 150

3SLS Three-stage least square. 150, 151

BMBF The German Federal Ministry of Education and Research. 105

CDF Cumulative Distribution Function. 142

CDM Crepon, Duguet, and Mairesse econometric model. vii, 14, 19, 37, 56–61, 82, 83, 113, 140, 197, 198, 206, 207

CIS Community Innovation Survey. 30, 32, 38, 39, 41–44, 46, 48, 53, 57, 61–63, 66, 69, 70, 74–76, 78, 105, 106, 108–112, 206

CME Coordinated Market Economies. 47, 48

DIW Institute of the German Economy. 2

EU European Union. 72, 142–144, 188, 191, 195, 203, 205

FA Factor Analysis. 152

FE Fixed Effects. 131, 189, 190

FIML Full Information Maximum likelihood. 150, 151

GDP Gross domestic product. 7

GLS Generalized Least Square. 151

ICT Information and Communication Technology. 11, 12, 19, 32, 51–54, 57–59, 62, 68, 70, 76, 79, 81, 102, 143, 202, 209, 210

IID Independent Identically Distributed. 166

ILS Indirect Least Square. 149, 150

IMR inverted Mills' ratio. 138, 141, 142, 177, 186–188, 190, 191, 195, 199

IT Information Technology. 22, 42, 52, 53, 70, 117

IV Instrumental variables. xvii, 136, 149, 150

LIML Limited Information Maximum Likelihood. 149, 150

LME Liberal Market Economies. 47, 48

LR Likelihood Ratio. 189, 190, 198

MFP Multi Factor Productivity. 11, 21

MIP Mannheim Innovation Panel. vii, 13, 44, 46, 60, 67, 73, 76, 80, 81, 101, 102, 105, 109–114, 116, 117, 119, 124, 126, 130, 143, 148, 157, 197, 209, 210, 268

OECD The Organization for Economic Cooperation and Development. 11, 106, 112

OLS Ordinary least square. 133, 149, 150

OR Odds Ratio. 176, 177, 187

PCA Principal Component Analysis. vii, 14, 152, 153, 184, 185, 193, 194, 196, 199, 200, 202, 207

PDF Probability Density Function. 142

R&D Research and Development. 5, 14, 22, 27, 28, 31, 32, 34–38, 40–43, 48–52, 55–58, 61–66, 68, 71–76, 78–82, 108, 110–113, 116, 117, 138, 142, 206

RE Random Effects. 131, 189, 190

SEM Simultaneous Equation Model. 135

SME Small and Medium Enterprises. 2, 37, 62, 63

SUR seemingly unrelated regressions. 150

TFP Total Factor Productivity. 5, 21, 22, 27, 52, 71, 77

VIF Variance Inflation Factor. 134, 190, 193, 196

ZEW The Centre for European Economic Research. 105, 108, 112, 113, 210