TECHNISCHE UNIVERSITÄT DARMSTADT

ICT, Productivity and Economic Growth
–
Empirical Results on Country Level

Vom Fachbereich Rechts- und Wirtschaftswissenschaften
der Technischen Universität Darmstadt

genehmigte

Dissertation

von

Dipl.-Wirtsch.-Inform. Mathias Rhiel
geboren in Langen (Hessen)

zur Erlangung des akademischen Grades
Doctor rerum politicarum (Dr. rer. pol.)

Erstgutachter: Prof. Dr. Jens Krüger
Zweitgutachter: Prof. Dr. Volker Caspari

Darmstadt 2018

Rhiel, Mathias : ICT, Productivity and Economic Growth – Empirical Results on Country Level

Darmstadt, Technische Universität Darmstadt

Dissertation veröffentlicht auf TUprints im Jahr 2018

Tag der mündlichen Prüfung: 05.07.2018

Acknowledgments

This dissertation was written while I was research assistant at the Chair of Empirical Economics at the Technische Universität (TU) Darmstadt, Germany. The progress and completion of my thesis would not have been possible without the support of many people, whom I sincerely thank with the following acknowledgements.

First and foremost, I am deeply grateful to my supervisor and doctoral advisor Prof. Dr. Jens Krüger. His knowledge, opinions, helpful advice and patience have always been of great assistance and value to me. I would also like to thank Prof. Dr. Volker Caspari for his support as my second supervisor. My interest in this research topic was originally sparked by his inspiring lectures. Additionally, I want to thank Prof. Dr. Michael Neugart for chairing the examination committee as well as the committee members Prof. Dr. Alexander Kock and Prof. Dr. Dirk Schiereck for their questions and insights during the defense.

Moreover, sincere thanks are due to my former colleagues Christian Berker, Dr. Benjamin Hampf, Dr. Stephan Hitzschke, Dr. Julian LeCrone, Prof. Dr. Günther Rehme and Sebastian Ruths Sion, with whom I had many interesting and fruitful discussions and who provided valuable feedback on my research. Vanessa G. Belew and Philip Savage provided excellent support by proof-reading my academic writings. I am especially grateful to my colleagues from the Chair of Empirical Economics for the very cooperative work atmosphere during the last years.

Furthermore, I would like to thank and acknowledge TU Darmstadt for the excellence working conditions and their support in helping me to participate in conferences and seminars at home and abroad.

Last, but definitely not least, I want to thank my family for the unconditional trust and support during all the years of education. An especially heartfelt expression of gratitude is owed to my wife Rebekka for her loving, trust and patience. All of this gave me the strength to persist through challenging times.

Abstract

This dissertation examines the relationship between Information and Communication Technology (ICT), productivity and economic growth. ICT, as one of the driving forces for productivity development and thus for economic growth, is of considerable interest to economics. Although the impact of ICT on productivity and economic growth has already been examined in a wide range of national studies (particularly for the U.S.), comparative studies for the rest of the world are scarce. For this reason, this thesis extends the existing available cross-country literature on this topic by three empirical studies which investigate the economic impact of ICT for a broad sample of countries at all development stages. The work focuses on the question of whether ICT contributes significantly and positively to productivity and economic growth globally or whether this only applies to individual countries or country groups.

Chapter 2 provides an overview of the current state of research on the contribution of ICT to productivity and economic growth. First, the economic impact of ICT is discussed from a theoretical perspective and reflected on from a productivity standpoint. Furthermore, the empirical approaches of growth accounting and estimating production functions, which are commonly used in the literature to investigate the research subject, are presented and discussed. Subsequently, an overview of the subject-related literature is given.

Chapter 3 revolves around the definition and quantitative measurement of ICT. In the literature, a quantitative measurement of ICT is mostly given monetarily in the form of capital stocks. The disadvantages of using capital stocks to measure ICT are pointed out and the use of non-monetary penetration rates as adequate proxy variables for ICT is motivated. By performing a statistical principal component analysis, the penetration rates are merged into a single variable that comprises most of the information. The thus constructed ICT infrastructure variable serves as a dependent or descriptive proxy variable for ICT in the respective empirical analyses.

Chapter 4 examines the economic and institutional determinants of ICT infrastructure, which explain its diffusion over a broad cross-section of more than 100 countries for the period 2002-2012. This follows the well-known fact that developed countries possess a higher level of ICT than developing countries. These differences in the distribution are examined in the literature under the term "global digital divide". The chapter follows an approach based on variable selection methods originating from machine learning research. This approach considers a broad set of candidate explanatory variables simultaneously and selects the most relevant ones. The ICT infrastructure variable is subsequently regressed to these selected variables. The results show that the identified determinants have a high degree of explanatory power to describe the diffusion of the ICT infrastructure.

Chapter 5 investigates the role of ICT in economic growth for the long-term period of 30 years (1980-2010). It is examined whether there is a positive and significant relationship between ICT and per capita income growth over a sample of more than 95 countries. The ICT infrastructure variable as constructed in chapter 3 is included in a commonly used cross-country regression model. To avoid the problem of endogeneity due to reverse causality between per capita income and ICT, an instrumental variable approach is applied. The results show that ICT infrastructure explains the per capita income growth during the investigation period positively and significantly.

Chapter 6 examines the role of ICT in productivity growth for more than 120 countries in the 2001-2012 period. There is particular interest in the research question of whether developing countries have also been able to obtain substantial productivity increases through the use of ICT. This is achieved by applying an extension of the non-parametric Multi-directional Efficiency Analysis (MEA) approach. The results show that ICT productivity increased worldwide over the investigation period. The results further reveal that developing countries benefit to a lesser extent from the productivity-enhancing effects of ICT in comparison to developed countries. A regression analysis also identifies factors that explain the differences in ICT productivity between countries.

Summary in German – Zusammenfassung

In der vorliegenden Dissertation wird der Zusammenhang zwischen der Informations- und Kommunikationstechnologie (IKT), Produktivität und Wirtschaftswachstum untersucht. Die IKT ist als ein Treiber der Produktivitätsentwicklung und somit auch des Wirtschaftswachstums von erheblichem Interesse für die Ökonomie. Obschon der Einfluss der IKT auf Produktivität und Wirtschaftswachstum bereits in einer Vielzahl nationaler Studien (insbesondere für die USA) untersucht wurde, gibt es nur wenige Studien die dieses Thema länderübergreifend analysieren. Aus diesem Grund wird durch diese Dissertation die bestehende wissenschaftliche Literatur von länderübergreifenden Untersuchungen zu diesem Thema um drei empirische Studien erweitert, welche die ökonomischen Auswirkungen der IKT für eine umfassende Anzahl von Ländern unterschiedlicher Entwicklungsstufen untersuchen. Im Fokus der Arbeit steht dabei die Frage, ob die IKT global einen signifikanten und positiven Beitrag zu Produktivität und Wirtschaftswachstum erbringt oder ob dies nur für einzelne Länder oder Ländergruppen zutreffend ist.

Kapitel 2 vermittelt einen Überblick über den aktuellen Forschungsstand in Bezug auf den Beitrag der IKT zu Produktivität und Wirtschaftswachstum. Zunächst wird der wirtschaftliche Einfluss der IKT aus theoretischer Sicht diskutiert und reflektiert, wie sich dieser auf die Produktivität auswirkt. Desweiteren werden die empirischen Ansätze des Growth Accounting und der Produktionsfunktionsschätzung vorgestellt und diskutiert, welche in der Literatur üblicherweise zur Untersuchung des Forschungsthemas Anwendung finden. Daran anschließend wird ein Überblick über die themenrelevante Literatur gegeben.

Kapitel 3 behandelt die Definition und quantitative Erfassung der IKT. In der Literatur wird eine quantitative Erfassung der IKT meist monetär in Form von Kapitalstöcken vorgenommen. Die Nachteile zur Beschreibung der IKT mittels Kapitalstöcken werden aufgezeigt und darauf aufbauend die Nutzung nicht-monetärer Penetrationsraten als adäquate Proxyvariablen für die IKT motiviert. Die Penetrationsdaten werden unter Durchführung einer statistischen Hauptkomponentenanalyse zu einer Variablen zusammengefasst, die den Großteil der Informationen aus den Penetrationsdaten in einer einzelnen Variablen bündelt. Die so konstruierte IKT-Infrastrukturvariable dient als abhängige bzw. beschreibende Stellvertretervariable für IKT in den jeweiligen empirischen Analysen.

In Kapitel 4 werden die ökonomischen und institutionellen Determinanten der IKT-Infrastruktur für den Zeitraum 2002-2012 über einen breiten Querschnitt von mehr als 100 Ländern untersucht, welche deren Verbreitung erklären. Dies geschieht vor dem allgemein bekannten Hintergrund, dass entwickelte Länder über ein höheres Maß an IKT verfügen als Entwicklungsländer. Diese Verteilungsunterschiede werden in der Literatur unter dem Begriff der „globalen digitalen Kluft“ untersucht. Dieses Kapitel verfolgt einen Ansatz, der auf Variablenselektionsmethoden aus dem Forschungsbereich des maschinellen Lernens basiert. Dieser Ansatz behandelt eine große Zahl erklärender Variablen simultan und wählt die relevantesten Variablen aus. Die IKT-Infrastrukturvariable wird anschließend auf diese selektierten Variablen regressiert. Die Er-

gebnisse zeigen, dass die so identifizierten Determinanten ein hohes Maß an Erklärungskraft zur Beschreibung der Diffusion der IKT-Infrastruktur besitzen.

In Kapitel 5 wird die Rolle der IKT für das langfristige Wirtschaftswachstum über einen Zeitraum von 30 Jahren (1980-2010) erforscht. Dabei wird die Frage untersucht, ob ein positiver und signifikanter Zusammenhang zwischen der IKT und dem langfristigen Wirtschaftswachstum über ein Datensample von mehr als 95 Ländern besteht. Zu diesem Zweck wird die in Kapitel 3 konstruierte IKT-Infrastrukturvariable einem erweiterten länderübergreifenden Wachstumsregressionsmodell hinzugefügt. Aufgrund vermuteter reverser Kausalität zwischen der Höhe des Pro-Kopf-Einkommens und der IKT-Infrastruktur werden auch Verfahren der Instrumentvariablenschätzung eingesetzt. Die Ergebnisse zeigen, dass die IKT-Infrastruktur das Wachstum des Pro-Kopf-Einkommens im Untersuchungszeitraum positiv und signifikant erklärt.

In Kapitel 6 wird die Rolle der IKT für das Produktivitätswachstum im Zeitraum von 2001 bis 2012 für mehr als 120 Länder untersucht. Es besteht ein besonderes Interesse an der Forschungsfrage, ob auch die Entwicklungsländer substanzielle Produktivitätssteigerungen durch den Einsatz von IKT erzielen konnten. Hierbei wird eine weiterentwickelte Form des nichtparametrischen Ansatzes der Multidirektionalen Effizienzanalyse (MEA) angewendet. Die Ergebnisse zeigen, dass die Produktivität der IKT über den Untersuchungszeitraum weltweit gestiegen ist, wobei die Entwicklungsländer nicht im gleichen Umfang von den produktivitätserhöhenden Effekten der IKT profitieren konnten, wie die entwickelten Länder. Die Durchführung einer Regressionsanalyse identifiziert zudem Faktoren, welche die Produktivitätsunterschiede der IKT zwischen den Ländern erklären.

Contents

List of Tables

List of Figures

1 Introduction

1.1 The Emergence and Diffusion of ICT

"[We] see the computer age everywhere...",[1] an observation often cited at the beginning of scientific articles and publications, many of which are thematically related to this dissertation. The phrase in question is part of a quote that dates back to the year 1987, in which computers were already to be found in most offices in the business world. A few years before, in January 1983, *Time* magazine has selected the personal computer as its *Man of the Year*.[2] In this context, the computer is viewed as a representative for the more comprehensive Information and Communication Technology (ICT). The term ICT generally refers to equipment and services related to broadcasting, computing and telecommunications, all of which capture, process and display information electronically (United Nations 2004).[3] Since 1987, the diffusion of ICT has been further intensified worldwide and has become an integral part of almost all economic activities and daily life. Today ICT is not just represented in the form of a computer on every office desk, but also in the smartphones in our pockets. No one doubts that ICT has radically changed life and society over the past decades.

The development of ICT is closely connected with the rapid progress of semiconductors.[4] The "birth of modern ICT" (Jorgenson and Vu 2016, p. 383) was marked by the invention of the transistor, a semiconductor device that acts as an electrical switch and encodes information in binary form (Acs et al. 2013). The first functioning transistor was presented to the public in 1947 by the Bell Laboratories (Ross 1998). In 1956, their inventors William Bradford Shockley, John Bardeen and Walter Houser Brattain received the Nobel Prize for Physics for this construction.

Additional progress in ICT emerged through the invention of integrated circuits by Jack Kilby of Texas Instruments in 1958 and Robert Noyce of Fairchild Semiconductor in 1959. An integrated circuit consists of multiple, up to million of transistors on a single chip based on silicon (Swedin and Ferro 2005). By enabling data storage and retrieval in binary form, it became also known as memory chip. In 2000, their invention was rewarded with the Nobel Prize for Physics.

Since the introduction of integrated circuits, the performance of computers has increased exponentially. Gordon E. Moore (co-founder of Intel) made a prescient observation, later known as Moore's Law. By plotting data on memory chips he observed that the transistor density on integrated chips doubles every 18-24 months, which implies an exponential growth rate of 35-45 percent per year (Jorgenson 2009). Moore's law can be illustrated by the development of transistors on Central Processing Units (CPU). A CPU takes instructions from a program and works as the brain of a computer. The processing capability of a computer can be roughly assessed on the basis of the number of transistors on its CPU chip. The number of transistors on Intel's CPU chip increased from 2250 in 1971 (Intel 4004 model) to 125 million in 2004 (Pentium 4

[1] These words are part of a quotation from Nobel Prize winner Robert M. Solow (Solow 1987), to which we return below.

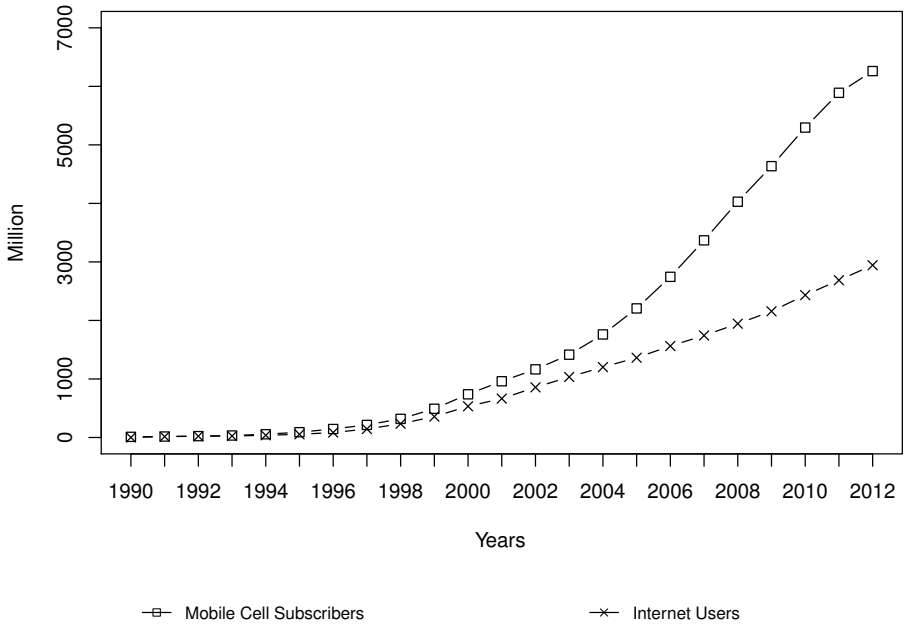
[2] http://content.time.com/time/covers/0,16641,19830103,00.html.

[3] A detailed description and definition of the term ICT is provided in chapter 3.

[4] This means technological change and product improvement in semiconductors and the steep and sustained decline in semiconductor prices.

Prescott), 291 million in 2006 (the Core 2 Duo Conroe model) and 19200 million in 2017 (the 32-core AMD Epyc model). This implies an annual growth rate of 41 percent over 1971-2017. Hence, the development of transistors on CPUs has followed Moore's Law since its introduction with astounding accuracy. According to Moore, this development and the validity of Moore's law will continue for the next 10 years.[5]

Figure 1.1: World Internet Users and Mobile Phone Subscribers, 1990-2012



Data Source. World Development Indicators (World Bank).

Just as the innovation in semiconductors and PCs was the basis for the progress of ICT, the emergence of the internet and mobile technology in the 1990s has driven the rapid diffusion of ICT applications across sectors and nations (Jorgenson and Vu 2016). In figure 1.1 it is shown that the number of internet users increased from 1.3 million in 1990 to almost 3 billion in 2012. In the same period, the number of mobile phone subscribers soared from 11 million to more than 6 billion. Equally remarkable is the worldwide diffusion of ICT into developing countries. The spread of internet and mobile phones has reached even the poorest and most isolated nations.[6]

1.2 The Economic Impact of ICT

Sustainable economic growth is of the highest priority for policy makers as it promises high standards of living. Examining the role and contribution of ICT as one of the driving forces for productivity development and thus for economic growth has been of considerable interest

[5] http://venturebeat.com/2015/05/11/intels-gordon-moore-speculates-on-the-future-and-the-end-of-moores-law/.

[6] Starting from nil in 1990, the penetration rate of mobile phones (per 100 inhabitants) in 2012, for example, was 25.3 in the Central African Republic, 22.8 in Burundi, 7.1 in Myanmar and 6.9 in North Korea (Source: World Bank Development Indicators).

to economists from an early stage. ICT is able to improve productivity in several ways: by increasing the amount of capital deployed per worker (i.e. capital deepening), by speeding up the aggregated productivity of an economy due to technological improvements in the ICT-producing sector, and as an enabler of products, processes and organizational innovation in the sectors that use ICT. Because ICT has the potential to improve productivity in a multitude of industries and sectors, it is often considered a "general purpose technology". A general purpose technology is characterized by "the potential for pervasive use in a wide range of sectors and by their technological dynamism" (Bresnahan and Trajtenberg 1995, p. 84) and in that it may lead to substantial productivity gains in the economy as a technology that enables complementary innovations.

The role of ICT to productivity and economic growth has been investigated in numerous studies in more than 25 years of research since the 1980s.[7] In the 1980s and early 1990s, the debate was dominated by the phenomenon of the so-called "productivity paradox", which describes the "discrepancy between measures of investment in information technology and measures of output at the national level" (Turban et al. 2002, p. 592). The term emerged from the observation of a significant productivity slowdown in OECD countries since the early 1970s, which lasted about 20 years despite high investments in ICT (Macdonald et al. 2000) and extensive technical progress in computer power.

The issue received public attention through a book review by Robert Solow, which was published in The New York Times in July 1987. This article included the statement "we see the computer age everywhere except in the productivity statistics" (Solow 1987). Solow's remark in 1987 stirred up the discussion about the impact of ICT on productivity and growth and led to an intense effort to measure the economic impact of ICT. Since his famous quote, the productivity paradox of information technology has been known as the Solow paradox (see e.g. Jorgenson and Stiroh 1995).

The empirical literature on the impact of ICT on productivity and economic growth can be described in several dimensions. In particular, there are studies at different levels of aggregation that examine the impact of ICT at firm, industry or country level. Different empirical strategies have been pursued, such as the application of growth accounting techniques or the estimation of production functions. The investigations specify ICT in various ways and use different definitions of hardware and of software. Early studies, for example, usually consider information technology (IT) only and exclude communication aspects.

The phenomenon of the productivity paradox was originally studied on the highly aggregated country-level. Early country-level studies show low or no contribution of IT/ICT to productivity and economic growth in the U.S. (Oliner and Sichel 1994, Jorgenson and Stiroh 1995).[8] This situation changed considerably in the late 1990s. After two decades of productivity slowdown, the U.S. experienced a period of increasing productivity growth. The majority of researchers

[7] The work of Hardy (1980) can be seen as the starting point for a period of more than 35 years of research in this field.

[8] In hindsight, a possible reason for this result is seen in the low level of ICT investment, which was (proportional to the capital stock) too small to show economic effects (Sichel 2001). The nominal share of IT investment as a percent of total business investment grew from 2.6% in 1970 to 3.5% in 1980, to 9% in 1990 and 22% in 1999 (Dedrick et al. 2003).

agree on the importance of ICT for U.S. growth resurgence (see e.g. Jorgenson and Stiroh 2000, Oliner and Sichel 2002). The technological progress in the ICT-producing sector in the mid-1990s has led to rapid price decline of ICT products. These price declines triggered higher investment in ICT and substitution of less productive inputs in the ICT-using sectors, especially in the U.S. (Jorgenson 2005). The share attributable to ICT in U.S. growth performance went from 43% for the period 1970-1995 to 59% for the period 1995-2000 (Jorgenson et al. 2008). Oliner and Sichel (2000, p. 21) infer that "information technology accounted for about two-thirds of the step-up in labor productivity growth between the first and second halves of the decade."

Since the majority of the studies confirm a positive and substantial contribution of ICT to productivity and economic growth in the U.S. since the mid-1990s, Solow's productivity paradox seems to have been resolved. A comparison of U.S. results with other countries, however, reveals differences in the economic impact of ICT. It was recognized, for example, that the EU could not benefit from the productivity increasing effects of ICT to the same extent as the U.S. (van Ark and Inklaar 2005).[9]

While the economic impact of ICT has been widely explored in the U.S., comparative studies for the rest of the world are scarce. The reason for this is mainly due to the limited availability of national and comparable data. Hence, there are only few studies in the literature that examine the economic impact of ICT for a wide range of countries. The results of these studies are also contradictory. There are studies which find a positive growth effect of ICT over a broad country sample (Papaioannou and Dimelis 2007) as well as studies which find positive and significant effects only for the group of developed countries (Dewan and Kraemer 2000). Other studies cannot find an economic impact of ICT, either for the group of developing countries or for the group of developed countries (Pohjola 2002).

The results of these investigations show that the findings from the empirical literature for the U.S. do not necessarily apply to other countries. In other words, the U.S. studies are not sufficient to assess the economic impact of ICT globally. Thus, a comprehensive assessment of the productivity and growth contribution of ICT requires research at the cross-country level that includes a sufficient amount of countries, especially at all stages of development.

1.3 Research Objectives

This dissertation empirically examines the relationship between ICT, productivity and economic growth. In so doing, we extend the existing available cross-country literature on this topic by empirical studies which investigate the economic impact of ICT for a broad sample of countries at all development stages. These empirical analyses address the fundamental question of whether ICT contributes significantly and positively to productivity and economic growth or if this is only the case for individual countries or groups of countries. The research objective of these empirical analyses is therefore to provide a global assessment about the economic impact of ICT.

A first objective of this dissertation is to develop an appropriate measurement of ICT. The reason for the rarity of cross-country studies is obviously the lack of available data on ICT. These data

[9] This phenomenon has been named as the productivity gap between the U.S. and the EU (van Ark et al. 2008) or "Atlantic divide" (Timmer et al. 2003).

are difficult to obtain, especially for a sufficiently large number of developing countries. It is therefore in our research interest to examine how the spread of ICT can be measured for a broad sample of countries at all levels of development. In particular, we are interested in the question of which indicators or proxies are appropriate to measure ICT quantitatively.

Subsequently, this dissertation addresses the question of the determining factors of ICT. It is a well-known fact that ICT is not spread equally across all countries. Developed countries possess a higher level of ICT than developing countries. These differences have already been considered under the term of 'global digital divide' in the literature. Against this background of a positive impact of ICT on the national economies, we are therefore interested in what determines the 'global digital divide'. In the context of a possible macroeconomic growth effect of ICT, the issue of what determines ICT is also of interest for policy makers.

The digital revolution in the respective countries took place at different times and at different speeds. The beginning of the economic effects of ICT can be dated back to the early 1980s. For this reason, this dissertation examines whether ICT had a significantly positive contribution to the long-term growth of labor productivity worldwide. In this context particular attention needs to be paid to the possible problem of endogeneity, which exists due to potential reverse causality between the income levels of countries and their respective ICT levels.

Apart from the question of the global effects of ICT, the dissertation provides an analysis of ICT-specific productivity changes. The different and sometimes contradictory results from the cross-country literature suggest that there are differences between countries with regard to the effect of ICT on productivity. We will examine these differences in detail to discover differing patterns across groups of countries. We are particularly interested in groups classified according to their development status. Furthermore, we are also interested in finding factors that explain the country differences in the productivity changes of ICT.

1.4 Outline of the Dissertation

The following gives a brief overview of this dissertation's structure. Chapter 2 provides an overview of the current state of research on the contribution of ICT to productivity and economic growth. At first, the economic impact of ICT is discussed from a theoretical perspective while reflecting how ICT affects productivity. Second, the two major empirical approaches that are largely used in the literature to measure the impact and contribution of ICT on productivity and economic growth are briefly introduced. These are growth accounting and the approach using production function estimation. These two methods and their advantages and disadvantages are discussed. Third, an overview of the literature is given that examines the impact and contribution of ICT on productivity and economic growth in a myriad of studies since the 1980s. The literature is classified according to the aggregation level and categorized into studies at the disaggregated firm level, studies at the industry level and research at the aggregate country level.

Chapter 3 revolves around the definition and quantitative measurement of ICT. The chapter starts with a description of the terminology and definition of ICT. Based on this definition, the situation of available data is subsequently reviewed. In the literature, IT and ICT are

usually measured in monetary terms using capital data. We discuss the disadvantages of using capital data to describe ICT and motivate the usage of non-monetary penetration rates of ICT infrastructure as appropriate proxies for ICT. These penetration levels are available for a broad number of countries and overcome several of the disadvantages of monetary measurement. Having constructed an appropriate variable from the penetration rates, we descriptively analyze the distribution and development of ICT infrastructure.

Chapters 4 to 6 pursue the research objectives briefly outlined in this introduction. Chapter 4 examines the determinants of ICT infrastructure diffusion. Most of the studies (at the macro level) share the same approach of first making theoretical assertions about factors influencing ICT, followed by identifying appropriate indicators for these factors and running regressions of the dependent ICT variable on the identified explanatory variables. Depending on the theoretical assertions, this consequently leads to diverse findings in the literature. We therefore pursue an approach based on variable selection methods originating from machine learning research that considers a broad set of candidate explanatory variables simultaneously and selects the most relevant ones. By using the so-called Lasso and several of its more advanced variants, we investigate economic and institutional determinants of ICT infrastructure for the period 2002-2012 for a broad cross section of more than 100 countries.

Chapter 5 investigates the role of ICT in economic growth for the long-term period of 30 years (1980-2010). Although the impact of ICT to economic growth is assessed in the literature, only few investigations cover a period of more than 20 years and a sufficient number of countries at different stages of development. This chapter therefore examines whether there is a positive and significant relationship of ICT and long-term economic growth across countries. For this purpose, the ICT infrastructure variable as constructed in chapter 3 is included in a commonly used cross-country regression model. To avoid the problem of endogeneity due to reverse causality between GDP per capita and ICT, we apply an instrumental variable approach.

Chapter 6 examines the role of ICT in productivity growth. Most of the previous research has found ICT investment to be associated with significant productivity gains for developed countries but not or to a lesser extent for developing countries. Nonetheless, developing countries have also increased investments in ICT (infrastructure). An important research objective is therefore to examine whether developing countries achieved significant productivity gains through ICT. The empirical approach followed in this paper relies on an extension of the non-parametric Malmquist total factor productivity index that enables variable-specific analyses of productivity change across countries, respectively country groups. This allows us to analyze the differences in the input-specific productivity gains of a country's development stages and to make an explicit statement about the contribution of ICT to productivity.

Chapter 7 puts results in perspective by highlighting the main insights we have gained from our analyses and by showing how the results in chapters 4, 5 and 6 relate to each other. The study concludes by pointing out some lessons for (development) policy, outlining possible extensions to our analyses and discussing several related areas that deserve further attention.

2 Contribution of ICT to Productivity and Economic Growth

In this chapter, we present the current state of research on the contribution of ICT to productivity and economic growth. The chapter is divided into four section. In section 2.1, we discuss the impact of ICT on productivity and economic growth from a theoretical perspective and reflect how ICT affects the productivity of an economy. We also consider the literature on the question of whether ICT can be regarded as a general purpose technology. In section 2.2, we briefly introduce the two major empirical approaches that are widely used in the literature to measure the impact and contribution of ICT on productivity and economic growth. These are growth accounting and the approach using production function estimation. We discuss these two methods and their advantages and disadvantages in the context of this field of research. This also serves to classify the studies carried out in the subsequent literature review. In section 2.3, we provide an overview of the literature that examines the impact and contribution of ICT on productivity and economic growth. The literature is classified according to the aggregation level and categorized into studies at the disaggregated firm level, studies at the industry level and research at the aggregated country level. In section 2.4, we conclude this chapter by summarizing the key results of each section.

2.1 Impact of ICT to Productivity and Economic Growth

This section addresses the theories on the impact of ICT on productivity and economic growth. ICT affects economic growth in general and productivity in particular in various ways. Basically, there are two types of ICT effects on productivity; direct and indirect effects.

First, ICT is part of produced goods (such as computers, network infrastructure) and services (such as cloud storage). Technological progress and productivity growth in ICT-producing sectors have a direct effect on the aggregate productivity of an economy proportional to the size of the ICT sector (see Jorgenson et al. 2002 and 2008, van Ark et al. 2008).

Indirectly, ICT also affects productivity in the sectors in which it is used. For instance, ICT is used in production in the form of CNC machines.[10] In industrial applications, the use of software enables the automation of processes. An indirect effect is thus created by capital deepening in the ICT-using sectors as a result of investment, which helps to increase productivity and GDP growth.

The effects mentioned should not be considered separately, but they interact with each other. The technological progress in the ICT-producing sector in the mid-1990s led to rapid price declines for ICT products. These price declines triggered higher investment in ICT and substitution of less productive inputs in the ICT-using sectors, especially in the U.S. (Jorgenson 2005). Research on this topic was undertaken by, inter alia, Stiroh (2002b), Jorgenson et. al (2008) and van Ark et al. (2008) and will be addressed in section 2.3.

As a further (also indirect) effect, ICT can increase productivity due to spillover effects and externalities beyond the ICT-producing sector. As Cardona et al. (2013, p. 111) points out,

[10] The abbreviation CNC stands for Computer Numerical Control and describes the automation of machine tools by using computers that execute pre-programmed sequences of machine control commands.

investments in ICT "allow faster information processing and firms to think of new ways of communicating with suppliers or customers or arranging new distribution systems. Internal processes can be streamlined, reducing capital needs through better utilization of equipment and reduction in inventories and thereby space requirements. Increased communication and more timely and widespread transfer of information reduces coordination costs, the number of supervisors required reducing labor costs and facilitate better decision making."

Since ICT is able to generate, store and transmit information, it enables the ICT-using sectors to reduce information asymmetries (Lechman 2015), which is one of the major causes of high transaction costs, uncertainty and therefore market failure (Wolf 2001). A reduction of information asymmetry again enhances the efficiency of resource allocation (Akerlof 1970). By facilitating communication, ICT further promotes efficient processes of collaboration and thus the creation of new knowledge (Forman and van Zeebroeck 2012).

Because ICT reduces information asymmetry and thus affects a multitude of industries and sectors by making them more productive, the impact of ICT goes beyond conventional capital equipment. For this reason, ICT is often considered an "enabling technology" (Jovanovic and Rousseau 2005). According to Bresnahan and Trajtenberg (1995), this enabling role of ICT qualifies ICT as a General Purpose Technology (GPT). GPTs are technological innovations that affect production and/or innovation in many sectors of an economy. Well known examples of GPTs in economic history include the steam engine, electricity and the internal combustion engine (Gordon 2012). The main characteristics of a GPT are the following (Bresnahan and Trajtenberg 1995):

1. "Pervasiveness" of the technology: due to its broad applicability, the technology should be applicable for a broad range of users.
2. Inherent potential for technical improvements: the GPT allows continuous improvements and experimentation and facilitates innovation in using sectors through co-inventions.
3. Innovation spanning: the GPT should ease the inventing process and create new products or processes.

ICT appears to comply with all the characteristics of a GPT. The first two properties of a GPT are confirmed inter alia by industry studies on U.S. data which show that ICT investment in several (non ICT-producing) sectors has led to higher (total factor) productivity (Baily and Lawrence 2001, Stiroh 2002b, Bosworth and Triplett 2007). Other authors find that ICT has a significant and positive impact on innovation (Becchetti et al. 2003, Bertschek et al. 2013), which also confirms the second and third properties of GPTs.

The impact of a GPT on productivity and economic growth is not likely to be observed immediately after its invention. The three characteristics of a GPT jointly point to a time consuming process. The historical analysis of GPTs by David and Wright (1999) reveals three main development stages. In the first stage, an increase of productivity growth in the GPT-generating sector is observable. In the subsequent second stage, a significant capital increase in goods embedding GPT is noticed, stimulated by a price reduction of these goods. In the final third stage, the

GPT-using sectors reorganize their production. Thus, it takes time (according to David and Wright (1999) up to decades) to generate and observe spillovers. This may explain the fact that only a weak or no ICT contribution to productivity growth could be found until the mid-1990s, well expressed by the Solow Paradox. Using time-lagged data (on firm level), studies show that the effects of ICT on productivity is stronger over longer periods, which confirms the time delay of spillover effects due to ICT (see e.g. Brynjolffson and Hitt 2003, Greenan and Mairesse 2000).

As we will see in section 2.3, the first two properties of GPTs can be empirically confirmed for ICT. In addition to empirical studies, ICT has also been compared with other GPTs from the past. Jovanovic and Rousseau (2005) compare electricity and ICT in terms of their ability to generate economic growth. They conclude that, while electricity is more pervasive, ICT is more able to generate improvements and to promote innovation in all economic sectors. The authors conclude that these findings, in connection with falling ICT prices, will further increase the pervasiveness of ICT in the future.

Gordon (2000, 2002, 2012, 2016) is sceptical about the role of ICT as GPT. In comparison to previous industrial revolutions, the consequences of the "IT revolution" would be much more limited and, in his opinion, exhausted. The first industrial revolution (1750-1830) is originated by the first GPT, the steam engine. The second industrial revolution (1870-1900) was spun by the development of two GPTs, the electricity and the internal combustion engine as well as the development of running water with indoor plumbing. The third industrial revolution, also referred to as the "digital revolution", is mostly mentioned as being spurred on by development and diffusion of ICT. While, according to Gordon, the effects of the second industrial revolution lasted 81 years (from 1891 to 1972) in the U.S., those of the IT revolution lasted only 8 years (from 1996 to 2004). He states that the productivity-enhancing effects of ICT faded away by 2004, because since 2000 improvements have been made mostly in the performance of entertainment and communication devices, which "[d]o not fundamentally change labour productivity or the standard of living in the way that electric light, motor cars or indoor plumbing changed it" (Gordon 2012, p. 2).

To summarize, ICTs affect productivity both directly and indirectly. The direct effects result from productivity growth in the ICT-producing sector. Due to capital deepening and as an enabler of products, processes and organizational innovation, ICT also affects productivity indirectly in the ICT-using sectors. Several authors have qualified ICT as a GPT. However, there is controversy in the literature about whether ICT is a GPT, for which reason this question is also discussed as the "GPT hypothesis" in ICT-related literature. Because this section is merely intended to provide an overview on the impact of ICT to productivity and growth, the controversy on the GPT hypothesis is outlined, but not further discussed, as it is not relevant in the context of this dissertation. The work of Cardona et al. (2013) provides a detailed analysis of the research on the GPT hypothesis.

2.2 Empirical Approaches

In the literature there are two major methods that are commonly used to measure impact and contribution of ICT to productivity and economic growth. Since we also discuss the literature in

the review of section 2.3 in the context of the methods used, we briefly introduce the two main methods in the following.

The approaches to assess the impact of ICT on productivity growth can generally be distinguished to be parametric or non-parametric. The most common method is growth accounting, introduced by Abramowitz (1956) and Solow (1957). The growth accounting approach has been used and established in several studies to quantify the contribution of ICT to output and productivity growth, inter alia Inklaar et al. (2005), Jorgenson et al. (2005), Timmer and van Ark (2005) and Jorgenson and Timmer (2011). The growth accounting approach is introduced in subsection 2.2.1.

The estimation of production functions represents the parametric approach to investigate the contribution of ICT to productivity growth. An advantage of production function estimation over growth accounting lies in the performing of statistical significance tests of the estimation model and its determinants. The significance of the impact of ICT on productivity and economic growth was examined inter alia by Brynjolffson and Hitt (1995), Dewan and Kraemer (2000), Röller and Waverman (2001) and O'Mahony and Vecchi (2005). This approach of production function estimation is introduced in subsection 2.2.2.

In general, productivity describes the ratio of output(s) and inputs(s) required to generate the output(s). To measure productivity, different measures for inputs and outputs can be applied. The most common productivity measure is labor productivity, which relates output to labor input. Labor productivity is usually calculated by using the number of employees or the number of hours worked as labor input. Other measures use quality-adjusted hours worked (labor service) to additionally account the productivity effect of human capital. Similarly to labor productivity, capital productivity relates output to capital input, usually measured by capital stocks. Both labor productivity and capital productivity are partial factor productivity measures. The simultaneous use of labor and capital inputs yields Total Factor Productivity (TFP)[11] measures. The output measure is usually either gross output or value added. Gross output includes intermediate inputs, whereas in value added they are subtracted from gross output.

2.2.1 Growth Accounting

Growth accounting is based on the seminal papers by Abramovitz (1956) and Solow (1957) on technical change and the aggregate production function. The approach is extensively discussed in Aghion and Howitt (2007). Growth accounting provides a well-established approach to examine which part of the output growth of a country or an industry can be explained by growth in the inputs, while the residual is interpreted as a measure for the rate of unobservable technological progress. Growth accounting can differentiate between the different types of capital input. Thus, growth accounting allows a distinction between ICT and non-ICT capital. Several studies have made use of this approach, as will be discussed in section 2.3.

The growth accounting approach employs properties of production theory to determine empirical measures of the parameters of a production function by constructing economically defined index

[11] Total Factor Productivity (TFP) is also known as Multifactor Productivity (MFP).

numbers. Therefore, the validity of some neoclassical assumptions have to be postulated, like competitive factor markets, efficient producers, separability of inputs (see Aghion and Howitt 2007).

In the growth accounting approach, it is assumed that the overall economic output Y_{it} of a country or industry i at time t can be described by a production function with the input factors labor L_{it}, capital K_{it} and level of technology A_{it} as:[12]

$$Y_{it} = A_{it} \cdot F\left(K_{it}, L_{it}\right). \tag{1}$$

Most growth accounting exercises assume the existence of an aggregate production function of the Cobb-Douglas type, with constant returns to scale. This implies that we can write:

$$Y_{it} = A_{it} \cdot \left(K_{it}^{\alpha} L_{it}^{1-\alpha}\right). \tag{2}$$

Taking the (natural) logs of equation (2) and derivative with respect to time, gives us a function where rates of change in output are expressed in terms of rates of change in capital, labor and technological progress:[13]

$$\Delta \ln Y_{it} = \Delta \ln A_{it} + \alpha \Delta \ln K_{it} + (1-\alpha)\, \Delta \ln L_{it}. \tag{3}$$

The term $\Delta \ln A_{it}$ denotes the growth rate of total factor productivity, which comprises all effects on growth that cannot be explained by the factors labor and capital. The term α denotes the output elasticity with respect to capital, the term $(1-\alpha)$ denotes the output elasticity with respect to labor.

Subtracting $\Delta \ln L_{it}$ from both sides of equation (3) allows us to express the relationship between labor productivity growth and growth of TFP:[14]

$$\Delta \ln y_{it} = \Delta \ln A_{it} + (1-\alpha)\, \Delta \ln k_{it}, \tag{4}$$

where y_{it} denotes the output per worker (Y_{it}/L_{it}) and k_{it} denotes the capital per worker (K_{it}/L_{it}). Hence, growth in labor productivity can be increased by capital deepening and TFP growth.

Under the assumption that factor markets and product markets are perfectly competitive (so that factors are paid a return equal to their marginal product) and constant returns to scale, the output elasticities of capital and labor can be approximated by the respective factor income shares. The capital share v_{it}^{K} is derived as $v_{it}^{K} = \frac{r_{it} K_{it}}{Y_{it}}$, where K_{it} is the amount of capital, Y_{it} the national income and r_{it} the user cost of capital or marginal product of the capital. Hence, $r_{it} K_{it}$

[12] Here we assume a production function with Hicks-neutral technical progress.

[13] This procedure is commonly used in the growth literature to approximate growth rates. Formally, this approximation is given by $\frac{\Delta \ln Y_{it}}{\Delta t} = \frac{\dot{Y}_{it}}{Y_{it}} \approx \Delta \ln Y_{it}$. Barro and Sala-i-Martin (1992, 1995) justify this approach when studying economies in discrete time periods.

[14] The transition from equation (3) to eq, (4) is carried out by the following intermediate step: $\Delta \ln Y_{it} - \Delta \ln L_{it} = \Delta \ln A_{it} + (\alpha - 1)\, \Delta \ln L_{it} + (1-\alpha)\, \Delta \ln K_{it}$.

denotes the capital income. Analogously, the labor share v_{it}^L is derived as $v_{it}^L = \frac{w_{it}L_{it}}{Y_{it}}$, where L_{it} is the amount of labor employed, and w_{it} the labor compensation or marginal product of labor. Consequently, $w_{it}L_{it}$ represents aggregate labor income. Under these assumptions equation (3) can be rewritten as the following equation for output (or value added) growth:

$$\Delta \ln Y_{it} = \Delta \ln A_{it} + \bar{v}_{it}^L \Delta \ln L_{it} + \bar{v}_{it}^K \Delta \ln K_{it}. \tag{5}$$

The bars over the shares indicate that the respective shares are typically derived by averaging the weights over the two periods for which the growth is accounted. Note that by assuming constant returns of scale it is well-known that the weights add up to 1, i.e. $v_{it}^L = 1 - v_{it}^K$ as well as $\bar{v}_{it}^L = 1 - \bar{v}_{it}^K$.

Subtracting $\Delta \ln L_{it}$ from both sides of equation (5) allows to express the relationship in terms of per capita variables:

$$\Delta \ln y_{it} = \Delta \ln A_{it} + \bar{v}_{it}^K \Delta \ln k_{it}. \tag{6}$$

It is assumed, so far, that there is only one type of capital. To examine the role of ICT, a distinction between ICT and non-ICT capital is made:

$$\Delta \ln y_{it} = \Delta \ln A_{it} + \bar{v}_{it}^I \Delta \ln k_{it}^I + \bar{v}_{it}^N \Delta \ln k_{it}^N, \tag{7}$$

where $\Delta \ln k_{it}^I$ resembles the ICT capital deepening and $\Delta \ln k_{it}^N$ the non-ICT capital deepening per worker. The contribution of each input (except for term A) to labor productivity growth is derived by weighting each of the factor growth rates by their respective income shares $\bar{v}_{it}^I$, $\bar{v}_{it}^N$. Since TFP is not observable, the rate of change of TFP can be obtained as the residual from the specification of equation (7):

$$\Delta \ln TFP_{it} = \Delta \ln A_{it} = \Delta \ln y_{it} - \bar{v}_{it}^I \Delta \ln k_{it}^I - \bar{v}_{it}^N \Delta \ln k_{it}^N. \tag{8}$$

Hence, $\Delta \ln TFP$ describes the proportion of the output growth that cannot be attributed to the growth rates of inputs (labor and capital) and thus remains an unexplained residual. For this reason, TFP is called the Solow residual. It resembles "a host of unobservable factors that affect the improvement in overall efficiency of how output is produced" (Cardona et al. 2013, p. 133). Thus, effects of technical improvements are captured by TFP, which cannot directly be captured by quantity changes of capital and labor. For this reason, the TFP or Solow residual reflects a "measure of our ignorance" (Abramovitz 1956, p. 11).

The standard growth accounting approach has been modified in several studies. In recent publications, a distinction between ICT and non-ICT capital was made with respect to different depreciation rates of ICT-capital (see e.g. Collechia and Schreyer 2002). Instead of a raw labor indicator, some authors use a human capital variable to control for different educational levels and to distinguish between skilled and unskilled labor force (see, for instance, O'Mahony and Vecchi 2005, Papaioannou and Dimelis 2007).

Growth accounting provides a well-established and commonly used approach to examine how much output growth of a country or industry can be explained by growth in the different types of capital input. The approach allows for the quantification of the proximate sources of growth.

Despite the benefits, growth accounting raises some methodological problems. A main critique is that growth accounting simply seperates productivity growth into components without accounting for the underlying sources of growth (Baily 2002). Accordingly, the method is limited in its potential to draw policy conclusions. The growth accounting approach requires the consideration of inputs in the form of capital. In the case of ICT, this is not without problems. Due to technological progress in the mid-1990s, the information processing capacity of ICT increased at an exponential rate. The large quality improvements of IT equipment have led to rapid price declines (see e.g. Jorgenson 2001, Corrado and van Ark 2016). In the standard growth accounting, this development has two effects. Firstly, quality improvements lead to an increase in the TFP in ICT-producing sectors. Secondly, declining prices for IT goods of the same quality lead to a bias in the growth accounting results for the ICT-using sectors. To obtain unbiased results, the ICT prices have to be quality-adjusted (see section 3.2). An underestimation of the quality improvements in the price index, for example, would result in an overestimation of the productivity effect in growth accounting the ICT-using sector. In summary, the correct measurement of input(s) and output is crucial, since any error in measuring, for example the stock of ICT capital, will immediately affect the measured TFP index (i.e. the rate of change in TFP).

2.2.2 Production Function Estimation

The econometric approach of estimating a production function avoids the postulation of a theoretically based relationship between production elasticities and income shares. The output elasticities of input factors are directly estimated. In contrast to the growth accounting approach, assumptions about the producer behavior, and competitive factor markets do not have to be imposed. Furthermore, productivity estimates based on a production function may deviate from the strict concept of constant returns to scale.

The estimation procedure is used to determine whether the variables explain productivity growth significantly. In any case, a production function must be specified, which has commonly the form (see e.g. Brynjolffson and Hitt 1995, Dewan and Kraemer 2000):

$$Q_{it} = F\left(K_{it}^{N}, K_{it}^{I}, L_{it}; i, t\right), \tag{9}$$

where Q_{it} denotes the output of firm, industry or country i in period t. K_{it}^{N} is non-ICT capital, K_{it}^{I} is ICT capital and L_{it} is labor input. As functional form for $F(\cdot)$ the Cobb-Douglas production function is usually adopted. Applying and log-linearizing the Cobb-Douglas production function and so accepting the existence of constant returns to scale, one can derive a regression equation which can be estimated as follows (see e.g. Brynjolffson and Hitt 1995, Dewan and Kraemer 2000):[15]

[15] The use of the Cobb-Douglas production function is most common in the literature regarding the impact of ICT to productivity. There are also studies on firm level consider the translog function in addition to the Cobb-Douglas production function (see e.g. Brynjolfsson and Hitt 1995).

$$\ln Q_{it} = \alpha + \beta_1 \ln K_{it}^N + \beta_2 \ln K_{it}^I + \beta_3 \ln L_{it} + controls + u_{it}, \tag{10}$$

where Q_{it} typically represents the value added in firm or industry studies, while in country studies the GDP is commonly used. Furthermore, *controls* is a placeholder for a set of control variables, and u_{it} denotes the error term. The coefficients β_1, β_2 and β_3 of equation (10) resemble the output elasticities and correspond to the respective shares v of equation (5) in the growth accounting approach. The focus of most analyses, estimating production functions, is on the estimation and interpretation of the output elasticities, which measure the increase in output associated with a small increase in the corresponding input. For example, the output elasticity of ICT capital, β_2, represents the average percentage output increase associated with a 1% increase in ICT capital. The other elasticity parameters have analogous interpretations.

In most of the studies (mentioned in the subsequent section 2.3) the production functions are estimated with panel data, where the time period t is commonly measured in years. In studies on the firm level, usually time dummies or region and industry controls are added to the model (instead of the placeholder *controls* in equation (10)). Several studies control for the observational unit in fixed-effects models to capture any time invariant idiosyncratic productivity effect. Since it can be assumed that some firms, industries or countries are per se more productive, this unobserved heterogeneity, e.g. due to better management practices or market structure, can be captured by these models.

One benefit of estimating an elasticity instead of postulating it by economic theory is the possibility of testing its statistical significance. However, these tests on significance have their limitations due to possible endogeneity. It is plausible to assume that investment in ICT drives productivity but can also be a result of productivity and economic growth. Hence, the critique is that productivity estimations determine a correlation rather then causal effect on productivity (see e.g. Cardona et al. 2013). To address the problem of potential endogeneity, the simple regression model can be modified by using dynamic panel data models that utilizes lagged values of the ICT variable as instruments. On the firm level, this modification has been made e.g. by Brynjolffson and Hitt (1995). Alternatively, the estimation can be performed with a first-stage diffusion estimation, as done by Czernich et al. (2011).

Another benefit is that the estimation of production functions does not necessarily need the specification of ICT capital. Some studies use other variables, such as the penetration rate of telephone lines or the penetration of broadband internet (Czernich et al. 2011), as proxies for ICT capital.

2.3 Literature Review

In this section we provide an overview of the literature that examines the impact and contribution of ICT on productivity and economic growth. This necessarily incomplete and selective literature review focuses on empirical studies with output or productivity as the dependent variable. We exclude related topics that are not relevant in context of this dissertation. These are, for example, studies on consumer surplus of ICT, studies that discuss firm performance indicators or studies

that measure the effect of ICT on employment. We refer to literature surveys regarding these issues, such as Dedrick et al. (2003) and Cardona et al. (2013).

Empirical studies concerning the contribution of ICT to productivity and economic growth can be classified in several categories. Firstly, studies can be distinguished according to the definition, such as IT (hardware/software), communication (internet/broadband/mobile) and the combined form of ICT. Secondly, studies can be distinguished according to the empirical approaches of the non-parametric growth accounting approach and the parametric estimation of production functions, as presented in subsections 2.2.1 and 2.2.2. A third categorization results from the classification of studies according to the aggregation level. There, studies at the disaggregated firm level, studies at the industry level and research at the aggregated country level are distinguished.

In the following, we will present the research field according to the three aggregation levels and present their core results. Within each subsection, the research field is presented in chronological order. This serves to provide an overview of the research field and to identify gaps in the literature, which will be discussed in the subsequent summary.

2.3.1 Firm-Level Studies

The phenomenon of the productivity paradox was originally based on aggregated country-level. Since the beginning of research in this field, however, analyses have been published at the firm or plant level. In these early years, researchers usually examined only IT and excluded communication aspects. Since companies have an interest in exploring their return on IT investment, some publications are also available in the management literature.

The research in this field began in the 1980s and was motivated by the phenomenon of the productivity paradox and the well-known quote of Robert Solow. Previous studies at firm level could not show that IT investments lead to payoffs (see e.g. Loveman 1994, Brynjolfsson and Hitt 1996). Dedrick et al. (2003) states that this is due to inadequate data integration of IT investment and small sample sizes by that time.

Since 1993, studies with larger sample sizes have been published, such as Brynjolfsson (1996), Brynjolffson and Hitt (1995, 1996, 1998) or Lichtenberg (1995). These include data for more than 300 large U.S. companies within the period 1987-1994 and measure the contribution of IT capital investment and IT labor to output. These studies (like almost all firm-level studies) estimate a production function, derived from a Cobb-Douglas production function. The firm output is measured as value added (per employee), the input set includes labor hours, the IT and non-IT capital stock. As a result they estimate the marginal output elasticity of IT capital, which means the increase in value added associated with a 1% increase in IT investment. Each of these studies confirms a positive and significant contribution of IT investments to firm productivity. Moreover, they find investments in IT to have higher gross marginal returns than non-IT investments.

In addition to the studies of U.S. companies, investigations have been published for companies in other countries. For example, Greenan and Mairesse (2000) analyze the impact of IT investments

on productivity for French companies. The results are consistent with those of the U.S. studies by Brynjolfsson and Hitt (1996) and Lichtenberg (1995).

There are several explanations for the fact that previous studies could not find any relationship between IT investments and productivity. Schreyer (2001) suggests that the small amount of IT capital in the early 1990s could be an explanation for the missing impact. Dedrick et al. (2003) suggests that firms had to learn over time how to apply IT capital more productively.

Brynjolfsson and Hitt (2000) emphasize that organizational changes in firms need to be made in order to implement IT efficiently. They state that "a significant component of the value of information technology is its ability to enable complementary organizational investments such as business processes and work practices; [...] these investments, in turn, lead to productivity increases by reducing costs and, more importantly, by enabling firms to increase output quality in the form of new products or in improvements in intangible aspects of existing products like convenience, timeliness, quality, and variety" (p. 24). The costs of these organizational investments are firm-specific and would in some cases even exceed the investment in IT. In addition to the organizational changes, other complementary or simultaneous efforts are mentioned which have an influence on the productivity benefits of IT. Bresnahan et al. (2002) and Brynjolffson et al. (2002) referred to the skills of the workforce in this context.

The results of the studies reveal that since the mid-1980s higher IT investment can be associated with higher firm productivity. The productivity effects of IT investments vary between different companies. Brynjolffson and Hitt (1995) estimate that these "firm effects" account for about half of the productivity benefits.

More recent studies, such as Brynjolfsson and Hitt (2003) find that the effect of computerization is greater in the long-run (five years or more) than in the short-run (one year). Analyzing a panel of 527 U.S. firms in the period 1987-1994, they find that the productivity and output contributions associated with the level of computerization are up to five times greater over long periods. The authors suggest that the observed contribution is accompanied by relatively large and time-consuming investments in complementary inputs, such as organizational capital.

Recent studies also examine communication aspects and thus ICT. Van Reenen et al. (2010) use firm-level data of 13 European countries in the observation period of 1998-2008 to investigate the impact of ICT capital on labor productivity. They find that a 10% increase of ICT capital is associated to an increase of 0.23%-0.9% in output.

2.3.2 Industry-Level Studies

While firm-level studies focus on a specific industry or compare sectors (e.g. manufacturing vs. service), there are several studies based on aggregated industry-level data. In the late 1990s, after two decades of productivity slowdown, the U.S. experienced a period of increasing productivity growth. At the same time, industry-level data became available for the U.S. and several industry-level studies were published. These studies at industry-level focus on whether payoffs from IT

have taken place in a large number of industries or only in specific sectors (e.g. the IT-producing industry). In most cases, these studies use the empirical approach of growth accounting.[16]

In the context of the U.S. productivity revival in the late 1990s, a number of studies have found that labor productivity has accelerated in many industry sectors (see e.g. Jorgenson and Stiroh 2000). Comparing industrial sectors in the 1990s, Stiroh (2002b), for example, finds labor productivity shifts in two thirds of the 61 industries. He also finds that IT-intensive industries (with a higher than average level of IT capital) have 1.3% higher labor productivity acceleration than other industries.

Gordon (2000), however, finds an increase in labor productivity only in the durable goods manufacturing sector and most of that in the IT-producing industry.[17] For the 1995-1999 period, Jorgenson (2001) attributes even two-thirds of the growth in TFP to the IT-producing industry. Overall, there is considerable agreement in the literature that TFP has increased in the IT-producing industries (besides the aforementioned, see e.g. Jorgenson and Stiroh 2000, Oliner and Sichel 2002). This TFP increase in the IT-producing sector has contributed to the TFP growth in the U.S. proportional to its size in the economy. This is evidence for the direct productivity-enhancing effect of IT as mentioned in section 2.1.

While there is consensus about TFP growth in the IT-producing industry, there is a controversial debate in the literature on whether there has been an acceleration of TFP growth in the IT-using industry. For the post-1995 period, most studies find TFP growth in both the IT-producing as well as IT-using industry (see e.g. Jorgenson and Stiroh 2000, Oliner and Sichel 2002). Motivated by this controversy, Triplett and Bosworth (2004, 2006) and Bosworth and Triplett (2007) explicitly examined the service industry. Analyzing productivity in 27 industries in the service sector with the growth accounting approach, Triplett and Bosworth (2006) find that productivity improvements in the service industry can be explained by TFP growth and IT capital deepening. In their analyses of the post-1995 U.S. performance, they find that service industries (such as wholesale, retail and finance), which account for about 80% of the increase in IT capital in the U.S., are responsible for the acceleration in aggregated TFP after 1995. Corrado et al. (2007) also find a positive connection between IT capital deepening and productivity acceleration in the service sector in the late 1990s U.S. Similar findings for the role of the IT-using sector in the EU is provided by the investigation of van Ark et al. (2008). Hence, there is evidence for the indirect productivity-enhancing effect of IT due to spillovers between IT-producing and IT-using industries as mentioned in section 2.1.

Growth accounting allows to assess the contribution of ICT to labor productivity in percent. By adding up the effect of ICT capital deepening ($\bar{v}_{it}^{I} \Delta \ln k_{it}^{I}$) and TFP growth in the ICT-producing industries (weighted by the industry share) and dividing it by labor productivity growth, the contribution of ICT to labor productivity in percent can be determined. This also allows to compare the results of accounting studies. For the U.S., Jorgenson et al. (2008) find that the

[16] There are only few studies using a different approach, like Dimelis and Papaioannou (2011). The authors use system GMM and the pooled mean group panel data estimators to investigate the growth impact of ICT in the U.S. and the EU industries over the period 1980-2000.

[17] In a later study (also covering the year 2000), Gordon (2001) also finds acceleration in labor productivity in other sectors.

contribution of ICT to labor productivity grew from 11% in the period 1959-1973 to 43% in the period 1973-1995. In the post-1995 period the contribution further increases to 59% (1995-2000). Although labor productivity in the U.S. remained relatively high at an average annual growth rate of 2.5 in the period 2000-2006, the contribution of IT capital deepening as well as TFP growth in the IT-producing sector declined to 38%. Simultaneously, the TFP outside the IT-producing sector and thus in the IT-using sectors increased. Hence, the results seem to indicate for the U.S. that after 2000, the (direct) effects of the IT-producing sector declines and the contributions from TFP and capital deepening in IT-using sectors become more relevant. This is evidence for the time delay of spillover effects due to IT and hence indicators for the GPT hypothesis of I(C)T.

2.3.3 Country-Level Studies

Research at aggregated country level addresses two issues. On the one hand, it determines the contribution of IT, Communication or ICT to productivity and economic growth in certain economies. On the other hand, the research examines whether this contribution differs between countries.

The phenomenon of the productivity paradox was originally based on aggregated country-level data. Similar to the investigations at firm and industry level, early country-level studies also show low or no contribution of IT/ICT to productivity and economic growth (Oliner and Sichel 1994, Jorgenson and Stiroh 1995). As one of the first, Oliner and Sichel (1994) investigated the influence of ICT[18] capital on labor productivity in the period 1970-1992 with aggregated data for the U.S. They find only a small contribution of ICT capital to labor productivity, due to the low level of ICT capital at that time. According to Sichel (2001), the proportion of IT investments in proportion to the capital stock was too small to show substantial economic effects. While the share of IT capital in U.S. capital investments was 3.5% in 1980 and 9% in 1990, the share of IT capital increased to 22% during the 1990s.

Jorgenson and Stiroh (1995, 1999) complement the studies by Oliner and Sichel (1994). They argue that the massive price decline for computers in the 1980s and 1990s lead to a systematic underestimation of IT capital. By using a constant-quality price index for computing equipment, the authors compensate the effect of the price decline. Jorgenson and Stiroh (1995) show that the impact of IT on U.S. productivity is higher but not major. For the period 1985-1992 they find IT investments to be associated with 0.5 percent incremental economic growth. For the period 1973-1995, Jorgenson and Stiroh (2000) estimate a 13% ICT contribution to the annual GDP growth rate of 3.0 and a 27% ICT contribution to the 1.4 annual labor productivity growth rate.

In contradiction to their paper in 1994, Oliner and Sichel (2000) find the increase in productivity growth since the mid 1990s is due to capital deepening, mainly in ICT equipment. Furthermore, the growth of labor productivity is also due to efficiency gains in the production of ICT. They conclude that "information technology accounted for about two-thirds of the step-up in labor

[18] They consider information processing equipment, containing computer and peripheral as well as communications equipment.

productivity growth between the first and second halves of the decade" (p. 21). Other investigations from the late 1990s confirm the relevance and contribution of IT/ICT for the growth acceleration in the U.S., such as Jorgenson et al. (2002), Oliner and Sichel (2002) and Daveri (2003).

Van Ark et al. (2008) investigate the period 1973-2006 and examine the contribution of ICT capital deepening and TFP to labor productivity growth in the EU. In their study, they document that the average annual growth rate of hourly labor productivity in the EU was 2.4% over the period 1973-1995, which is twice as high as the rate in the U.S over the same period. In the subsequent period of 1995-2006, this revolves. In the U.S, the average annual growth rate in this period was 2.3%, while it was only 1.5% in the EU. Consequently, there was a productivity gap between the U.S. and EU in both periods.[19] A possible explanation for the increased productivity gap between the EU and U.S. in the post-1995 period can be found in the role of ICT.

The argumentation of Jorgenson and his co-authors is that the EU could not benefit from the effects of ICT capital deepening and TFP growth in the ICT producing sector to the same extent as the U.S. A possible reason for this may be the lower (average) level of IT investment in the EU compared to the U.S, as Daveri (2002) suggests. In his study, he identifies significant differences between the different EU countries. He finds high investment rates for the Netherlands, U.K. and Spain, medium investment rates in ICT for France and Germany and low rates in Greece, Italy, Portugal and Spain. According to Daveri (2002), there is a high correlation between investment rates and ICT's contribution to productivity growth (by using growth accounting). Other investigations, such as Inklaar et al. (2008), also find that the differences in the ICT investment level are an explanation for the post-2000 productivity gap.

The reason why there is less investment in ICT in the EU and why there is a lower return on investment is investigated by van Ark and co-authors. In several papers they compare the aggregated labor productivity gap between the EU and U.S. in the different time periods. They find ICT capital deepening in the EU to be considerably lower than in the U.S. According to van Ark et al. (2008), the contribution of ICT to the growth of labor productivity in the EU decreased from 72% (1980-1995) to 36% (1995-2004), while in the U.S. the contribution of ICT doubled from 33% to 59%. Also, van Ark et al. (2008) identify large differences between the individual countries of the EU (as similar to Daveri 2002).[20]

While the contribution of IT/ICT to productivity and economic growth has been widely explored in the U.S., comparative studies for the rest of the world are scarce. The reason for this is mainly due to the limited availability of national and comparable data. Hence, there are only few cross-country studies in the literature.

In addition to the previously mentioned studies for selected countries of the EU by van Ark and co-authors, there are also a number of studies considering countries in the OECD. They

[19] Van Ark. et al. (2008) hypothesize that the cause of the decline in hourly labor productivity is to be found in labor market rigidities.

[20] While Finland and the U.K. show similar labor productivity growth rates as the U.S., other countries such as Italy and Spain are far below. A sectoral analysis by van Ark et al. (2008) reveals that the share of the service sector in the EU is on average lower than in the U.S. Countries with a larger service sector also tends to have a greater ICT contribution to labor productivity growth (such as the Netherlands and the U.K.).

also use the growth accounting approach. As one of the first, Schreyer (2000) investigated the contribution of ICT investment to productivity growth in the G7 countries[21] (as subset of the OECD) in the period 1980-1996. He finds that ICT investment significantly explains growth in all seven countries, though the magnitude differs across the countries. Colecchia and Schreyer (2002) examine the contribution of ICT capital investment to economic growth in nine OECD countries in the 1980-2000 period. They find that in the early-1990s, ICT contributes between 20-50% to economic growth. The contribution increased in the late-1990s to 30-90%, whereby the contribution rates differ between the countries. They find high rates for Australia, Finland and Canada. Low rates were found for Germany, Italy and Japan.

Studies with larger data samples find differences in the contribution of ICT to productivity and economic growth across countries of different development status. Examining the pre-1995 period, Dewan and Kraemer (2000) find positive and significant returns from IT-capital in developed countries, but no substantial returns for developing countries. They suggest that the gap was due to a low IT capital stock (relative to GDP) in developing countries and to the lack of complementary assets, such as infrastructure and human capital. Pohjola (2002) does not find any significant relationship between ICT and GDP growth in the two subgroups of developing and developed countries.

While the above-mentioned cross-country studies use growth accounting, the studies of the post-1995 period almost all estimate production functions. These studies also come to different conclusions. Papaioannou and Dimelis (2007) find positive growth effects of ICT in both developing and developed countries, whereby the impact in developed countries was found to be higher. Yousefi (2011) investigates the period 2000-2006 and finds a major role of ICT in growth of high and upper middle countries, but not in lower middle income countries. Hence, there is mixed evidence of the impact of ICT on productivity and economic growth in developing countries.

Among the above-mentioned authors of cross-country studies, Yousefi (2011) has the largest data sample, with 62 countries. The study by Becchetti and Adriani (2005) has an even larger sample (up to 92 countries). Instead of measuring ICT monetarily in terms of capital investment, the authors use communication components (such as telephone lines and internet hosts) as a proxy for ICT. These components are also available for a range of developing countries and a large investigation period.

According to information technology, there are also several studies available that are related to the economic impact of communication technologies. Telecommunication infrastructure, for example, is investigated by Röller and Waverman (2001). They examine 21 OECD countries in the period 1970-1990 and find evidence for a significant and positive impact of telecommunications infrastructure to economic growth. The effect of broadband infrastructure on economic growth was examined by Koutroumpis (2009). He investigates the economic impact of broadband penetration in 22 OECD countries for the period 2002-2007. As a result Koutroumpis (2009) finds a significant causal positive link between the broadband penetration and economic growth, especially when a critical mass of infrastructure is present. In a similar study, Czernich

[21] The sample contains Canada, France, Western Germany, Italy, Japan, United Kingdom and the United States.

et al. (2011) find that a 10-percentage point increase in broadband penetration raised annual per capita growth by 0.9-1.5 percentage points.

2.4 Summary

In this chapter, we presented the current state of research on the contribution of ICT to productivity and economic growth. From a theoretical point of view, ICT affects productivity and growth directly and indirectly. Direct effects occur due to technological improvements in the ICT-producing sector and thus to the aggregated productivity of an economy proportional to the sector share. Due to capital deepening and as an enabler of products, processes and organizational innovation, ICT also affects productivity indirectly in the ICT-using sectors. Some authors have examined whether ICT is a GPT. Although ICT seems to satisfy all the characteristics of a GPT, there are controversial positions in the literature.

The impact and contribution of ICT on productivity and economic growth is examined in the literature mainly by using two methods. These are the non-parametric growth accounting approach and the parametric production function estimation. Growth accounting examines how much output growth of a country or industry can be explained by growth in the different types of capital input. It employs properties of production theory to determine empirical measures of the parameters of a production function by constructing economically defined index numbers. By using the parametric approach of estimating a production function, output elasticies of input factors are directly estimated by statistical methods.

Both approaches have their advantages and disadvantages. Growth accounting provides a well-established and commonly used approach that allows for the quantification of the proximate sources of growth, but without accounting for the underlying causes. The growth accounting approach requires the consideration of ICT input in the form of capital, which can lead to biases in the case of incorrect measurement. The parametric approach avoids most of the neoclassical assumptions of the non-parametric growth accounting. Furthermore, it enables the estimated elasticities to be tested for their statistical significance. However, the approach obliges its user to specify a functional form of a production function.

We reviewed the empirical literature on the contribution of ICT to productivity and economic growth on firm, industry and country level. Besides these aggregation levels, the studies also differ according to their definition of IT, communication and combined form of ICT as well as the applied empirical method. The role of ICT to productivity and economic growth has been investigated in an abundance of studies since the 1980s. The literature review reveals that the contribution of ICT to productivity and economic growth has been investigated primarily in the U.S. Studies from the 1980s and early 1990s show no or only low contribution of IT to productivity and economic growth. This changed considerably in the late 1990s, when the U.S. experienced a period of increasing productivity growth after two decades of productivity slowdown. The majority of researchers agree on the importance of ICT for the U.S. growth resurgence. A main reason for that is seen in the capital deepening since the mid-1990s, which was promoted by the massive price decline for IT/ICT equipment. This price decline triggered higher investment in ICT and substitution of less productive inputs in the ICT-using sectors.

Several authors investigated the effects of TFP growth in IT-producing industries and role of IT capital deepening in IT-using industries. The overall pattern of studies on industry level show a high contribution of IT investments to TFP growth, especially in the period from 1995-2000. Further studies point to the role of IT in explaining productivity growth in the IT-using sectors. They show that TFP growth has also increased in the ICT-using sector, especially in industries that use IT more intensively. This provides evidence for the hypothesis that I(C)T is a GPT.

A comparison of U.S. results with other countries reveals differences in the impact of ICT on productivity and economic growth. A key result is that the EU could not benefit to the same extent from the effects of ICT capital deepening and TFP growth in the ICT-producing sector as the U.S. Moreover, there are considerable differences between EU countries. These differences are also shown by studies comparing countries at different stages of development. While most post-1995 investigations show that ICT has a significant positive impact on developed countries, some studies show that ICT has only a minor or no impact on developing countries.

While the economic impact of ICT has been widely explored in the U.S., cross-country studies covering a sufficient number of countries at all stages of development are rare. The results of these studies are also contradictory and show that the findings from the empirical literature for the U.S. do not necessarily apply to other countries. Furthermore, the reasons for the differences between developing and developed countries in the impact of ICT on productivity and economic growth are still largely unexplored. Since there are only a few studies which include a sufficient number of countries at all stages of development, this dissertation extends the current research field with three scientific contributions on the country level.

3 Information and Communication Technology

This chapter addresses the definition and quantitative measurement of ICT. The chapter is divided into four sections. In section 3.1, we describe the terminology and definition of ICT. In section 3.2, we review the situation of available data for measuring ICT. We discuss the disadvantages of using capital stocks to describe ICT and motivate the usage of non-monetary penetration rates of ICT infrastructure as an appropriate proxy for ICT. Subsequently, we construct the ICT variable from the penetration rates. In section 3.3 we conduct a descriptive analysis of the constructed variable. Section 3.4 summarizes this chapter.

3.1 Terminology and Definition of ICT

The acronym ICT stands for 'Information and Communication Technology', a phrase that has been used by academic researchers since the 1980s. Sometimes, ICT is used synonymously with IT ('information technology'). In contrast to IT, however, the acronym ICT additionally covers the area of communication and thus a broader, more comprehensive list of components related to connected computers and digital technologies.

A clear and unequivocal definition of what precisely ICT refers to cannot be found in the scientific literature. The available literature covers ICT in the technological, economic, occupational and cultural dimensions. The complexity of the ICT concept derives partly from the influences of the terms 'information age' (see for instance Castells 2011), 'digital age' (also 'digital revolution', see e.g. Brynjolfsson and McAfee 2012) and 'information society' (see e.g. Webster 2014). In determining and defining ICT in the context of this dissertation, the scope is limited to the economic and technological aspects.

It can generally be assumed that the first definition of information technology is that of Leavitt and Whisler from 1958, right at the beginning of the information age:

> "The new technology does not yet have a single established name. We shall call it information technology. It is composed of several related parts. One includes techniques for processing large amounts of information rapidly, and it is epitomized by the high-speed computer. A second part centers around the application of statistical and mathematical methods to decision-making problems; it is represented by techniques like mathematical programming, and by methodologies like operations research. A third part is in the offing, though its applications have not yet emerged very clearly; it consists of the simulation of higher-order thinking through computer programs" (Leavitt and Whisler 1958, p. 41).

Comparing the definition with the more modern definition of Davis and Hamilton from 1993, it can be argued that the definition from the 1950s is essentially still applicable today. In addition to the description of hardware and software, the aspect of telecommunications has been added:

> "Information technology refers broadly to the technology of computers and electronic communications as applied to processing, transfer, and storage of information. It

> encompasses computer hardware, data communications, software, and a large variety of input and output devices. Local area and wide area communications network for information transfer are also included" (Davis and Hamilton 1993, p. 21).

Many other definitions describe the concept of information technology in a similar way, including hardware and software components as well as communication technology. The common naming of information and communication technology does not seem to be absurd. These terms are also referred to as Siamese twins, meaning that information is inconceivable without communication and vice versa. ICT can be seen as an extended term for information technology (IT) which stresses the role of unified communications. According to Technopedia, an IT Dictionary for Computer Terms and Tech Definitions, ICT "refers to all the technology used to handle telecommunications, broadcast media, intelligent building management systems, audiovisual processing and transmission systems, and network-based control and monitoring functions".[22] Thus, the term ICT does not merely describe an extension of the IT concept, but also refers to the convergence of telecommunications (telephone lines and wireless signals) with computer networks (including software used) using a single unified system of cabling or link system.

The Organization for Economic Co-operation and Development (OECD) has been striving since the mid-1990s to establish and enforce an internationally standardized and accepted definition of the ICT sector. Rather than classifying ICT-related products, the OECD initially defines ICT-related sectors and industries on the basis of the international standard classification of activities (ISIC). A first delimitation of ICT was defined in 1998 by the WPIIS [23] (Working Party of Indicators for the Information Society) on the basis of the ISIC classification and published in 2000 (OECD 2000). From the perspective of the OECD, ICT refers to the combination of manufacturing and services industries that capture, transmit and display data and information electronically (OECD 2002).

Since 2002, the OECD has reconvened discussions regarding the sectoral definition of ICT and has clearly delineated manufacturing from services industries. The current delimitation from 2006 is based on the ISIC classification Rev. 4 and categorizes the ICT industry into the subdivisions of manufacturing, trade and services industries.[24] In addition to ICT, a delimitation of a content and media sector was published in 2007, which contains industries that are concerned with the production, publication and/or electronic distribution of media content.[25]

In 2003, the OECD also developed a classification of ICT goods that was complemented by a classification of ICT services in late 2007 as well as content and media products in 2008. The

[22] https://www.techopedia.com/definition/24152/information-and-communications-technology-ict.

[23] This is an OECD working group that was set up in 1997 under the newly established ICCP (Information, Computer and Communications Policy) statistical panel.

[24] The current sectoral definition of ICT from the OECD, based on the ISIC classification Rev. 4, is given in table B1 of the appendix.

[25] The OECD defines media content as follows: "Content corresponds to an organized message intended for human beings published in mass communication media and related media activities. The value of such a product to the consumer does not lie in its tangible qualities but in its information, educational, cultural or entertainment content" (OECD 2009). This content and media sector includes products such as printed or otherwise transported media, moving images, television or radio content, software games or music (OECD 2009). In the revision of 2008, the content and media sector was grouped together with the ICT sector under the term information economy.

classification uses the Central Product Classification (CPC) code of the UN. In the latest version, the classification of ICT goods, services and media products are combined in an information economy product classification that was released by the WPIIS in 2010. For the detailed sectoral and product definitions we refer to the recent publication in OECD (2011).

The elaborate definition of the ICT sector, products and services by the OECD is seldom used outside the official statistics. Actually, most of the scientific works concerning ICT in the economic sense do not provide a clear and unambiguous specification of that term. It appears to be assumed that the definition and delimitation of ICT is obvious and does not need to be defined explicitly. Accordingly, the subjects of research differ widely in the economic and scientific literature. For example, Oliner and Sichel (2002) use the sum of final computer hardware, communication equipment and software sales as well as the sales in the semiconductor sector (from various, partly unpublished sources) to reflect the ICT. Jorgenson and Timmer (2011), by contrast, utilize the aggregated output of companies that belong to the ICT sector.[26] These also include manufacture of televisions and postal services, and are surveyed by national statistics offices in the EU. Brynjolffson et al. (2002) consider the investment in computers as a proxy of IT. Other proxies are e.g. the number of internet hosts (Guillén and Suárez 2001, Kiiski and Pohjola 2002), the number of personal computers (Chinn and Fairlie 2007) or the internet diffusion rates (Wunnava and Leiter 2009).[27]

In summary, we can conclude that a concrete and unambiguous definition of ICT does not exist. The establishment of a consistent and universal definition seems difficult. The example of the OECD shows that it is challenging to define ICT as a list of components. A list of ICT components is exhaustive, and it continues to grow over time. While computers and telephones have existed for decades, other components such as smartphones or e-commerce have only been prevalent for a few years or will only be released in the years ahead, such as nanotechnology. ICT continues to insinuate and alter itself in the ever-changing globe. For this reason, a definition of ICT in the form of an abstract concept seems to be more appropriate.

In an abstract form, ICT can be described as a diverse set of technologies and resources used to communicate, create, disseminate, store and manage information. These technologies have the characteristic that they provide access to information through telecommunications. The plumbing analogy of Davenport and Prussak (2000) can be used for better understanding: In a plumbing system of pipes and storage tanks, water is stored in the storage tanks and flows through the pipes. In this analogy, information technology is represented by the storage tanks and communication technology by the pipes. Using information technology, information (stored, non-flowing water) is communicated (flowing water) though communication technology.

[26] The sectoral definition is based on the Statistical Classification of Economic Activities in the European Community, Rev. 1.1 (2002) (NACE Rev. 1.1) and includes codes 30-33 and 64, which are 'Manufacture of office machinery and computers', 'Manufacture of electrical machinery and apparatus n.e.c.', 'Manufacture of radio, television and communication equipment and apparatus', 'Manufacture of medical, precision and optical instruments, watches and clocks, Post and telecommunications'.

[27] The various indicators and proxies used to describe ICT will be outlined in detail at a later stage in section 3.2.

3.2 Data Situation and Measurement of ICT

We now review the situation of available data to measure ICT. As previously mentioned in section 2.3, ICT is commonly considered by using data on IT/ICT capital. From the beginning of research activities in this field, capital data was one of the few available indicators, because it was collected from official bodies for fiscal and tax reasons.

Firm-level studies mostly use data from private data suppliers. For example Brynjolffson and Hitt (1995, 1996, 2000) and Lichtenberg (1995) used data from the International Data Group (IDG) that cover IT-spending of over 300 large U.S. firms. Until 1995, the Computer Intelligence InfoCorp also provided computer capital data, as used for example by Bresnahan et al. (2002).

Studies at the aggregated industry level for the U.S. largely use data from the official U.S. statistics, provided since the mid-1990s. The U.S. Bureau of Economic Analysis (BEA) provides the U.S. industrial database, which contains (inter alia) data on expenditures for ICT equipment based on the North American Industry Classification System (NAICS). These data have been used in studies by Bosworth and Triplett (2007), Corrado et al. (2007) and Jorgenson et al. (2008), for example.

Most studies at the country-level used private data suppliers, such as the the International Data Corporation (IDC), World Information Technology and Services Alliance (WITSA) or the European Information Technology Observatory (EITO). These private data sources have been used, for example, by Schreyer (2000) and Daveri (2002). Others, such as Collechia and Schreyer (2002), collected the data on expenditures in IT and communication equipment from the national statistics. However, national statistics differ substantially in terms of available aggregates. Some countries (including the U.S.) only publish data from private investment, while other countries (such as Canada) also include government expenditure on ICT. The official national statistics also differ considerably in terms of delimitation of ICT. Some countries only account for purchased software, others countries also include own-account software (see Collechia and Schreyer 2003 for an overview). Thus, the differences in the ICT classification of the national statistical offices impede cross-national comparisons. Harmonized European data on ICT (and non-ICT) capital inputs is provided by EU KLEMS Productivity and Growth Accounts from the Groningen Growth and Development Centre. Their database contains data on capital formation as well as capital stock, differentiated by computing equipment, communications equipment and software. The EU KLEMS database is (inter alia) used by Dimelis and Papaioannou (2011).

The monetary consideration of ICT raises some problems in the context of the commonly-used methodological approaches. In the mid-1990s the information processing capacity of IT increased at an exponential rate due to technological progress in the production of IT goods. The large quality improvements of IT equipment led to rapid price declines (see e.g. Jorgenson 2001, Corrado and van Ark 2016). In other words, the capacity of computers at that time was significantly higher than that of computers at the same price ten years before. In the literature, using traditional growth accounting, this development has been reflected in TFP growth in the IT producing sectors and IT capital deepening in IT/ICT-using sectors (see section 2.3).

In order to capture these effects of price declines in ICT equipment, quality-adjusted price indices have been developed, for example from the BEA. These price indices should be constant

quality deflators that reflect price changes for a given performance of IT/ICT investment goods. The construction of price indices is crucial. If the quality improvements are not fully reflected in a price index, this leads to a bias in the measurement of productivity growth. In growth accounting, for example, all effects of quality improvement in ICT that are not fully captured by the price index appear in TFP. Price deflators for IT had suffered in the early days from inappropriate measurement and were usually assumed to underestimate the quality improvements due to incomplete quality adjustments (Cardona et al. 2013, Baily et al. 1998, Griliches 1994). In the context of studies at country level, the importance of adequate price indices is particularly high, as they also have to deal with different national currencies. In the literature, various ways of constructing a price index are presented, such as the "hedonic" price index of Schreyer (2000), which captures the price change in various types of ICT capital goods in the OECD countries. Furthermore, the EU KLEMS database provides harmonized European data.

There are a number of alternatives to the monetary expression of ICT. Several authors use the number of Personal Computers (PC) to consider ICT (Loveman 1994, Greenan and Mairesse 2000). One can argue that the number of PCs is correlated with other ICT spending, such that this variable is an appropriate proxy for ICT. This measure can be complemented by adding the expenditures for ICT staff (Brynjolfsson and Hitt 1995, 1996).

Other studies that investigate communication technologies use penetration levels. Hardy (1980) as well as Röller and Waverman (2001) use the number of telephone lines, Koutroumpis (2009) use the level of broadband penetration in 100 inhabitants. Becchetti and Adriani (2005) further use the number of internet users as well as the number of mobile phones.

In the previous section above we characterize ICT as technology that provides access to information through telecommunications. Spoken in the plumbing analogy: without pipes (communication infrastructure) there is no information flow. Hence, a communication of information without a communication technology is impossible. The penetration rate defines both the capacity and accessibility of information in an economy. Where communication infrastructure is rudimentary, communication between companies is limited. In these cases "[t]he transaction costs of ordering, gathering information, and searching for services are high." (Röller and Waverman 2001, p. 910).

In contrast to monetary indicators, infrastructure proxies do not require any specification of goods, services or sectors of ICT. New products resulting from technological changes, such as ICT services e.g. in the form of cloud computing, do not have to be incorporated into national statistics but immediately increase the demand for the required communication infrastructure (broadband connections in this case).

For these reasons, the penetration rates mentioned above are recommended as appropriate proxies for the measurement of ICT. These penetration rates have the further advantage that they are available for a broad set of countries. They are therefore suitable for investigations at country level, as intended in this dissertation.

A single encompassing and non-monetary variable can be found neither in the literature nor in publicly accessible databases.[28] Instead, five non-monetary variables are available, describing single isolated aspects of ICT.[29] These are:

1. **The number of (fixed) telephone lines**: The telephone network represents a basic ICT technology, as first interconnections were established via this technology carrier. For this reason, the number of telephone lines is available for a number of countries even as far back as the 1960s.

2. **The number of internet users:** This variable captures the individuals who use the internet via a computer, mobile phone, personal digital assistant, games machine, digital TV or other. Hence, the access can be via a fixed or mobile network.

3. **The number of broadband internet subscribers**: This variable captures the number of subscribers with a digital subscriber line, cable modem, DSL or other high-speed technology with a downstream speed equal to, or greater than, 256 kbit/s. Broadband technology is particularly used for large data volumes and high transmission rates.

4. **The number of mobile cell subscribers**: This variable captures the number of subscriptions to a public mobile telephone service using cellular technology, which provide access to the public switched telephone network. Included are postpaid as well as prepaid subscriptions.

5. **The number of Personal Computers (PCs)**: The variable is an estimate by the United Nations of the number of PCs based on the number of broadband connections and telephone lines.

The data for the former four variables are from the World Telecommunication/ICT Development Report of the International Telecommunication Union (ITU).[30] The variables can be downloaded from World Bank database[31] and are available as both absolute as well as proportional values for a broad set of countries since 2001. Before that time, values for the number of broadband internet subscribers (per 100 people) are rarely available because of the relatively new broadband technology. The number of PCs is part of the Human Development Report ('Table 12: Innovation and Technology') and only available as averaged value over the period 2002-2009.

Figure 3.1 illustrates the development of the first four variables from 2001 to 2012.[32] We calculate the worldwide average for each variable and year. Furthermore, we calculate the average values for each development group of countries to illustrate the differences in penetration. We use the

[28] Since commercial data resources are not part of our analysis, we focus on the availability of appropriate variables in publicly available databases.

[29] There are also other non-monetary variables available, such as the number of secure internet servers. However, these variables are not suitable as proxies for capturing ICT, since they do not provide information about the intensity of ICT usage in a particular country. The number only indicates the number of frequent server locations.

[30] For further description and definition of these data, see homepage of the International Telecommunication Union: http://www.itu.int/en/ITU-D/Statistics/Pages/publications/handbook.aspx.

[31] http://databank.worldbank.org/data.

[32] Since the number of PCs is only available as an average for the period 2002-2009, we are unable to illustrate any development at this point.

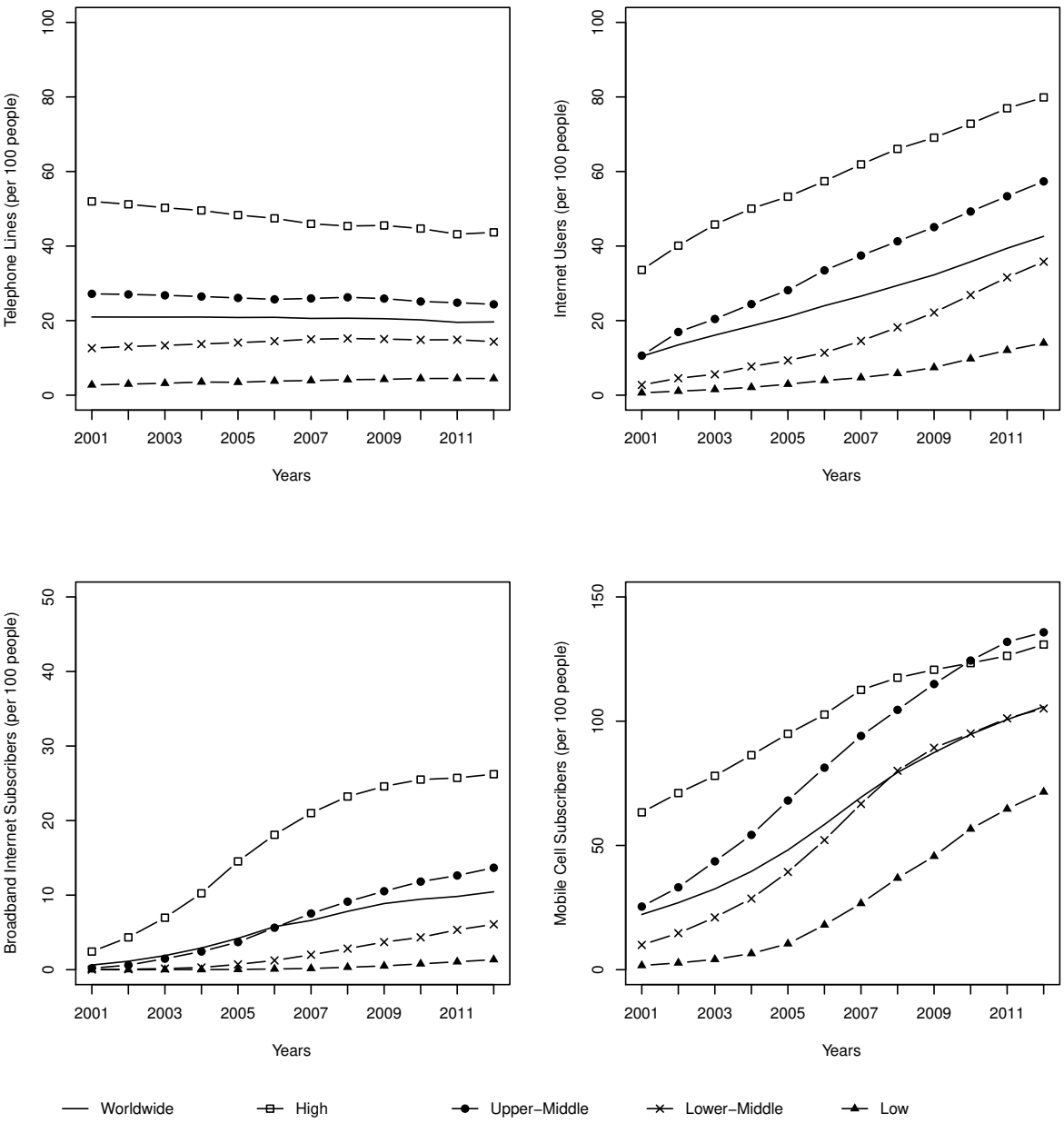
World Bank Atlas method to classify the countries with a population of more than 30,000 by income in four categories. For the (fiscal) year 2007, low-income economies are defined as those with a gross national income (GNI) per capita of $875 or less; lower middle-income economies are those with a GNI per capita between $876 and $3,465; upper middle-income economies are those with a GNI per capita between $3,466 and $10,725; high-income economies are those with a GNI per capita of $10,725 or more.[33]

The first chart (top left) of figure 3.1 shows the development of the number of telephone lines (per 100 people). In the investigation period, worldwide penetration of telephone lines is relatively constant at 20%. Countries with low incomes have the lowest rates. In this income group, the penetration rate over the entire period is 3.8% on average. Since 2001, however, the number of telephone lines (per 100 people) has grown by 62.4%. This is different in the case of high-income countries. These countries have the highest penetration rate of 47.3% on average. This is more than twice as high as the global average. Over the entire period, however, the number of telephone lines (per 100 people) decreases by 16% to an average of 43.7 telephone lines per 100 people. The countries with lower middle income are below the worldwide average with an average penetration rate of 14.2, whereas the number of telephone lines per 100 people has increased (+13.7%). With 25.9 telephone lines per 100 people, the countries with upper middle income are above the worldwide average and decreases above that over the period (-10.4%). Throughout the investigation period, the number of telephone lines remains constant across all country groups. This confirms the role of the telephone network as a basic ICT technology. However, the decreasing rate in the two upper income groups indicates that this technology has been partially replaced by other (e.g. mobile phone technologies).

The second chart (top right) of figure 3.1 shows the development of internet users (per 100 people). In contrast to telephone lines, the relative number increases over time. On average, the relative number of internet users per 100 people worldwide increases by more than four times from 10.4 to 42.6. Although the number of internet users is highest in high-income countries and above the global average, the growth over the investigation period is the lowest. The number of internet users in the lower middle and low-income groups is below the global average. It is remarkable that the growth in both groups is stronger in the first half of the investigation period than in the second half. In the group of low-income countries, growth decelerates from 531% (2001-2006) to 197% (2007-2012). Growth in lower middle-income countries is 317% (2001-2006), followed by a reduction to 146% (2007-2012).

[33] Both current and historical classification by income can be downloaded from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

Figure 3.1: Development of the ICT Infrastructure Components



Note: Shown are the curves of the averaged variable values. The curves illustrate the global average as well as that of the countries with high, upper-middle, lower-middle and low incomes. The countries are grouped using the World Bank's Atlas method following the classification of the (fiscal) year 2007.

Broadband connections are not as widespread as telephone lines (see figure below left in figure 3.1). The worldwide average is 5.8 connections per 100 people. Again, high-income countries have the strongest average penetration of this technology. Starting from an average of 2.4 broadband internet subscribers in 2001, the number increases to 26.2 in 2012. The resulting curve in figure 3.1 indicates a saturation. The group of countries with upper middle income also shows saturation tendencies. Until 2006, broadband penetration in this group is lower than the global average. By 2012, the average number of broadband internet subscribers in this group increases

to 13.7. The low and lower middle-income groups have an average of 1.3 and 6.1 broadband internet subscribers in 2012. The growth rate remains stable and linear in the second half of the investigation period.

We finally examine the number of mobile cell subscribers (per 100 people) in the bottom right of figure 3.1. In relation to the other indicators, these increase sharply over the investigation period. The global average number of mobile cell subscribers in 2001 is 22.2 and increases to 105.8 by 2011. It is remarkable that the scale of this indicator exceeds 100. This implies that there are multiple subscriptions per person in several countries. In 2012, the number of mobile cell subscribers in high-income countries is 130.9, while the number in upper middle-income countries is even higher, averaging 135.8 mobile cell subscribers. The number of mobile cells also increases in the group of countries with lower middle income and has been close to the global average since 2007. The growth of mobile cell subscribers is highest in the low-income countries. The growth of mobile cell subscribers is the highest in low-income countries. The proportion of mobile cells in this group increases by a factor of 42 and is thus higher than the lower middle-income countries (10), upper middle-income countries (5) and high-income countries (2).

Table 3.1 reveals that the variables are highly correlated with each other. This is not surprising, since users connect to the internet either via telephone line or via wireless and mobile technology. For this reason, the fraction of internet users overlaps with the fraction of telephone lines on the one hand and the mobile cell subscribers on the other hand. We can find a similar overlap between the fraction of internet users and mobile cell subscribers with the fraction of broadband internet subscribers. This overlap is of course also reflected by the high correlations between the single aspects of ICT infrastructure.

Table 3.1: Correlation Coefficients of ICT Variables

	(1)	(2)	(3)	(4)	(5)
(1) Telephone Lines	1.000				
(2) Internet Users	0.908	1.000			
(3) Broadband Internet Subscribers	0.897	0.925	1.000		
(4) Mobile Cell Subscribers	0.769	0.798	0.703	1.000	
(5) PCs	0.828	0.869	0.867	0.645	1.000

Note: All Variables are in values per 100 people and averaged over the years of 2002-2012. Pearson correlation coefficients are computed between each pair of variables using all complete pairs of observations on those variables. Based on 204 observations, the correlation coefficients of 178 complete pairs are computed.

Due to the strong correlation between the individual variables, they are well suited for performing a Principal Component Analysis (PCA). By using a PCA, the strongly correlated variables can be reduced to a single meaningful variable. Table 3.2 summarized the results of the PCA using averaged data for the decade 2002-2012.

As we can see in Table 3.2, the first principal component describes 93.2% of the entire variance. All variables load on the first principal component and all loadings have a positive sign. The other components add only a small amount to the explained variance. As a result of the PCA,

Table 3.2: Output of the Principal Component Analysis

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
Telephone Lines	0.284	-0.297	-0.517	0.738	-0.140
Internet Users	0.388	-0.414	-0.444	-0.665	-0.198
Broadband Internet Subscribers	0.095	-0.192	-0.133	-0.024	0.967
Mobile Cell Subscribers	0.824	0.541	0.159	0.022	0.049
PCs	0.285	-0.640	0.702	0.113	-0.056
Standard derivation	86.277	20.734	8.149	6.090	2.594
Proportion of Variance	0.932	0.054	0.008	0.005	0.001
Cumulative Proportion	0.932	0.986	0.994	0.999	1.000

Note: All Variables are in values per 100 people and averaged over the years of 2002-2012. Based on 178 observations, the Principal Component Analysis is conducted using the singular value decomposition, which examines the covariances/correlations between the individual variables.

the five variables of ICT infrastructure can be merged to a single variable that comprises most of the information.

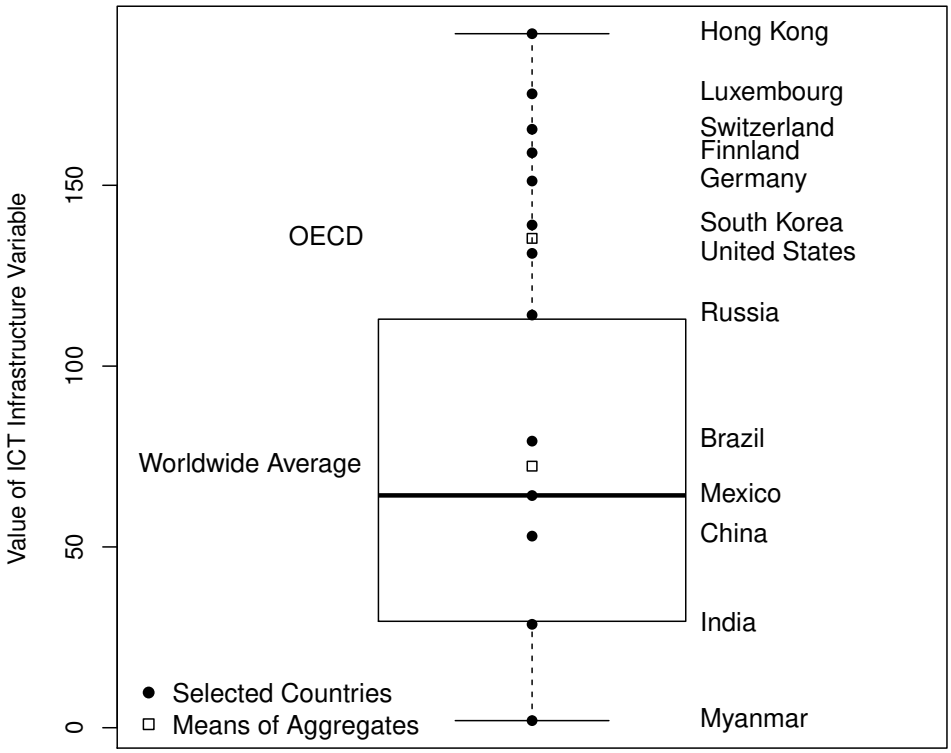
This procedure can also be applied for the calculation of ICT infrastructure in the individual years. In this case, only the former four variables are used, since data for the number of PCs is only available as average value over the period 2002-2009. Also in this variant, the variables load on the first principal component, whereby all loadings have a positive sign. We can observe these high variances on the first principal component in combination with constant signs for all available years 2001-2012. As a result of this annual variant, the four ICT variables can be merged to a single variable that comprises most of the information for a total of 148-167 countries (depending on the respective year). Figure A1 in the appendix illustrates the first component loadings of the PCA. Table A1 in the appendix shows the explaining proportion of the first component to the total variance.

3.3 Distribution and Development of ICT Infrastructure

We now proceed to the analysis of the variable for ICT infrastructure, which we constructed using a PCA in the previous section. In the context of this dissertation we will use this variable in two versions. Initially, we present the first version of the variable, which is composed of the averaged values of the components for the years 2002-2012. Subsequently, we describe the second version, in which the variable for ICT infrastructure is calculated for each year of our investigation period. This analysis provides information on the distribution and development of ICT infrastructure.

The boxplot in figure 3.2 illustrates the distribution of this ICT infrastructure variable, constructed for a total of 178 countries as the first principal component. The values of selected countries (right side of the boxplot) as well as the averaged values of all countries and OECD in particular (left side of the boxplot) are also indicated in the boxplot. The variable values range from approximately 2 (Myanmar) to a value slightly above 191 (Hong Kong). This implies huge cross-country differences in the stage of ICT infrastructure. The value for Mexico lies close to

Figure 3.2: Boxplot of the ICT Infrastructure Variable



Note: Distribution of the ICT infrastructure variable.

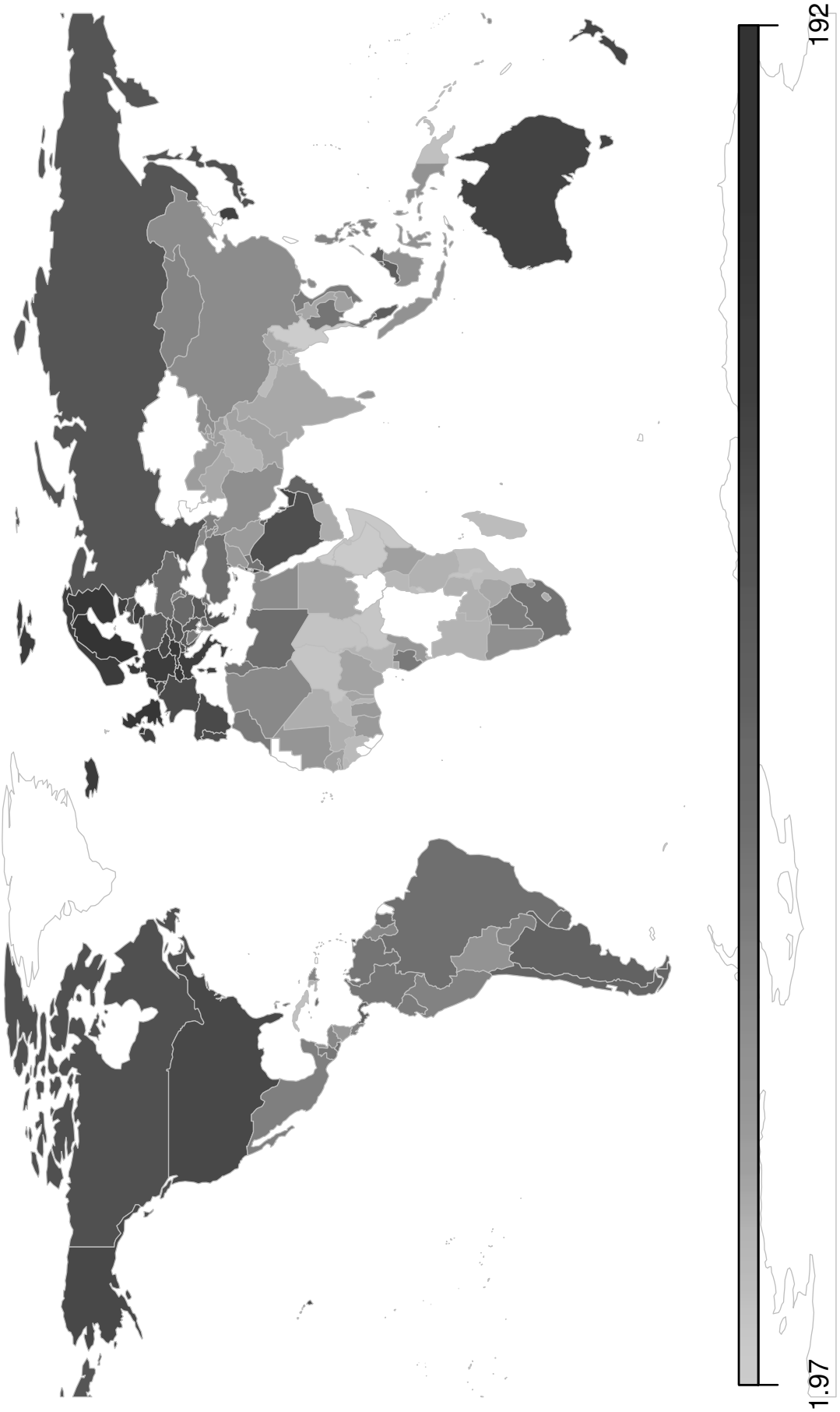
the median of 64.2, the value for Columbia and Tunisia near the mean of about 72. The box, with the 1st and 3rd quantiles as its lower and upper margins, has a comparably low position, which means that 75% of the countries have a value of ICT infrastructure below 112 while a few countries show fairly high values. These are either countries which are mere cities (Hong Kong, Luxembourg) or small advanced countries (Switzerland, Finland). While the worldwide mean value is in the middle of the box, the mean value of the OECD countries lies in the upper whisker between the positions of the U.S. and South Korea.

In figure 3.3 the global distribution of ICT infrastructure is plotted on a world map. Higher values of ICT infrastructure are represented by darker areas. These can be observed in North America (mean value of ICT infrastructure 127.90) and Europe (121.77). South, East (30.33) and West Africa (30.83) as well as Central Asia (34.37) present lower values.[34]

We now turn to the analysis of the second version in which the variable for ICT infrastructure is calculated for each year of our investigation period. Figure 3.4 illustrates the distribution and the development of the respective years in boxplots. It can be seen that there are substantial

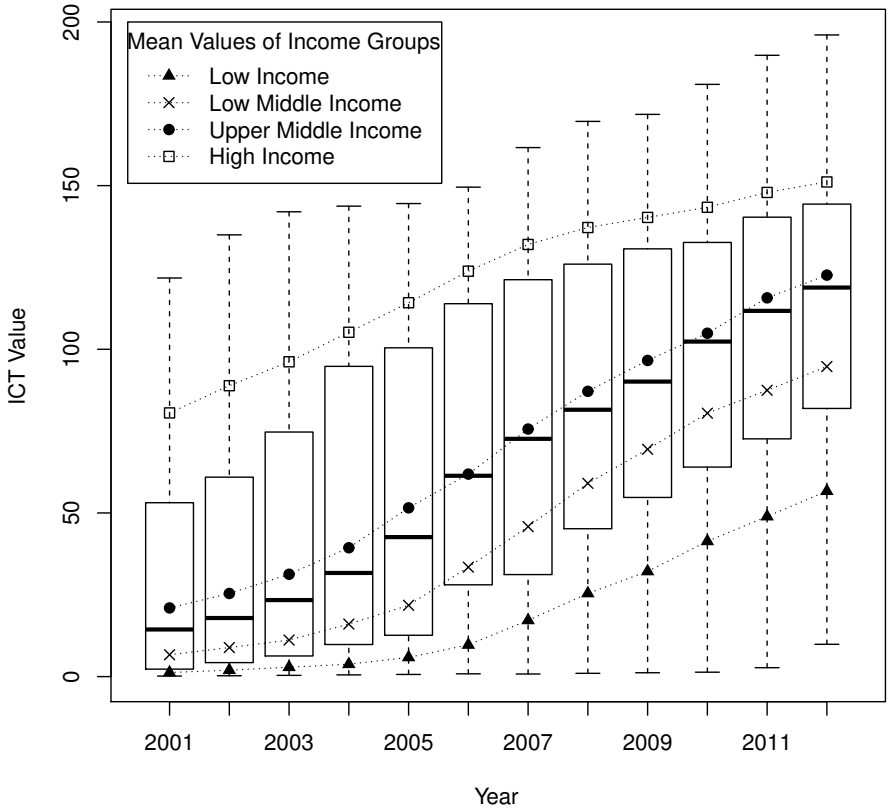
[34] An enlarged map illustrating the European distribution of the ICT infrastructure is given in figure A2 of the appendix.

Figure 3.3: Worldmap of the ICT Infrastructure Variable



Note: Global distribution of the ICT infrastructure variable. Countries with missing data represented by white color.

Figure 3.4: Boxplots of the ICT Variable in the Course of Time

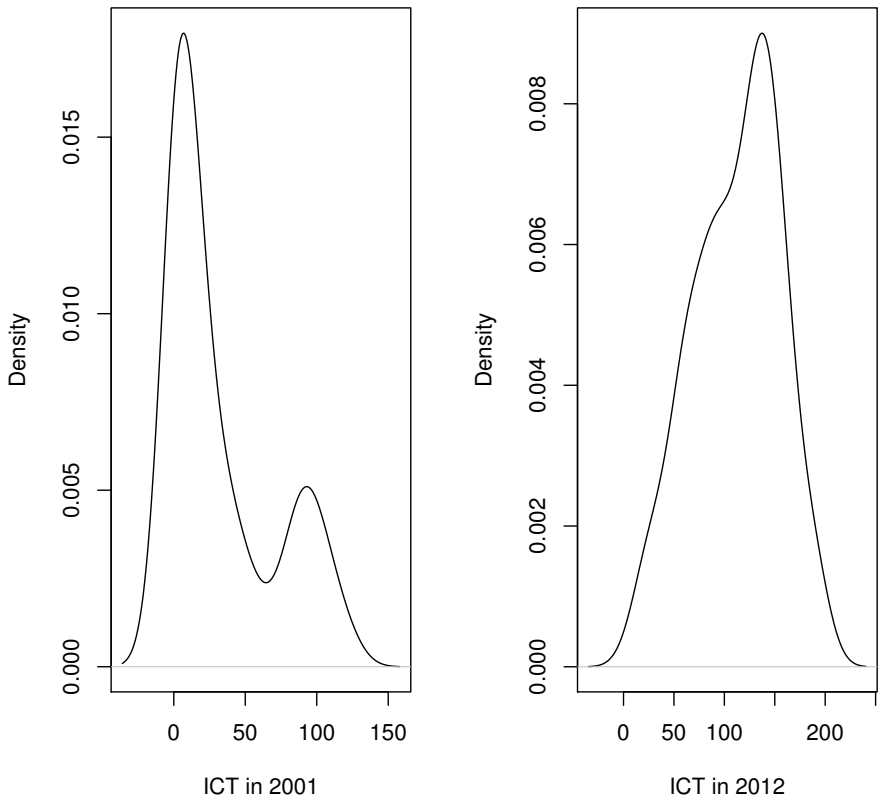


Note: Distribution of the ICT variable in the course of time.
Mean values of the income groups are shown as dashed lines.

cross-country differences in the stage of ICT infrastructure. Over the whole period 2001-2012 it is observable that there is on average an increase of the ICT infrastructure variable. Starting from the initial year 2001 the disparity between the ICT infrastructure values widenes steadily. While in 2001 the difference between the highest (121.78) and the lowest (0.08) ICT infrastructure value is 121.7, it increases to a value of 188.51 (difference between 198.38 and 9.87) in 2012. Countries with consistently high ICT infrastructure values are, inter alia, the countries in the OECD and Russia. The main reason for the increasing disparity over time are countries that were not able to improve their ICT.

It can be seen from the boxes (with the 1st and 3rd quantiles as its lower and upper margins) that the ICT values increase between 2001 and 2012 on worldwide average. The mean value increased from 29.77 in 2001 to 110.57 in 2012. The position of the boxes with respect to the whiskers reflects the changes in the distribution. It can be seen in 2001 that both the box as well as the median line have a comparably low position. Over the years, the position of box and median line rises so that the distributional skewness declines over time. A graphical comparison of the densities between 2001 and 2012 is given in figure 3.5. The figure illustrates that in 2001 the distribution is right-skewed. In 2012, the skewness considerably diminishes.

Figure 3.5: Density Plots of the ICT Variable in 2001 and 2012



By considering the courses of mean values by income groups in figure 3.4, the development stage of ICT worldwide becomes apparent. In general, countries with higher income levels also have higher values of ICT in contrast to countries with lower income levels. The mean values of the low-income group of countries lies on the lower whisker on every boxplot in the investigation period, the mean values of the high-income group of countries is on the upper whisker for every year. The mean ICT values of both, low middle as well as upper level-country groups, are positioned in the boxes. They are separated by the median of the respective years, with the mean value of the upper middle-income countries being closer to the median of the total distribution. In the course of time, the mean values of both country groups move away from those of the low-income group towards the group of countries with high income.

Especially remarkable are the curve shapes of the respective income groups, illustrating the saturation and catch-up process of the worldwide ICT distribution. Whereas there seems to be a saturation of ICT in the high income countries represented by lower growth rates of the ICT value over time, the low-income countries show a growing rate until 2006, followed by a linear growth until 2012. Both curves, for low middle as well as upper middle income, are s-shaped as diffusion rates usually are. Increasing growth rates of ICT values in the early years are followed by decreasing growth rates in the later years of the period.

3.4 Summary

In this chapter we addressed the definition and quantitative measurement of ICT. Although an initial definition of information technology dates back to the 1950s, it still seems to be valid today. Over time, the aspect of communication is also considered in relation to information technology. Today, the concepts of information technology and communication technology can be described as Siamese twins, because information seems inconceivable without communication and vice versa.

In the past, efforts have been made (such as those of the OECD) to establish a standardized and internationally accepted definition of IT/ICT sectors or products. The scientific literature, however, uses different delimitations of IT and ICT. Consequently, a concrete and standardized definition of ICT cannot be found. In context of this dissertation, we define ICT abstractly as a technology that provides access to information through communications.

Usually, IT and ICT are measured in monetary terms using capital data. These include expenditures in IT and communication equipment. The data come from both statistical offices as well as private data suppliers and were available only for the U.S. until the mid-1990s. Most studies at the country-level used private data suppliers. Data on expenditures in IT and communication equipment are only available for a limited number of countries (mostly OECD and EU). These originate from the national statistical offices and are not consistently defined due to different ICT definitions. In addition to the limited availability for a large number of countries, data on expenditures in IT and communication equipment must be quality-adjusted, which can lead to biased results in the case of inappropriate measurement.

As an alternative to monetary indicators, proxies can be used to measure the ICT level in countries. The literature has already used penetration levels of communication technology as proxy for ICT (see e.g. Becchetti and Adriani 2005). These penetration levels overcome some of the disadvantages of monetary measurement, as they do not have to be quality-adjusted, for example. These data are also available for a broad number of countries.

We can identify five indicators that are available to a variety of countries at different levels of development during the 2001-2012 period. These indicators are highly correlated and are therefore considered to be appropriate for the application of a PCA. The PCA reveals that the first principal component describes over 90% of the entire variance. As a result of the PCA, the variable of ICT infrastructure can be merged to a single variable that comprises most of the information.

It is evident that we can not be sure whether our variable mainly reflects the infrastructure, the equipment or the usage of ICT. It is also clear that infrastructure, equipment and usage mutually rely on each other. Therefore, we might have used these (and related) terms interchangeably, but decided to stick to the term ICT in the following discussion. The variable, constructed from a principal components analysis, can be interpreted as a proxy for either ICT equipment or ICT usage because of the close relation of both aspects. Given the findings of section 3.1 that communication technology is the necessary prerequisite for the exchange of information, the constructed variable is an appropriate proxy for measuring the ICT diffusion of economies.

We analyzed the variable in the first version, which is composed from the averaged values of the (single) components for the years 2002-2012, as well as in the second version, in which the variable for ICT infrastructure is calculated for each year of our investigation period. The analysis reveals substantial cross-country differences in the stage of ICT infrastructure. In general, countries with higher income levels also have higher values of ICT in contrast to countries with lower income levels. In the next chapter we will examine the determinants of ICT diffusion.

4 Determinants of ICT Infrastructure

4.1 Motivation

Since ICT is commonly suggested to be a determinant of macroeconomic growth it will be of increasing importance in the coming decades. In the previous chapter 3, we have constructed an ICT proxy variable from a PCA that has merged highly correlated penetration rates of ICT infrastructure to a single variable that comprises most of the information. This non-monetary proxy variable for ICT is available for up to 178 countries at different levels of development for the 2001-2012 period. Our analysis of this ICT variable has revealed substantial cross-country differences in the stage of ICT infrastructure. Due to the assumed productivity- and growth-promoting properties of ICT, it is therefore of special interest what determines ICT infrastructure and, thus, explains the differences in its diffusion.[35]

The question of what determines ICT infrastructure is also interesting from a policy point of view and has already been considered under the term of 'global digital divide' in literature. Studies that address this issue across countries are, inter alia, Hargittai (1999), Caselli and Coleman (2001), Kiiski and Pojola (2002) and Chinn and Fairlie (2007). The estimates reported in the received literature are based on different theoretical approaches and therefore lead to different sets of explanatory variables. Thus, depending on the theoretical stance certain variables are considered in some studies whereas others are entirely neglected. What is lacking is a comprehensive approach which considers a broad set of candidate explanatory variables simultaneously and uses modern methods for model selection to determine the optimal set of explanatory variables.

In this chapter we pursue such an approach based on variable selection methods originating from machine learning research. These methods, the so-called Lasso and its variants, are based on regularization instead of significance tests. Their application leads to parsimonious regression specifications using the most relevant explanatory variables and reaching a high degree of fit. Although no economic theory is involved in the variable selection procedure, theory is of course involved in assembling the set of candidate variables and in the interpretation of the results. We investigate economic and institutional determinants of ICT infrastructure for a broad cross section of more than 100 countries at very different stages of development.

The empirical approach followed in this chapter relies on two distinctive features which are novel to the literature. First, methods from the machine learning literature are used to select the relevant explanatory variables from a broad set of candidates. Second, in addition to common least squares regression, recent methods for robust regression estimation and semiparametric regression are used to validate the results against the influence of outliers in the data and to uncover nonlinear effects of the explanatory variables, respectively.

The analysis of this chapter proceeds by providing an extensive literature review in section 4.2, identifying a wide range of relevant explanatory variables. This is followed by the description of the database in section 4.3. The variable selection approach and the route taken for the empirical analysis are outlined in section 4.4. The results are presented and discussed in section 4.5. Section 4.6 summarizes this chapter.

[35] This chapter is based on Krüger and Rhiel (2016).

4.2 Previous Work

In this section we review the existing literature on the determinants of ICT infrastructure. This serves to demonstrate the state of research and to identify those determinants previously scrutinized. Basically, many empirical studies have been conducted both at the micro and the macro level to discover the determinants of ICT in general. The studies on the micro level examine the factors influencing the firm's investment behavior in ICT. At the macro level, the literature discusses cross-country differences in the adoption of information technology, mostly in context of the 'global digital divide' between advanced and less developed countries.[36] In this chapter we focus on the factors explaining the ICT infrastructure at the macro level.

Most of the studies share the same approach of first making theoretical assertions about factors influencing ICT, followed by identifying appropriate indicators for these factors. In a second step, the dependent ICT variable is regressed on the identified explanatory variables. The studies basically differ in the specification of the dependent variable and therefore in the concretization of the research object ICT. Common dependent variables used in the literature are ICT expenditure, number of internet users, adoption of internet by employees or ICT imports. Furthermore, the individual investigations differ regarding time coverage and country sample.

As a result, it is not surprising that quite diverse findings are reported in the literature. One the one hand, a set of common ICT-explaining variables is used in the studies. On the other hand, another group of variables is mentioned in the literature only in single occasions. Generally, the variables can be classified in the following categories, in decreasing order of importance in the literature: the economic wealth and structure of the countries, human capital, regulations, demographic factors and geographical/territorial factors. We now review the results of this literature structured along these categories.

Economic Wealth and Structure

Per capita income is the main and most widely used determinant of ICT (see e.g. Hargittai 1999, Kiiski and Pojola 2002, Norris 2001, Beilock and Dimitrova 2003). In the previous literature it was found that countries whose citizens are better off economically tend to have more ICT (see e.g. Hargittai 1999, Beilock and Dimitrova 2003). The underlying assumption is that countries with higher per capita income invest more in R&D and are therefore more able to discover and better in adopting ICT (Baliamoune-Lutz 2003). So, per capita income influences ICT indirectly. Next to education, income is an important determinant of computer ownership and internet use (OECD 2001). Per capita income is found to be positively and significantly related to ICT adoption by Caselli and Coleman (2001), Guillén and Suárez (2001), Kiiski and Pojola (2002), Baliamoune-Lutz (2003), Pohjola (2003), Chinn and Fairlie (2007), Wunnava and Leiter (2009). In contrast, Dasgupta et al. (2001) find the relationship to be non-significant.

Next to the level of per capita income, economic conditions are also characterized by income equality within a country, which "may have a negative effect on ICT diffusion because fewer people will be able to afford to pay for ICT products and services" (Wunnava and Leiter 2009,

[36] The term (global) digital divide is extensively discussed in Norris (2001) and Hargittai (2003).

p. 418). Hargiatti (1999) examined (among other variables) the impact of income equality on internet connectivity among OECD countries, but did not find a significant relation.

The sectoral composition of the economy has also been considered in the literature. As the underlying idea, the share of manufacturing and/or service sector are supposed to positively affect investment rates in ICT. Caselli and Coleman (2001) found no evidence supporting this assertion. Despite this, they found evidence for an inverse relationship to the share of the agricultural sector. A positive effect of employment in the service sector (as percent of total) and negative in public sector was found by Gust and Marquez (2004).

Human Capital

Next to the differences in the economic wealth of countries, human capital is frequently addressed in the literature. The basic idea for considering human capital as a determinant of ICT is that skilled and educated workers are more capable of learning how to use new technologies. Academic institutions especially play an essential role in adopting new technology (Guerrieri et al. 2011). While schools were among the first to introduce young people to ICT, these technologies provide the basis for research and education today and also promote their adoption in this way.

From the theoretical point of view, human capital seems to be one of the most essential factors positively influencing ICT adoption. Empirically, however, most of the authors found no clear evidence for this hypothesis. Wunnava and Leiter (2009) found significantly positive effects of tertiary enrollment on internet diffusion. Baliamoune-Lutz (2003) uses the education index from the UNDP Human Development Report as a variable for human capital, finding a positive effect on the diffusion of mobile telephones, but no effect on the diffusion on internet hosts, internet users or personal computers. Crenshaw and Robinson (2006) used tertiary education enrollments and Chinn and Fairlie (2007) chose the years of schooling as a determinant for human capital. Both also find mixed evidence for the role of human capital as a determinant of ICT.

Gust and Marquez (2004), using years of schooling as determinant for ICT expenditures, and Hargiatti (1999), relying on the education index from the UNDP Human Development Report, find a significant and positive effect on OECD countries. Kiiski and Pohjola (2002), however, came to a different result using the average years of schooling for the population over age 15, obtained from Barro and Lee (see Barro and Lee 2000), to discover the effect on internet diffusion in the OECD.

Thus, education does not seem to explain global differences in ICT robustly. In contrast to studies on the macro level, a positive relationship between ICT and employee qualification can be found on the micro level at Bayo-Mariones and Lera-Lopez (2007), as well as Haller and Traistaru-Siedschlag (2007).

Intuitively, a low level of education obstructs both the accessibility and distribution of ICT. A form of a particularly low education level is illiteracy. Literacy is required because of the text-based technologies of application software, world wide web and e-mail. However, the application range has been expanded in the last years. For example, video and voice communication applications do not necessarily need a higher level of literacy. The effect of literacy has been examined in studies by Baliamoune-Lutz (2003) as well as Chinn and Fairlie (2007). Baliamoune-Lutz

(2003) uses literacy rates of adults as an indicator of the initial level of education. She found no significant effect of literacy on ICT. Chinn and Fairlie (2007) neither find a significant effect of illiteracy rates on computer penetration rate nor internet penetration rate. In general, data on literacy[37] are limited. Behrman and Rosenzweig (1994) pointed out that a major problem in using literacy data for cross-country comparisons lies in differences of its definition. As another problematic issue, actual data on which literacy rates are based are often sparse and dated. Besides this critique on the definition and calculation, Barro and Lee (2013) found that literacy rates do not adequately measure the aggregate stock of human capital. Although frequently used, literacy rates do not seem to be an appropriate variable to capture human capital.

Besides general education and literacy, the knowledge of the English language is an important aspect of ICT usage. Because English is the most important language in the ICT domain, most of the software, internet sites and internet-supported communication is shaped in that language. In connection with higher education, most of the scientific and academic work is taught, written and published in English. Caselli and Coleman (2001) are not able to find a significant effect for the English language skills of the population on computer imports per worker. Kiiski and Pohjola (2002) measure English skills by the percentage of pupils in secondary education learning English from the European Commission. The lack of data reduces the number of observations to 17 countries, for which they significantly find a negative sign in the regression. Guillén and Suárez (2001) include a dummy variable to identify countries in which English is an official language or the most widely spoken language.[38] They find English to be positively related to the worldwide number of internet users and hosts.

In summary, it can be concluded that human capital is theoretically one of the most plausible factors for explaining ICT. However, the empirical evidence is rather mixed and can not robustly identify a relation in several studies using various indicators for human capital.

Regulation

The impact of regulation on ICT adoption is a widely discussed topic in the literature. The basic argument is that all kinds of regulations or constraints hinder individuals in acting optimally (Guerrieri et al. 2011). The regulation aspect is particularly relevant in interaction with the prosperity level of nation. The idea is that "richer countries have well-developed market economies and well-established legal systems, and as a result are able and willing to invest more in research and development and innovation" (Wunnava and Leiter 2009, p. 416).

Popular variables measuring the extent of regulation from the literature are indexes for property rights and civil liberties, used by Caselli and Coleman (2001), Norris (2001), Baliamoune-Lutz (2003) or Crenshaw and Robinson (2006). The results show mixed evidence. Caselli and Coleman (2001) find a positive influence of property rights on the computer imports per worker, but only for a specific set of 45 countries. Baliamoune-Lutz (2003) finds that property rights explain the diffusion of mobile telephones and internet hosts significantly. She finds neither an effect on the diffusion of internet users nor on the diffusion of personal computers. In her examination civil

[37] Both the adult literacy rate of the population over age 15 and illiterate population over age 15.

[38] They also include a dummy variable for Scandinavian countries, in which an unusually large percentage of the population knows English as a second language.

liberties only have a significant and positive relation on the diffusion of internet hosts. Crenshaw and Robinson (2006) find property rights explaining the diffusion of internet hosts significantly.

Dasgupta et al. (2001) included the aspect of competition policy in their analysis. They argue that measures of competition policy "affect both the supply of internet services and the intensity of their use by local firms" (Dasgupta et al. 2005, p. 3). This idea can be transferred from internet services to the entire ICT. As a proxy for government competition policy Dasgupta et al. (2001, 2005) use the variable 'Government Inhibition of Competition in the Private Sector' from World Bank database. This variable, varying from 1 (most inhibition of a competitive private sector) to 6 (least inhibition), indicates whether the country inhibits a "competitive private sector, either through direct regulation or by reserving significant economic activities for state-controlled entities" (Dasgupta et al. 2005, p. 3). The authors find evidence for their hypothesis that a low level of inhibition has a significant and positive effect on the diffusion of internet and mobile phone subscribers.

Another aspect examined in the literature is the market structure of the telecommunications sector. The basic idea here is that competition in the telecommunications market leads to reduced prices for access and use. The results are again ambiguous. Hargiatti (1999) found a negative influence of a telecommunications monopoly on the internet connectivity in industrialized countries. The evidence of such a negative influence could not be confirmed by Kiiski and Pohjola (2002) and only partially by Guillén and Suárez (2001, 2005).

Gust and Marquez (2004) establish a negative influence of regulation in the labour market on ICT spending. They use three indexes: an index of employment protection legislation (from the OECD), an index of regulatory burdens on startups (World Economic Forum) and an index on overall regulatory burdens (World Economic Forum). All three indexes had a negative and significant influence on the ICT expenditures of 13 industrialized countries during the period 1992-1999.

As can be seen, various aspects and variables exist concerning the subject of regulation. In conclusion, the subject of regulation is important for an explanation of the 'global digital divide'. However, a definite variable capturing the degree of regulation could not be identified thus far and is not within reach because of the multi-faceted nature and great diversity of regulatory measures.

Demographic Factors

As a further aspect, demographic factors have received attention in the literature. The hypothesis states that the age structure and the size of the urban population explain ICT. The underlying idea is that young people and the urban population in general tend to use more ICT because of network economies and firms being mostly located in cities or in their neighborhood. Concerning the age structure, no effort has been made to determine empirical evidence. Chinn and Fairlie (2007), however, suggest, that "the global digital divide would be even larger if developing countries had an age composition that was more similar to the United States" (Chinn and Fairlie 2007, p.18).

The share of cities in the production of national gross domestic product ranges from an average of 55% in the developing world to 85% in developed countries (Crenshaw and Robinson 2006). Therefore, it would be plausible that a higher degree of urbanization positively influences ICT diffusion. Both Dasgupta et al. (2001) and Crenshaw and Robinson (2006) find a positive effect of urban population. Chinn and Fairlie (2007), however, detect a negative effect.

Geographical / Regional Factors

To control for geographical and regional factors, several authors include respective dummy variables. The usage of these variables uncovers the influence of explanatory variables on ICT for a specific group of countries. Kiiski and Pohjola (2002) include dummy variables for "nordic" countries, "southern" countries as well as for Mexico and Turkey. Beilock and Dimitrova (2003) divide the world into six regions and test for differences in internet usage rates. The regions considered are highly developed nations, Latin America, formerly socialist nations, Middle East/North Africa, rest of Asia and Sub-Saharan Africa. Only the parameter estimate associated with the rest of Asia is found significant with a positive sign.

Interim conclusion from the literature review

In this section we reviewed the relevant literature on determinants of ICT infrastructure on the macro level. During the last 15 years, several attempts have been made to explain the 'global digital divide'. As to be expected, the results are quite diverse. Some variables, like the GDP per capita, were unambiguously identified as a major determinant of ICT. A variety of variables are mentioned in the literature only once. Surprisingly, some variables or groups of variables have shown no significant influence despite their clear theoretical relevance. Even though human capital is one of the most featured factors in theory, the empirical evidence could not be consistently established in several studies using various indicators of human capital.

Taken together, the bottom line is that the question of what the 'global digital divide' explains has not yet been conclusively answered. In the following, we will use the insights gained from the literature reviewed above to build up an encompassing database of candidate variables which are potentially relevant for explaining ICT infrastructure. For these candidate variables we undertake a specific variable selection approach to find the variables which are most relevant for explaining our indicator of ICT infrastructure. Next, we turn to the construction of this indicator and the description of the database in general.

4.3 Data

The data used for forming the dependent variable and constituing the set of candidate explanatory variables are assembled from various sources. We will first focus on the dependent variable for ICT infrastructure which we construct as the first principal component from a principal components analysis in section 3.2. The variables serving as candidates for the model selection procedure used in this chapter are described subsequently.

ICT Infrastructure Variable

We use the (log transformed) indicator for ICT infrastructure as explained and constructed previously in the sections 3.2 and 3.3, generated as an average of the respective variable values for the period 2002-2012. We denote this variable as $\ln IT$ in the following. We generate two further ICT infrastructure variables. Firstly, a variable generated as an average for the subperiod 2002-2006 ($\ln IT_1$). Secondly, a variable generated as an average for the subperiod 2008-2012 ($\ln IT_2$). This second choice for the dependent variable allows to investigate the degree of persistence by allowing $\ln IT_1$ as an explanatory variable.[39]

Explanatory Variables

As described in the literature review above, many variables appear to be used to describe the global differences in ICT or ICT infrastructure. In addition to the variables addressed in the literature, some other variables seem to be potentially relevant. Since the variable selection in the received literature sometimes seems to be arbitrary we pursue a different approach. This approach consists of compiling a large database of potentially relevant candidate variables and using a specific statistical approach (the so-called Lasso explained below) to select the relevant explanatory variables from this pool.

All collected data are freely available on the internet. They come from the World Bank[40], the Quality of Government Institute[41], the Barro-Lee dataset[42], the Heritage Foundation[43], the Penn World Table version 8[44], the International Monetary Fund[45], the database of the United Nations[46] and the dataset of Sala-i-Martin, Doppelhofer and Miller[47]. We averaged all variables over the years from 1980 to 2000 as far as possible and mostly transformed them by taking their natural logarithms. As far as values for the years 1980 to 2010 were available, we also calculated the average growth rate (in logged differences) and the standard deviation of the annual growth rates. In addition to these variables, we also include dummy variables to control for geographical localization[48] and the development stage[49] of a country. All variables, their data source and literature references are listed in table B2 of the appendix.

[39] The PCA is conducted using the singular value decomposition, which examines the covariances/correlations between the individual variables.
[40] http://databank.worldbank.org/data.
[41] http://qog.pol.gu.se/data/datadownloads.
[42] http://www.barrolee.com.
[43] http://www.heritage.org.
[44] http://www.rug.nl/research/ggdc/data/pwt.
[45] http://data.imf.org.
[46] http://data.un.org/DataMartInfo.aspx.
[47] The dataset was assembled for the paper of Sala-i-Martin, Doppelhofer and Miller (2004). It can be downloaded from https://www.aeaweb.org/aer/data/sept04_bace_data.zip.
[48] Dummy variables for: East Asia and the Pacific, Europe and Central Asia, Latin America and the Caribbean, South Asia, Sub-Saharan Africa. Data is provided by Barro-Lee and available on the internet at http://www.barrolee.com/.
[49] Dummy variables for: Advanced Economies, developing countries and developed countries, countries of the OCED and countries of the European Union. Data from the first three groups is provided by Barro-Lee and available on the internet at http://www.barrolee.com.

As a result we obtain a dataset of 72 different variables for a total of 178 countries.[50] Expressed in the categories of the variable classification (see section above), the dataset contains 21 variables for national economic wealth and structure, 12 variables representing human capital, 23 variables measuring the extend of regulations, 3 demographic and 9 geographical/regional variables.[51] These 72 variables are supplemented by their transformations[52] to reach a total of 148 candidate explanatory variables.

The method of variable selection we pursue in this work requires a dataset with complete observations. Possible approaches to meeting this requirement are quite drastic: The exclusion of all variables which have at least one missing entry would result in only 4 countries left in the dataset. Vice versa, removing all countries which have a missing entry for at least one variable leads to a "dataset" with no countries included at all. The manual sorting-out of those variables not covering a sufficient amount of countries or of those countries with an insufficient number of variables, requires considerable effort and is inevitably subjective. Thus, an automated algorithmic approach is desirable.

To reach a more objective decision, we compute the percentage of available variable values per country. We then exclude all countries below a certain threshold of available variables values. From the remaining countries we then reduce the resulting set of all variables with at least one missing entry. This procedure results in a complete dataset. Using a threshold of 0.835 we reach a dataset with 81 variables and 113 countries, containing all OECD countries but also many developing countries. This is the typical sample size also reached in many cross-country growth analyses. Countries and variables of the final dataset are described in table B2 and B3, some descriptive statistics of these variables in table B4 of the appendix. Also included in the appendix are the density plots of the three ICT infrastructure variables for the final dataset with 113 countries (see figure B1). The plots show that the two levels variables $\ln IT$ and $\ln IT_2$ have rather similar left-skewed densities. We used this dataset for the subsequent analysis utilizing a bundle of statistical methods which are described in the following section.

4.4 Method

In this work we intend to select explanatory variables from a large pool of candidate variables by methods originating from machine learning research (see e.g. Murphy (2012) for a recent comprehensive account of this field). The aim of the selection procedure is to find appropriate variables for explaining cross-country differences in ICT infrastructure. These variables are subsequently introduced into a regression analysis. For the estimation we use ordinary least squares with a heteroskedasticity-consistent covariance matrix as well as a robust regression estimator and a semiparametric generalized additive model (GAM) estimator.

[50] We reduced the initial dataset from 255 countries to 178 countries for which data for the dependent IT infrastructure variable are available. A list of the countries included in the initial as well as in the reduced dataset is shown in table B3 of the appendix.

[51] Four further variables do not match the variable classification.

[52] Where meaningful, the variables are logarithmized. In addition, we also calculated the growth rate and standard deviation of the annual growth rates (in logged differences) for a specific variable if this is reasonable.

As outlined above, several ICT influencing variables have been identified in the literature. The results are diverse and partly contradictory. In our database of candidate explanatory variables we record those variables identified in the literature as well as generally potentially relevant variables. Since there are nearly as many explanatory variables in the database as country observations, the above mentioned methods of variable selection are used to obtain a parsimonious model with stable and unbiased coefficient estimates. These methods are currently diffusing from the machine learning area into econometrics.[53]

The statistical approach pursued is based on a linear regression model stated for country i out of a cross section of n countries

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + u_i,\ i = 1, ..., n\,, \tag{11}$$

where y_i denotes the dependent variable (an indicator of ICT infrastructure in our case), $\boldsymbol{x}_i$ is the k-vector of explanatory variables (including a constant) and u_i is the usual error term.

The machine learning methods for variable selection rely on regularization, which amounts to adding a penalty term to the least squares target function. The motivation is that larger coefficient estimates tend to induce higher variability in the least squares fit. Whereas the OLS estimator is unbiased under the classical assumptions, regularization tolerates some bias in order to reduce the variance. The Lasso (least absolute shrinkage and selection operator) regression, proposed by Tibshirani (1996), performs a selection of variables by introducing a specific penalty term weighted by a factor $\lambda > 0$. This term penalizes the magnitude of the regression coefficients in the vector $\boldsymbol{\beta}$ and thereby leads to a complete removal of some variables from the set of candidate explanatory variables.

The Lasso estimator minimizes the target function

$$\sum_{i=1}^{n}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2 + \lambda \cdot \sum_{j=1}^{k}|\beta_j|\,, \tag{12}$$

where the usual least squares target function is augmented by the regularization term serving to penalize large magnitudes of the regression coefficients. The amount of regularization is controlled by the parameter λ which may be chosen by cross-validation methods or information criteria.

The specific form of the regularization term used here causes some coefficients to be forced exactly to zero and thus excludes the associated explanatory variables completely. Those variables increase the penalty term by their regression coefficients but are not able to reduce the residual sum of squares by a substantive amount. This has the beneficial side effect of also reducing multicollinearity. Multicollinearity is usually a problem in large cross-country data sets, because many variables reflect the general state of development of the countries and thus are highly

[53] For more on machine learning methods in an econometric context see inter alia Bajari et al. (2015a,b), Belloni et al. (2012), Doornik and Henry (2015), Kleinberg et al. (2015), Schneider and Wagner (2011) and Varian (2014).

correlated. The Lasso tends to select only those explanatory variables with mild multicollinearlity of each other (see Bajari et al. 2015a).

A refinement of the basic idea is the adaptive Lasso proposed by Zou (2006), augmenting the penalty term by weight factors, i.e.

$$\sum_{i=1}^{n}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2 + \lambda \cdot \sum_{j=1}^{k} w_j|\beta_j|. \tag{13}$$

In the modified formula w_j denotes the weight factor of the j-th regression coefficient. In this work, we rely on the standard error adjusted adaptive Lasso (SEA-Lasso) proposed by Qiang and Yang (2013). For the SEA-Lasso the weights are defined by $w_j = \hat{\sigma}_j/|\hat{\beta}_j|$, where $\hat{\beta}_j$ are the OLS coefficient estimates and $\hat{\sigma}_j$ the associated standard errors. With this weighting scheme the SEA-Lasso has the advantage of being scale-independent. Moreover, the adaptive variants have the so-called oracle property (see Zou 2006), as demonstrated by Qian and Yang (2013) for the SEA-Lasso. The oracle property means that asymptotically the adaptive Lasso consistently selects the right variables (those with $\beta_j \neq 0$) and leads to a $\sqrt{n}$-consistent asymptotically normal estimator.

Since we need OLS estimates for forming the weights w_j in the penalty term of the SEA-Lasso target function, we use the so-called Elastic Net before applying the SEA-Lasso for the final variable selection. This procedure is sensible here although we are faced with $n > k$. When k is not much smaller than n and we have considerable collinearity in the data, the OLS estimates would be very unstable and the standard errors tend to be overestimated. In this case the weights could be heavily biased. To deal with this problem, we perform a pre-selection of variables before applying the SEA-Lasso. For this pre-selection we use the Elastic Net (Zou and Hastie 2005), which combines the basic Lasso with traditional ridge regression (Hoerl and Kennard 1970). It can be implemented by minimizing the modified target function

$$\sum_{i=1}^{n}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2 + \lambda \cdot \left(\alpha \cdot \sum_{j=1}^{k} |\beta_j| + \frac{1-\alpha}{2} \cdot \sum_{j=1}^{k} \beta_j^2 \right), \tag{14}$$

where $\alpha \in [0, 1]$ denotes an additional parameter, controlling the relative importance of the two penalty terms. The parameter α determines whether the penalty term is more akin to the Lasso (in the case of $\alpha = 1$) or more that of a ridge regression ($\alpha = 0$). For $\alpha = \frac{1}{3}$ both penalties are equally weighted. With a pre-selection of variables by means of this procedure, we can reduce biases in the weights used for the SEA-Lasso.

Our statistical approach for variable selection and coefficient estimation can be summarized by the following three-stage procedure:

1. Application of the Elastic Net for the pre-selection of variables. The penalty weight λ is choosen by cross-validation (actually by using 10 randomly assigned folds, repeated 100 times and averaged).

2. Final selection from the pre-selected variables from the previous stage through the application of SEA-Lasso. The penalty weight λ is chosen here by the Bayesian information criterion.[54]

3. Re-estimation of the regression with the selected variables by OLS, the robust Koller-Stahel estimator and GAM regression.

The final re-estimation stage is motivated as follows. All variants of Lasso select a subset of variables and shrink all coefficients towards zero by penalizing their absolute values. As described, regularization lowers the variance with tolerance of some bias. To reduce this bias, we re-estimate the final specification by least squares and robust regression analysis. Actually, we use the OLS estimator with heteroskedasticity-robust standard errors with the correction of MacKinnon and White (1985). We also use the robust regression estimator provided by Koller and Stahel (2011), which combines the advantage of a high breakdown point[55] with high estimation efficiency.

To uncover nonlinear effects and for validating linear relations we use an additional semiparametric GAM estimator. The GAM is formally stated as

$$y_i = s_1(x_{i1}) + ... + s_h(x_{ih}) + u_i,\ i = 1, ..., n\,, \tag{15}$$

where h denotes the number of selected explanatory variables from the previous stages and the functions $s_j(\cdot)$ are represented by splines. We use Wood's penalized likelihood approach as described in Wood (2001, 2006) for the computation in combination with thin plate regression splines to avoid the choice of knot locations.

The variable selection methods may neglect explanatory variables associated with coefficients of small magnitude which may simply be a consequence of the scaling of the variables (see Chernozhukov et al. 2015, p. 487). To counteract this tendency, we standardize the explanatory variables for use in the first two stages. In the subsequent estimation of stage 3, we us the original (not standardized) variables.

As an alternative mode of analysis we apply a bootstrap version of Lasso, the so called bolasso (see Bach 2008), instead of the first two stages. This variant runs the Lasso for several bootstrap replications of a given sample, on the basis of a residuals bootstrap. This method has proved to be a consistent model selection method under a wider range of conditions than the basic Lasso. We use a soft variant, keeping all variables that are selected in 90 percent of the bootstrap replications. Quite naturally, we find fewer variables using this method. These variables, however, can be viewed as the core predictors that are found to be robustly correlated with the dependent variable in at least 90 percent of the bootstrap replications. In addition, the bolasso can also be used as a device to combat the uniqueness problem of the Lasso in the presence of discrete regressor variables (see Tibshirani 2013).[56]

[54] For a detailed description see Qian and Yang (2013, pp. 298).

[55] The breakdown point is defined as the smallest fraction of contaminated observations in the sample that can lead to an arbitrarily large deviation of the estimator.

[56] All computations are programmed in R using the following packages: `glmnet` and `seaLasso` (for the variable selection), `car`, `lmtest` and `sandwich` (for the least squares regression with the computation of variance inflation factors and the heteroskadasticity-robust standard errors), `robust` (for the robust regressions) and `mgvc` (for the estimation of the GAM). The bootstrap Lasso is implemented in the package `mht`.

4.5 Results

We now turn to the presentation of the results from the variable selection procedure. This section is divided into three subsections, discussing the results for explaining the three variables introduced above in turn. The guiding idea is to use explanatory variables from a period before the period for which the ICT variable is constructed in order to reduce simultaneity bias.

4.5.1 Explaining ICT Infrastructure During 2002-2012

The regression results obtained with the three-stage procedure for the variable selection with $\ln IT$ (country means during 2002-2012) as the dependent variable are shown in table 4.1. The table contains the results of the OLS estimation with the heteroskedasticity-robust standard errors and those of the robust KS regression (reported in parentheses below the regression coefficients are the p-values of the standard t-tests). In addition, we present the results for the semiparametric GAM regression with the nonparametrically entered variables indicated by $s(\cdot)$ (for these variables we report the equivalent degrees of freedom (edf) jointly with the p-values of the F-tests for the joint significance of the spline terms in parentheses). The regressions rely on a total of $n = 113$ observations. In the case of the dummy variables such as region dummies (e.g. EU) we again report the regression coefficients with the p-values of the associated t-tests in parentheses.

Considering first the linear regression results in the first three columns of the table we find not all explanatory variables significant here. This is not a contradiction since the variable selection approach relies on regularization and not on significance testing. We find statistically significant coefficient estimates (at 5 percent level of significance) for the explanatory variables access to electricity (Elec_m_log), the European Union dummy (EU, OLS estimator only), gross fixed capital formation (gfcf_m_log), investment freedom (inv_freedom), the South Asian dummy (South.Asia), the Sub Saharan Africa dummy (Sub.Saharan.Africa, KS estimator only), urban population (UrbanPop_m_log, KS estimator only) and expenditure-side real GDP at chained PPPs (in millions 2005 US$) per person (RGDPP_m_log, KS estimator only). Expressed in the categories of the variable classification, the variable selection contains three variables explaining national economic wealth and structure, one variable measuring the extend of regulations, one demographic and two geographical/regional variables.[57] The 'm' in the abbreviation indicates the respective variable as averaged over the years of 1980 to 2000, 'sd' denotes the standard deviation of the respective variable.

[57] Access to electricity is not counted as it does not fit to one of the categories.

Table 4.1: Regression Results for the Three-Stage Procedure (dependent variable is $\ln IT$)

	OLS	KS		GAM
c	0.183	-0.043	c	4.260
	(0.757)	(0.914)		(0.000)
Elec_m_log	0.250	0.186	s(Elec_m_log)	2.789
	(0.001)	(0.000)		(0.000)
EU	0.101	0.098	EU	0.079
	(0.048)	(0.054)		(0.127)
Europe.and.Central.Asia	0.084	0.077	Europe.and.Central.Asia	0.060
	(0.206)	(0.166)		(0.300)
free_corrupt_m	0.002	0.001	s(free_corrupt_m)	1.000
	(0.376)	(0.615)		(0.400)
free_corrupt_m_log	0.107	0.131	s(free_corrupt_m_log)	1.000
	(0.240)	(0.098)		(0.339)
gfcf_m_log	0.243	0.199	s(gfcf_m_log)	1.000
	(0.001)	(0.002)		(0.000)
inv_freedom_m	0.003	0.003	s(inv_freedom_m)	1.530
	(0.030)	(0.008)		(0.027)
POP_sd	-6.037	-5.202	s(POP_sd)	2.046
	(0.214)	(0.087)		(0.099)
pyr_m_log	0.059	0.038	s(pyr_m_log)	1.000
	(0.398)	(0.490)		(0.225)
RGDPP_m_log	0.124	0.202	s(RGDPP_m_log)	2.962
	(0.230)	(0.002)		(0.024)
RGDPW_m	0.000	0.000	s(RGDPW_m)	1.000
	(0.398)	(0.732)		(0.559)
South.Asia	-0.411	-0.384	South.Asia	-0.439
	(0.000)	(0.000)		(0.000)
Sub.Saharan.Africa	-0.141	-0.171	Sub.Saharan.Africa	-0.205
	(0.090)	(0.008)		(0.006)
UrbanPop_m_log	0.121	0.117	s(UrbanPop_m_log)	1.000
	(0.166)	(0.019)		(0.024)
R^2	0.948	0.951	R^2	0.955
n	113	113	n	113

Note: Reported are the regression coefficients or the equivalent degrees of freedom (edf) in the case of the spline variables indicated by $s(\cdot)$. Stated in parentheses are p-values of the t-statistics or the F-statistics for the significance of the respective splines. In the case of OLS regressions the adjusted R^2 is reported. Renaud and Victoria-Feser (2010) explain the R^2 measure used in the case of the KS regressions.

In alphabetical order, the variable of access to electricity (Elec_m_log) is the first variable significantly explaining the ICT infrastructure during 2002-2012. This result is not surprising as ICT goods and services need power supply for their operation. The presence of electricity can be seen as an essential prerequisite for ICT infrastructure. The coefficient estimate is an elasticity and its value indicates that ICT infrastructure is inelastic with respect to access to electricity. Surprisingly, we could not find a consideration of this fundamental variable in the literature. Related is the study of Chinn and Fairlie (2007, 2010), using a variable to capture the electric power consumption (kWh per capita) for analyzing cross-country differences in computer and

internet penetration. In their study, they find no relationship between per capita electricity use and ICT penetration.

The EU dummy indicates that countries of the European Union on average have higher values of ICT infrastructure compared to the whole country set. In contrast, South Asian and Sub Saharan countries have values below-average, which is indicated by the negative sign of the respective coefficient estimates. For this fact we can find ample evidence in the literature. Individuals in high-income countries may have a higher ability to pay for personal computers or broadband services (Czernich et al. 2011) and tend to have higher degrees of internet penetration (Hargittai 1999). The geographical dummy variables approximately match with a high income (Europe) or low (South Asian and Sub Saharan) income levels.

Next, the gross fixed capital formation (gfcf_m_log) also belongs to the group of highly significant variables in both the OLS and KS regression results. The interpretation of the estimation coefficient can be ambiguous. On the one hand, investments in ICT infrastructure are part of the gross fixed capital. Hence, the amount of gross fixed capital formation is increased through higher investments in ICT. On the other hand, investments in certain goods or services increase investments in IT simultaneously. This is in particular the case with goods/services which need ICT infrastructure as a complementary product. These goods/services can be found in smart devices, household electronics, digital media, the automobile industry as well as in industrial products of the mechanical engineering sector or logistics (OECD 2011). Public investment such as the establishment and development of tolling systems or e-government services also requires ICT infrastructure as a crucial basis. Despite these obvious relationships, the role of gross fixed capital formation in relation to ICT infrastructure has not been examined widely in the literature.

Also significant in both the OLS and KS regression results is investment freedom (inv_freedom_m). This variable (provided by the Heritage Foundation) is represented by an index that indicates whether a country allows individuals and firms to move capital across countries' borders without restriction as well as capital flows internally (score of 100) or with restrictions on investment (score below 100).[58] Countries with a higher score of investment freedom are suggested to attract investors and therefore more (both domestic and foreign direct) investment.[59] As previously mentioned, part of these investments concerns products using IT/ICT infrastructure as complementary products.

Only significant in the KS regression is the GDP per person (RGDPP_m_log). As already mentioned in the section above, per capita income was found as the major and mostly identified determinant of ICT in the literature. The fact that the Lasso also selects per capita income to explain global differences in the diffusion of ICT is assuring for this result.

Also only significant in the KS regression is the urban population, measured as percentage of total population (UrbanPop_m_log). This result supports the hypothesis that the urban population

[58] Possible restrictions might be rules for foreign and domestic investment, payments, transfers, and capital transactions, restricted access to foreign exchange, labor regulations, corruption, red tape, weak infrastructure, and political and security conditions. For more information on the calculation see the Website of the Heritage Foundation: http://www.heritage.org/index/investment-freedom.

[59] See, for instance, Azman-Saini et al. (2010) for an overview of the discussion, a brief review of the literature and an empirical investigation of the international evidence.

tends to adopt more ICT, not least because of possible network economies. The positive effect of urban population is in accordance with the work of Crenshaw and Robinson (2006) as well as Dasgupta et al. (2001).

Also selected but not significant are the dummy variables of Europe and Central Asia, the score of freedom from corruption (free_corrupt_m(_log)), the standard deviation of the population (POP_sd), the variable of average years of primary schooling attained (pyr_m_log) and the output-side real GDP at chained PPPs per worker (RGDPW_m). Although these variables are not significant in the regressions, the coefficient signs are plausible in this context.

It is remarkable that the Lasso did not select even one human capital variable to explain ICT infrastructure in these regressions. Although we included several variables in the database, none was regarded as a major explanatory variable. This reflects the findings in the literature, where human capital is regarded as one of the most relevant ICT explaining factors in theory, whereas empirical evidence could not be established in several studies using various variables.

From the total of 14 selected variables, eight were found to be significant. From these, three variables describing geographical factors, two describe the economic status and structure and one variable each is included in the categories of demographic factors and regulation. Thus, we find the main areas of relevant influence factors also well represented here. With these variables we can explain about 95 percent of the variation (measured by the adjusted R^2)[60] in the log IT variable, averaged over the period of 2002-2012.

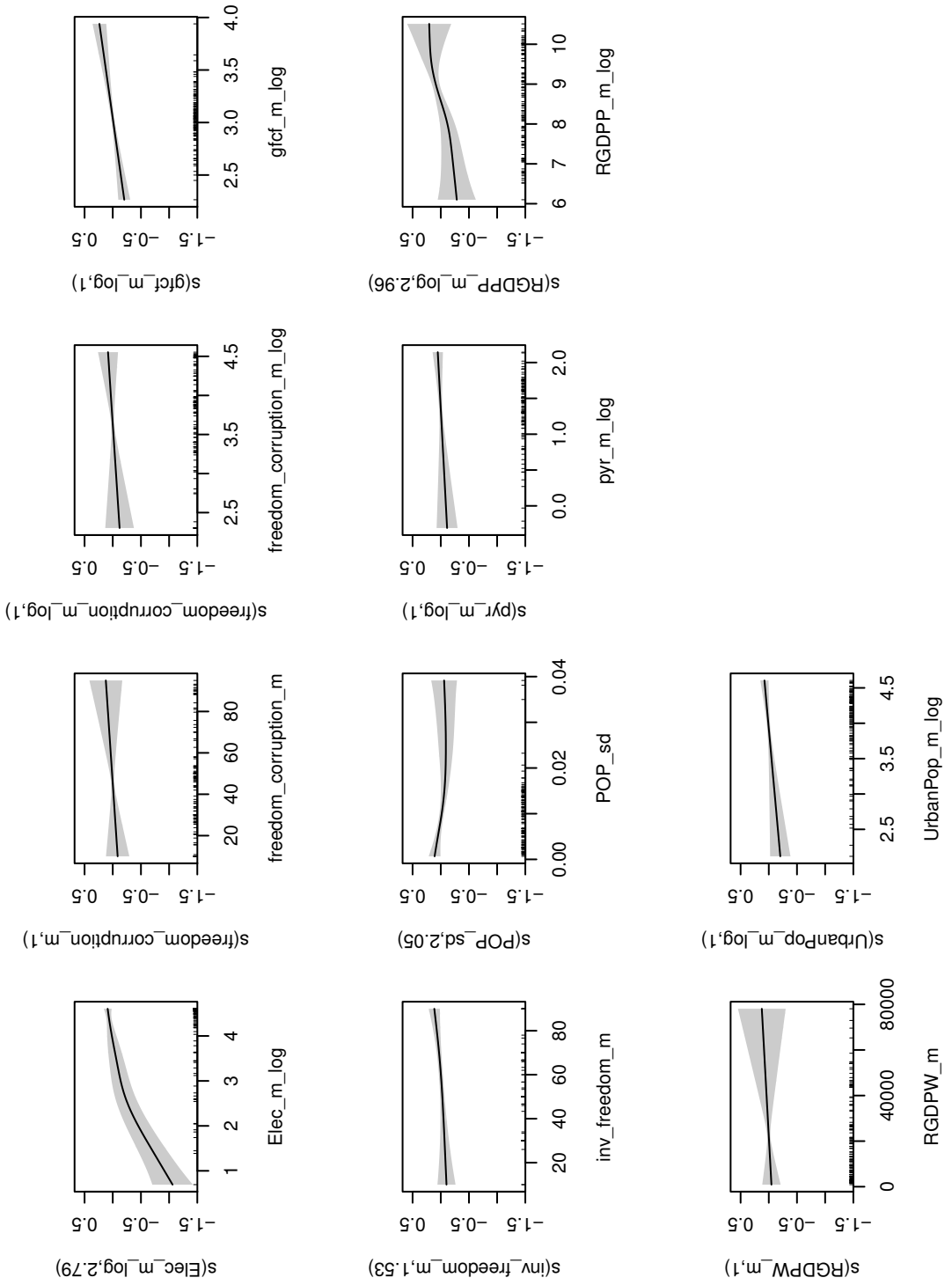
It may be suspected that multicollinearity is a major problem with such a large number of explanatory variables. This is, however, not the case since we find a condition number of about 15 based on the standardized matrix of explanatory variables and there are only very few variance inflation factors which may be viewed as large. This again shows the ability of Lasso-type procedures to successfully avoid multicollinearity.

Associated with the GAM regression results is figure 4.1, showing the plots of the (centered) spline terms for the selected variables. In the panels of the plot the tick marks at the abscissa (so-called rugs) indicate the positions of the data points of the respective explanatory variable. The gray shaded areas indicate the 95 percent confidence intervals. The equivalent degrees of freedom (edf) values substantially larger than one reveal nonlinear effects of Elec_m_log (access to electricity as percent of population), POP_sd (the standard deviation of population) and RGDPP_m_log (log GDP per person). The other variables appear to have a linear association with $\ln IT$. This assertion can be quickly verified by simply trying to draw a straight line through the gray-shaded 95 percent confidence intervals which is actually possible for the variables deemed linear.

At first in figure 4.1, the curve of variable Elec_m_log is concavely curved. The nonlinear effect shows that countries with a better electricity supply tend to have a more developed ICT infrastructure but this effect is driven by the large heterogeneity of the electricity supply variable across countries (see the rugs at the bottom of the right-hand panel of the figure). The association is weaker (the curve flatter) for the more advanced countries with a better electricity support system clustered at the upper end of the scale with values above four (approximately corresponds

[60] In the case of the KS regression, Renaud and Victoria-Feser (2010) explain the kind of R^2 measures used for the assessment of fit.

Figure 4.1: GAM Results for the Three-Stage Procedure (dependent variable is $\ln IT$)



to a 54% access of population to electricity) for this variable. A nonlinear effect of Elec_m_log is not surprising in this case. In general, urban areas were the first to be electrified because many customers hosted in a relatively small area.[61] The share of urban areas to national GDP ranges from an average of 55% in the developing world to 85% in developed countries (Crenshaw and Robinson 2006). In addition, urban populations tend to adopt more ICT (internet and computer) because of network economies. For these reasons it can be assumed that an initial electrification of urban (and mostly more industrialized areas) has a greater impact on the diffusion of ICT infrastructure than an electrification of rural (mostly not industrialized) regions. On closer inspection, the curve of variable Elec_m_log may be decomposed in two straight lines. At the value of approximately 2.5 the curve describes a kink which corresponds to 12% of population having access to electricity. Below this threshold, an increase of electrification has a stronger impact on the level of ICT infrastructure than above.

In the plot of the variable POP_sd, the solid line describes a mildly regressive curve. The rugs show that most of the data points have a value below 0.018, corresponding to a standard deviation about 1. In both theory and literature no connection between ICT infrastructure and the standard deviation of population is discussed. Moreover, since the edf value is only slightly larger than 2, we will not further elaborate on this issue.

At last, the edf value of RGDPP_m_log indicates a nonlinear effect. In the plot, the solid curve is s-shaped. As pointed out above, per capita income is an important determinant of computer ownership and internet use (OECD 2001). Hargittai (1999) as well as Beilock and Dimitrova (2003) argue that countries whose citizens are better off economically tend to have more ICT. Based on the assumption that countries with higher per capita income invest more in R&D and are therefore better able to discover and use ICT (Baliamoune-Lutz 2003), per capita income influences the ICT indirect. The curve of variable RGDPP_m_log shows a progressive course up to a level of approximately 9. This value corresponds to an expenditure-side real GDP per person of about 8100 US$. Up to this value, an increase of the GDP per person leads to larger effect on ICT infrastructure. Beyond this level saturation seems to take force.

Since we have dummy variables and other discretely-coded variables in our set of explanatory variables the uniqueness problem raised by Tibshirani (2013) may be an issue. We combat this problem by going a step further and employing the bootstrap Lasso procedure as described above to peel out those explanatory variables which are selected in 90 percent out of 10000 bootstrap replications of the Lasso. This device also delivers us the more robust explanatory variables. As to be expected, we obtain a substantially reduced set of selected variables. The final regression results are reported in table 4.2 and figure 4.2.

A first view of the results shows that the explanatory power of these regressions is somewhat reduced but remains well above 0.9. All dummy variables are now discarded by the model selection procedure. The remaining selected variables are all highly significant with one exception (UrbanPop_m_log in the case of the OLS regression). The finding that GDP per person (RGDPP_m_log) belongs to the group of robust explanatory variables again supports previous results finding per capita income to be a major determinant of ICT.

[61] Due to larger distances between customers in few inhabited, rural areas the further electrification causes marginal returns to diminish and thus drives average returns down.

In contrast to the regression of the three-stage procedure in table 4.1 the index of freedom from corruption (free_corrupt_m_log) belongs to the group of significant and even robust explanatory variables explaining global differences in the diffusion of ICT infrastructure. The basic idea of the score is that "corruption erodes economic freedom by introducing insecurity and uncertainty into economic relationships".[62] The index is provided by the Heritage Foundation and is mainly derived from Transparency International's Corruption Perceptions Index (CPI).[63] Multiplying the CPI by 10, the score of freedom from corruption ranges from 0 (very corrupt government) to 100 (very little corruption). The Heritage Foundation uses qualitative information from internationally recognized and reliable sources to determine the freedom from corruption score for countries that are not covered in the CPI.[64] The basic idea of the link between the score of freedom from corruption and ICT infrastructure is actually the same as for investment freedom. The fact remains that corruption, in conjunction with the consequences of insecure and uncertain economic relationships, discourages and sometimes prevents investment. Hence, complimentary investments in ICT infrastructure are also not undertaken.

The GAM regression results show that the effects of free_corrupt_m_log and RGDPP_m_log are clearly linear. As before, nonlinear effects can be uncovered for the variables Elec_m_log and UrbanPop_m_log. For Elec_m_log the associated figure 4.2 shows a similar curve shape for the bolasso procedure as for the three-stage procedure. However, at a value of about 4 the curve again becomes steeper. In this range of variable values above 4, we have a strong accumulation of rugs. The interpretation of this finding proves difficult, however, because a degressive curve shape (as in figure 4.1) is more plausible from a theoretical point of view. The reason why the increase in access to electricity above a level of approximately 55% should lead to a larger effect on ICT infrastructure than a level below is not clear and speculative.

The curve of UrbanPop_m_log is shaped like a wave. Different parts of the curve (intervals of 2.6-3.2 and 3.5-4) show that an increase in urban population in these intervals has a greater impact on ICT infrastructure than in the other intervals. It is remarkable that the curve weakly decreases from a value of 4.0 onwards. A further increase in urban population at a level of approximately 55% has a slightly diminishing effect on the level of ICT infrastructure. In this range the number of observations heaps up and the gray-shaded 95 percent confidence interval narrows. The diminishing effect of urban population on the level of ICT infrastructure can be explained by congestion effects.

[62] Source of the cite: http://www.heritage.org/index/freedom-from-corruption.

[63] The index in turn is composed by several data from various sources. The methodology of the CPI is described by Lambsdorff (2005).

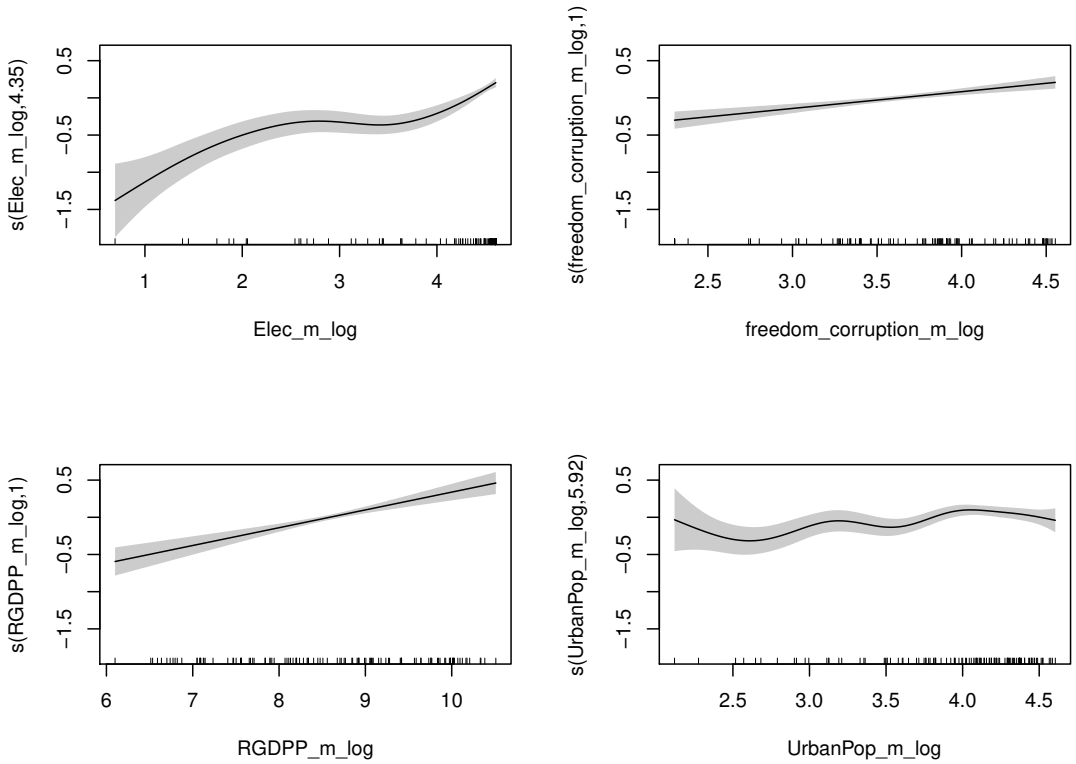
[64] For this purpose they use the following sources in order of priority: Transparency International, Corruption Perceptions Index, U.S. Department of Commerce, Country Commercial Guide, Economist Intelligence Unit, Country Commerce, Office of the U.S. Trade Representative, National Trade Estimate Report on Foreign Trade Barriers; and official government publications from each country.

Table 4.2: Regression Results for the bolasso Procedure (dependent variable is $\ln IT$)

	OLS	KS		GAM
c	-0.637	-0.691	c	4.218
	(0.005)	(0.000)		(0.000)
Elec_m_log	0.294	0.275	s(Elec_m_log)	4.352
	(0.000)	(0.000)		(0.000)
free_corrupt_m_log	0.195	0.174	s(free_corrupt_m_log)	1.000
	(0.001)	(0.000)		(0.000)
RGDPP_m_log	0.267	0.293	s(RGDPP_m_log)	1.000
	(0.000)	(0.000)		(0.000)
UrbanPop_m_log	0.165	0.161	s(UrbanPop_m_log)	5.922
	(0.117)	(0.009)		(0.002)
R^2	0.916	0.918	R^2	0.939
n	113	113	n	113

Note: Reported are the regression coefficients or the equivalent degrees of freedom (edf) in the case of the spline variables indicated by $s(\cdot)$. Stated in parentheses are p-values of the t-statistics or the F-statistics for the significance of the respective splines. In the case of OLS regressions the adjusted R^2 is reported. Renaud and Victoria-Feser (2010) explain the R^2 measure used in the case of the KS regressions.

Figure 4.2: GAM Results for the bolasso Procedure (dependent variable is $\ln IT$)



In this subsection we found several variables to explain ICT infrastructure during 2002-2012. In the initial three-stage procedure, a total of 14 variables have been selected explaining about 95 percent of the variation in the log IT variable. Among these variables 8 were found to

be significant with three variables for geographical factors, two describing economic status and structure and one variable are included in each of the categories of demographic factors, regulation and miscellaneous. In the associated GAM regression we could reveal nonlinear effects of Elec_-m_log, POP_sd and RGDPP_m_log. In the subsequent bootstrap Lasso (bolasso) procedure we get a reduced set of more robust explanatory variables. This procedure selects four variables with an explanatory power of well above 0.9 in all regressions. Furthermore assuring is the close correspondence of the OLS and the robust KS regression estimates. The following GAM regression detects nonlinear effects of Elec_m_log and UrbanPop_m_log.

Our results show the power of a wide variety of variables for explaining cross-country differences in ICT infrastructure. We are able to reestablish the empirical evidence for per capita income as significant and robust variable describing global differences in ICT infrastructure. Although we examined many empirical studies, we could not find evidence for Elec_m_log in the literature. This issue is quite remarkable as electricity can be viewed as a fundamental infrastructural prerequisite for ICT. This is even surprising as we found Elec_m_log to be a very robust and significant explanatory variable across all estimates.

4.5.2 Explaining ICT Infrastructure During 2002-2012

Turning to the results with $\ln IT_2$ as the dependent variable we find that only two explanatory variables are selected by the three-stage procedure. Recall that $\ln IT_2$ is the log average over the period 2008-2012. This allows us to put $\ln IT_1$ (the log average over the previous period 2002-2006) into the set of candidate explanatory variables. As shown in table 4.3 we see that the ICT infrastructure variable is characterized by persistence since $\ln IT_1$ appears as a strongly significant explanatory variable associated with a positive coefficient estimate. The coefficient estimate of about 0.34 (smaller than one) is indicative for the presence of conditional convergence of the ICT infrastructure across countries.[65] It seems that this persistence captures almost the entire amount of explanatory power of the other variables which were selected in the previous subsection. An exception is the variable Elec_m_log which remains strongly significant with a positive coefficient estimate although of a reduced magnitude. The coefficient estimates with the OLS and the robust KS estimates are rather similar, the overall explanatory power also remains substantial.

[65] Subtracting $\ln IT_1$ from both sides results in the change of $\ln IT$ on the left-hand side and a negative coefficient $0.34 - 1$ on the right-hand side. This is the indication of conditional convergence investigated in cross-section growth empirics, see e.g. Barro and Sala-i-Martin (1991, 1992).

Table 4.3: Regression Results for the Three-Stage Procedure (dependent variable is $\ln IT_2$)

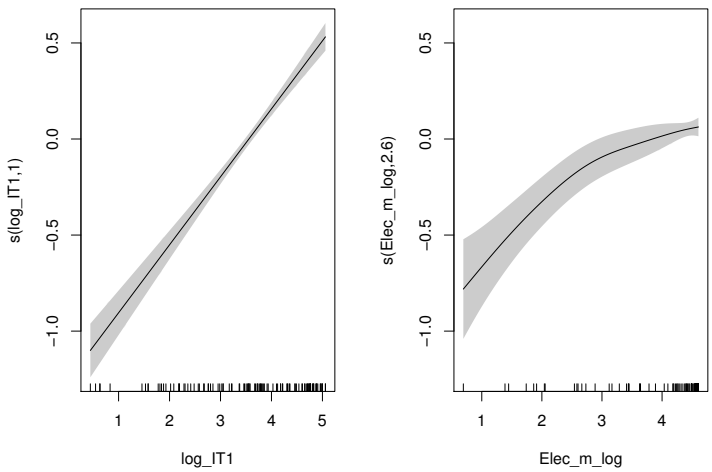
	OLS	KS		GAM
c	2.641	2.695	c	4.557
	(0.000)	(0.000)		(0.000)
log_IT1	0.337	0.344	s(log_IT1)	1.000
	(0.000)	(0.000)		(0.000)
Elec_m_log	0.175	0.155	s(Elec_m_log)	2.601
	(0.000)	(0.000)		(0.000)
R^2	0.920	0.919	R^2	0.926
n	113	113	n	113

Note: Reported are the regression coefficients or the equivalent degrees of freedom (edf) in the case of the spline variables indicated by $s(\cdot)$. Stated in parentheses are p-values of the t-statistics or the F-statistics for the significance of the respective splines. In the case of OLS regressions the adjusted R^2 is reported. Renaud and Victoria-Feser (2010) explain the R^2 measure used in the case of the KS regressions.

The GAM estimates point to a linear influence of log_IT1 and a nonlinear effect of Elec_m_log as can be seen from figure 4.3. The nonlinear effect shows a similar degressive course as in figure 4.1, again indicating that countries with a better electricity supply tend to have a more developed ICT infrastructure.

The results of the bolasso procedure lead to exactly the same variable selection and therefore to the same results as the three-stage procedure. Therefore, we need not show the corresponding table and figure at this point. While 14 variables (among 9 significant) were necessary to obtain an explanatory power of about 0.9 for the regressions of period 2002-2012, only two variables already achieve an explanatory power of 0.92 for the subperiod of 2008-2012. Clearly, many of the effects are already incorporated in the lagged IT variable as an explanatory variable.

Figure 4.3: GAM Results for the Three-Stage Procedure (dependent variable is $\ln IT_2$)



4.5.3 Explaining ICT Infrastructure Growth During 2002-2012

Finally, we consider the results with $\Delta \ln IT$ as dependent variable, the average growth rate of ICT infrastructure during 2002-2012. As shown in table 4.4, the Lasso selects 8 variables to explain the growth rate of IT, whereas 6 variables are statistically significant coefficient estimates (at 5 percent level of significance). These are $\ln IT_1$ (the log IT variable averaged over the previous period 2008-2012), again Elec_m_log (access to electricity as percent of population), for the first time the index of financial freedom, the score of freedom from corruption, the standard deviation of the capital stock (KS estimator only) and (also for the first time) the index for property rights. Hence, Lasso found next to some familiar variable, also some new variables. The explanatory power of both, the OLS and the robust KS regression is again substantial. Describing the average growth rate of IT during 2002-2012, variables from the category concerning the extend of regulation dominate the selection.

Table 4.4: Regression Results for the Three-Stage Procedure (dependent variable is $\Delta \ln IT$)

	OLS	KS		GAM
c	2.656	2.616	c	1.000
	(0.000)	(0.000)		(0.000)
log_IT1	-0.700	-0.705	s(log_IT1)	1.000
	(0.000)	(0.000)		(0.000)
Elec_m_log	0.183	0.169	s(Elec_m_log)	2.911
	(0.000)	(0.000)		(0.000)
EMP_m	0.000	0.000	s(EMP_m)	1.927
	(0.548)	(0.073)		(0.099)
EMP_sd	2.005	2.254	s(EMP_sd)	1.000
	(0.272)	(0.086)		(0.046)
financ_freedom_m	0.003	0.002	s(financ_freedom_m)	1.000
	(0.046)	(0.023)		(0.006)
free_corrupt_m_log	0.145	0.159	s(free_corrupt_m_log)	1.000
	(0.004)	(0.000)		(0.001)
K_sd	1.139	1.078	s(K_sd)	1.451
	(0.052)	(0.025)		(0.011)
prop_rights_m_log	-0.169	-0.151	s(prop_rights_m_log)	1.887
	(0.030)	(0.003)		(0.000)
R^2	0.961	0.963	R^2	0.966
n	113	113	n	113

Note: Reported are the regression coefficients or the equivalent degrees of freedom (edf) in the case of the spline variables indicated by $s(\cdot)$. Stated in parentheses are p-values of the t-statistics or the F-statistics for the significance of the respective splines. In the case of OLS regressions the adjusted R^2 is reported. Renaud and Victoria-Feser (2010) explain the R^2 measure used in the case of the KS regressions.

The negative regression coefficient of log_IT1 indicates a catch-up effect. Countries with a lower level of ICT infrastructure have higher growth rates in ICT infrastructure and therefore catch-up to countries with highly developed ICT infrastructure. A similar reason might be for the presence of the capital stock's standard deviation. Due to the catch-up in investments, the

standard deviation of the capital stock in less developed countries is higher than in well-developed countries.

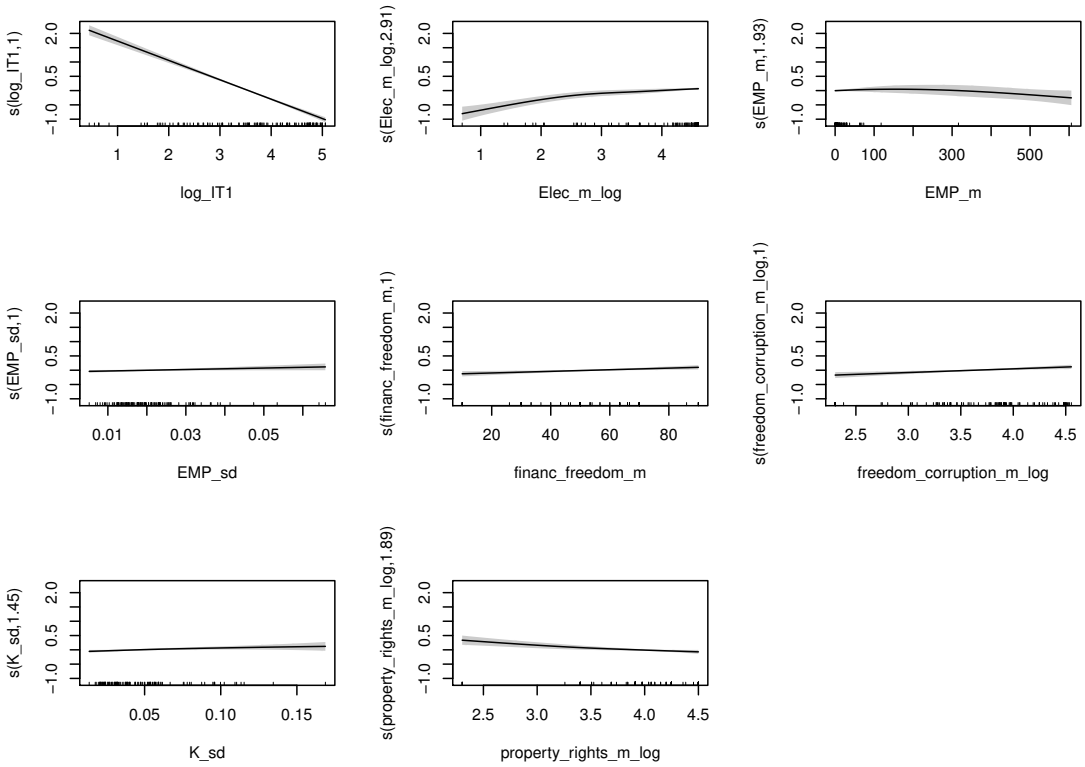
The variable of access to electricity (Elec_m_log) also significantly explains the average growth rate of IT. Again, the fundamental importance of electricity as an infrastructural prerequisite of ICT is pointed out.

The index of financial freedom appears for the first time. This index is calculated and provided by the Heritage Foundation and a measure of banking efficiency, independence from government control and interference in the financial sector. The basic idea of the score is that the financial environment ideally has a minimum of governmental interference, a minimum of regulation of financial institutions and an independent central bank supervision. The index of financial freedom scores an economy's financial freedom by analyzing: the extent of government regulation of financial services, the degree of state intervention in banks and other financial firms through direct and indirect ownership, the extent of financial and capital market development, government influence on the allocation of credit and openness to foreign competition.[66] The index ranges from 0 (repressive, private financial institutions are prohibited) to 100 (negligible government interference). The link between financial freedom and ICT infrastructure is actually similar as with the variable of investment freedom. The more extensive the government interference in banking and financing environment, the less credits are lent and deposits are accepted. Thus, less investment take place in general. Among these unrealized investments are those in ICT infrastructure, but also investments in those products that require ICT infrastructure.

The variable of property rights also appears for the first time. It is also calculated and provided by the Heritage Foundation and indicates "the ability of individuals to accumulate private property, secured by clear laws that are fully enforced by the state. It measures the degree to which a country's laws protect private property rights and the degree to which its government enforces those laws. It also assesses the likelihood that private property will be expropriated and analyzes the independence of the judiciary, the existence of corruption within the judiciary, and the ability of individuals and businesses to enforce contracts."[67] The score ranges from 0 (no private property) to 100 (private property is guaranteed by the government). Intuitively one might think that a higher score in a country's property rights has a positive effect for the growth rate of ICT infrastructure in a country. However, the regression coefficient of variable prop_rights_-m_log is negative (see table 4.4). We could interpret the negative coefficient as an indication of a catch-up effect of countries with a low score of property rights and higher growth rates of ICT infrastructure to the group of well developed countries. But this assertion cannot be answered with certainty here. In the literature, Crenshaw and Robinson (2006) find property rights significantly positive, predicting global internet diffusion in the period of 1995-2000. In their analysis on computer imports per worker during 1970-1990, Caselli and Coleman (2001) find considerable evidence that computer adoption is enhanced by good property-rights protection.

[66] Source of this information: http://www.heritage.org/index/financial-freedom.
[67] Source of the cite: http://www.heritage.org/index/property-rights.

Figure 4.4: GAM Results for the Three-Stage Procedure (dependent variable is $\Delta \ln IT$)



Again, the edf values in table 4.4 point to linear influences of nearly all variables except of Elec_-m_log. As can be seen in table 4.4 the nonlinear effect of Elec_m_log shows a degressive course, which indicates that countries with a better electricity supply tend to have a more developed ICT infrastructure.

Table 4.5: Regression Results for the bolasso Procedure (dependent variable is $\Delta \ln IT$)

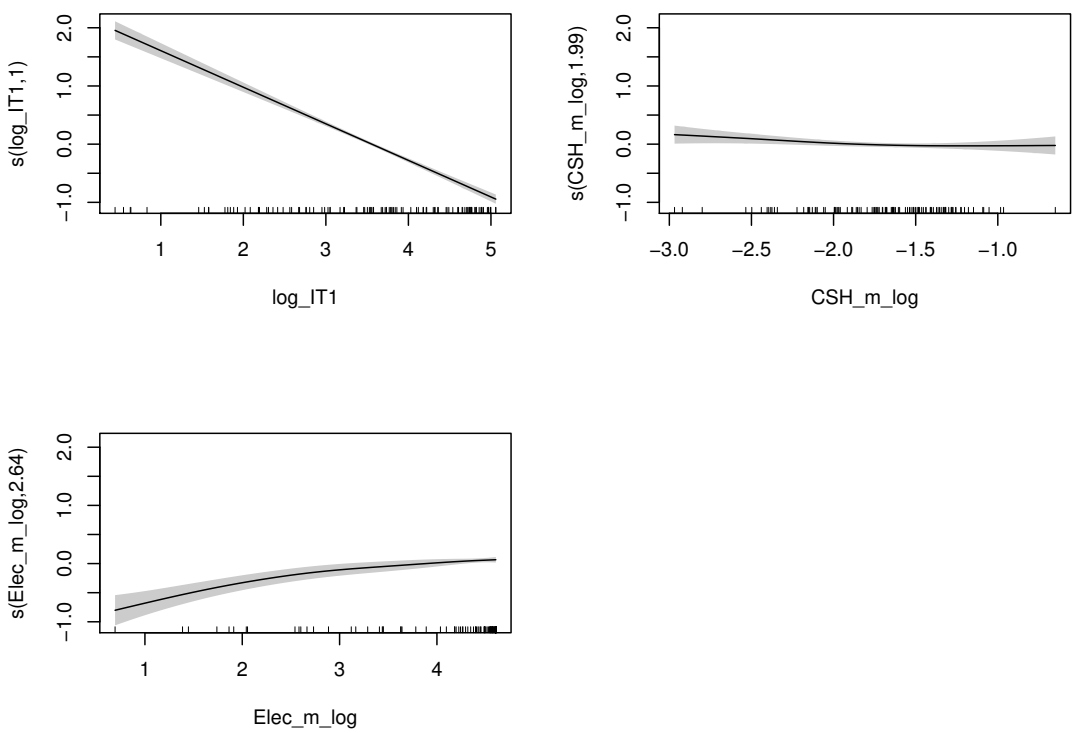
	OLS	KS		GAM
c	2.391	2.447	c	1.000
	(0.000)	(0.000)		(0.000)
log_IT1	-0.636	-0.630	s(log_IT1)	1.000
	(0.000)	(0.000)		(0.000)
CSH_m_log	-0.101	-0.105	s(CSH_m_log)	1.995
	(0.021)	(0.021)		(0.096)
Elec_m_log	0.171	0.150	s(Elec_m_log)	2.636
	(0.000)	(0.000)		(0.000)
R^2	0.951	0.955	R^2	0.955
n	113	113	n	113

Note: Reported are the regression coefficients or the equivalent degrees of freedom (edf) in the case of the spline variables indicated by $s(\cdot)$. Stated in parentheses are p-values of the t-statistics or the F-statistics for the significance of the respective splines. In the case of OLS regressions the adjusted R^2 is reported. Renaud and Victoria-Feser (2010) explain the R^2 measure used in the case of the KS regressions.

The results of the bolasso procedure (see table 4.5) show a reduced set of (now) more robust explanatory variables to describe the average growth rate of ICT infrastructure during 2002-2012. These are $\ln IT_1$ (the log IT variable averaged over the previous period 2008-2012), again Elec_-m_log (access to electricity as percent of population) and (for the first time) the share of gross capital formation (at current purchasing power parity), a variable of the category describing the economic status and structure. These three variables obtain an explanatory power of about 0.95, which is is (again) substantial.

While $\ln IT_1$ and Elec_m_log have already occurred in the preceding three-stage procedure, CSH_m_log (share of gross capital formation) was selected by bolasso as robust explanatory variable for the first time. Like $\ln IT_1$, the regression coefficient of CSH_m_log has a negative sign. Similar to the variable of property rights, we can suppose a potential indication of a catch-up effect. In general, the gross capital formation also includes investments in ICT infrastructure as well as investments in products requiring a certain degree of ICT infrastructure. Countries with a high share of gross capital formation have already realized complementary investments in ICT infrastructure. For countries with a comparatively low share of gross capital formation it means that a development of gross capital formation is associated partially with investments in ICT infrastructure. Hence, the growth rate of ICT infrastructure is comparatively higher for these countries.

Figure 4.5: GAM Results for the bolasso Procedure (dependent variable is $\Delta \ln IT$)



The GAM regression results indicate only a nonlinear effect of Elec_m_log. As before, the curving of the spline in figure 4.5 has the same shape as in the preceding GAM regressions. It

should be recalled here that this effect is driven by the large heterogeneity of the electricity supply variable across countries (see the rugs at the bottom of the right-hand panel of the figure). The association is weaker (the curve flatter) for the more highly developed countries with a better electricity support system clustered at the upper end of the scale.

In the variable selection explaining ICT infrastructure growth during 2002-2012, variables concerning the extent of regulation play a dominant role. This in in contrast to the first two subsections, where variables explaining national economic wealth and structure as well as geographical/regional variables are selected by Lasso. Similar to the preceding regressions, the access to electricity is selected in both the three-stage procedure and bolasso to explain ICT infrastructure growth during the decade. Once more, the importance of the infrastructural prerequisites of ICT infrastructure is pointed out. As in all other regressions, no human capital describing variables are selected by one of the variable selection methods.

4.6 Summary

In this chapter we have investigated economic and institutional determinants of ICT infrastructure. The analysis discussed in this chapter reveals that a set of explanatory variables, selected from a wide array of candidate variables, is very well able to explain cross-country differences in ICT infrastructure for a broad sample of more than 100 countries at all stages of development.

We can in particular show that real income per capita, electricity usage, urbanization, indicators of regulatory and institutional aspects as well as regional dummies are major determinants of ICT infrastructure. The explanatory variables are selected from a broad set of candidate variables by variants of the Lasso approach which have been developed in the machine learning literature.

Jointly, these variables achieve a very high degree of explanatory power. We find the results to be robust to heteroskedasticity and outlying observations. The former is assured by using a heteroskedasticity correction of the standard errors, while the latter is checked by comparing the least squares coefficient estimates to those of a robust regression estimator. We also applied a semiparametric GAM estimator and uncovered nonlinear effects for some explanatory variables, i.e. electricity usage. The vast majority of the explanatory power, however, originate from the linear effects of the regressors.

Although several human capital indicators are included in the set of candidate variables, interestingly none of them were selected. In a further analysis, splitting the sample period into two subperiods we can also establish conditional convergence of the ICT variable, which may be taken as evidence against the 'global digital divide'.

5 The Role of ICT in Long-Term Growth

5.1 Motivation

In this chapter, we examine the research objective of whether there is a positive and significant relationship between ICT and long-term economic growth across countries. This relationship has been the subject of numerous studies since the well-known quotation from Solow ("You can see the computer age everywhere but in the productivity statistics", Solow 1987). Most studies have investigated this relationship for the U.S.; only rare research addresses this subject in a global framework. The reason for this is undoubtedly the lack of available ICT data, especially for developing countries (see section 3.2). However, we have constructed a variable of the ICT infrastructure in section 3.2 which allows us to cover more than 100 countries at any stage of development. In the previous chapter 4 we saw that per capita income is a major determinant of ICT infrastructure. Against the background of the worldwide spread of ICT (see section 3.3) and the reputation of ICT to be a GPT (see section 2.1), we now examine whether ICT is in turn a determinant of per capita income growth.

Since the digital revolution has taken place at varying speeds in different countries, we are particularly interested in whether the impact of ICT on economic growth has taken place in the long term. We therefore investigate the effect over a period of 30 years (1980-2010) since its first appearance in the scientific literature. By doing this, we extend the empirical cross-country literature on this topic with an investigation which – for the first time – covers a period of more than 25 years and more than 95 countries at different stages of development.

To investigate the role of ICT in economic growth, we add the ICT infrastructure variable to a cross-country linear growth regression model. We use the specification of Mankiw, Romer and Weil (1992), hereafter denoted as MRW, which augments the growth model of Solow (1956) by additionally considering human capital accumulation. In turn, the addition of the ICT proxy represents an extension of the MRW model. From the consideration of ICT in the growth regression model, we expect two insights. Firstly, we are interested in determining whether the ICT variable positively and significantly explains the per capita growth during the investigation period. Secondly, we are interested in investigating whether the inclusion of ICT increases the proportion of variance explained in the growth model. We expect the latter in the context of ICT to be a (potential) GPT, which affects technological progress.

By adding the ICT variable to the MRW model, a potential endogeneity problem raises, since the ICT variable used is only available for the last decade of our research period. On the one hand, we have found in the literature review in section 2.3 that the majority of researchers find empirical evidence that ICT has a significant and positive effect on GDP per capita. On the other hand, in the previous chapter 4 we found that real income per capita is one of the major determinants of ICT infrastructure. Therefore, it can be suggested that ICT (infrastructure) and economic development are jointly determined, which induces reversed causality and, thus, endogeneity. In OLS regressions, the suspected endogeneity can lead to biased and inconsistent estimates. In order to prevent this, we apply an instrumental variable (IV) approach. This approach allows for a consistent estimation of the MRW model.

The following investigation begins in section 5.2, in which we give a short overview of the relevant literature that is related to this chapter. Based on this literature overview, we identify further research demands and motivate our research approach. The subsequent section 5.3 addresses the empirical framework. This includes an introduction to the MRW estimation framework, which we augment by the aspect of ICT. In the further course of section 5.4 we present the employed data, and in particular the instrumental variables used. In section 5.5 we analyze and interpret the empirical results of the different (IV-) estimators applied. Based on the results, we test for the existence of endogeneity. Finally, the results will be summarized in section 5.6.

5.2 Contribution of ICT to Economic Growth

Since Robert Solow's remark in 1987, the relationship between information technology and growth has been investigated in an abundance of studies, which we reviewed in section 2.3. The literature review reveals that the contribution of ICT to productivity and economic growth has been investigated primarily in the U.S. In these studies, the majority of researchers agree on the importance of ICT for the U.S. growth resurgence since the mid-1990s (see e.g. Jorgenson et al. 2002, Oliner and Sichel 2002 and Daveri 2003). The studies conducted for other countries reveal that the findings from the empirical literature for the U.S. do not necessarily apply to other countries. Authors like Daveri (2002), van Ark et al. (2008) and Inklaar et al. (2008) find substantial differences in the extent of the productivity-enhancing effects of ICT between the U.S. and the EU as well as considerable differences between EU countries. These differences are also shown by studies comparing countries at different stages of development (see e.g. Papaioannou and Dimelis 2007).

In the following we will review those studies which are thematically related to the investigation in this chapter. These are studies that examine the impact of ICT on productivity and economic growth across countries. At country-level, the literature on this issue can be divided into two streams. In one stream, the studies employ the growth accounting technique to estimate the contribution of ICT investments (in percentage points) to GDP growth. In the other stream, the studies estimate production functions to assess the effects of ICT on economic growth. The latter test the estimated elasticities on its statistical significance (see subsection 2.2.2). Against the background of the research question of this chapter whether ICT has a significant influence on economic growth, we therefore limit our presentation to those studies that also test for this significance. Consequently, this excludes those studies that use a growth accounting approach, as it does not provide the ability to test for statistical significance. In our brief literature overview, we also include studies that only partially cover ICT (e.g. telephone, broadband internet). The short presentation of the studies is given in chronological order of publication.

The work of Hardy (1980) is one of the first in this research field. He measures the role of telephone usage on economic development for a group of 60 countries in the period 1960-1973. The regressions show a significant impact of telephone lines on GDP per capita. Madden and Savage (1998) examining a sample of 27 European countries during the period 1990-1995, reveal a strong positive relationship between telecommunication infrastructure investment and economic growth. Datta and Agarwal (2004) investigated the role of telecommunication infrastructure

in economic growth for a sample of 22 OECD countries in the 1980-1992 period. Using panel data and a dynamic fixed-effects method, their results show that telecommunications are both statistically significant and positively correlated with growth in real GDP per capita growth for these countries. The results are robust even after controlling for investment, government consumption, population growth, openness, past levels of GDP and lagged growth. Becchetti and Adriani (2005) examined the impact of ICT on the level and growth rate of per capita income for up to 92 countries by applying a MRW model to panel data for two periods (1991-1997 and 1983-1997). As in our analysis, the authors use penetration levels of communication technology as proxy for ICT. The results of their analysis confirm a significant and positive role of ICT diffusion in explaining levels and growth rates of income per worker. Papaioannou and Dimelis (2007) use an adjusted GMM estimator, based on Arellano and Bond (1991), on a sample of 22 developed and 20 developing countries, covering the 1993-2001 period. They find a positive and significant ICT growth effect in both subsamples, whereby they find the impact to be higher for developed countries. Koutroumpis (2009) uses a macroeconomic production function with a micro-model to investigate the effect of broadband penetration on economic growth for 22 OECD countries in the 2002-2007 period. He finds a significant causal positive link between the broadband penetration and economic growth, especially when a critical mass of infrastructure is present. Venturini (2009) investigates the impact of ICT capital on GDP growth in the U.S. and the EU-15 members over the 1980-2004 period. He estimates a production function framework employing a cointegration procedure. The results show that ICT is a robust, long-run determinant of income levels for these modern knowledge-based societies. Röller and Waverman (2001) examine 21 OECD countries in the period 1970-1990 by estimating a micro-model for telecommunications investment with a macro production function. They find a strong causal relationship between telecommunications infrastructure and productivity above a certain threshold of telecommunications services. Czernich et al. (2011) estimate the effect of broadband infrastructure on economic growth by using an annual panel of 25 OECD countries in 1996-2007 period. They find that the introduction and penetration of broadband has a significant positive effect on economic growth. Their IV regression results suggest that a 10 percentage point increase in the broadband penetration rate resulted in a 0.9-1.5 percentage point increase in annual per capita growth. Vu (2011) identifies a strong association between ICT and growth by using a cross-country GMM dynamic panel analysis for the period 1996-2005. He also finds the marginal effect of ICT penetration is larger when at its lower level. Yousefi (2011) investigates the contribution ICT capital to economic growth using time-series cross-country data of a total of 62 countries in the period 2002-2006. He finds a significant and positive effect of ICT capital to per capita GDP growth. The estimated coefficients indicate that a 1% increase in ICT capital results in, on average, 0.22% increase in the rate of output growth in the investigation period.

Furthermore, there are a number of studies that do not find a significant influence of ICT on economic growth. Dewan and Kraemer (2000), analyzing panel data of 36 countries over the 1985-1993 period, reveal that returns from IT capital investments are not statistically significant for the developing countries. Pohjola (2002), examining data on a sample of 43 countries over the period of 1985-1999, find no significant correlation between ICT investment and economic growth. Jacobsen (2003), using data from 84 countries over 10 years between 1990-1999, finds

no significant growth effect from computer penetration, but confirmed a significant positive link between mobile phones and growth.

Summarizing this brief review of the relevant literature, it can be concluded that there are different and sometimes even contradictory results regarding the contribution of ICT to economic growth. There are obviously four reasons for this:

1. the studies use different measures to determine ICT,
2. different methods (such as functional forms of a production function) are used,
3. the studies differ in terms of the periods analyzed,
4. the studies differ in respect of the country coverage.

We consider the latter two points in more detail. Regarding the country coverage, the presented studies differ remarkably from 16 (Venturini 2009) to 92 (Becchetti and Adriani 2005) considered countries. Most studies only cover countries in the OECD and thus countries of a high level of development. However, this does not imply that ICT also impacts significantly on the economic growth of other countries, e.g. developing countries.

The investigations that take a broad sample of countries from all stages of development into account only cover a relatively short period of time. However, it is well known that the digital revolution has taken place at different times and at different speeds in the respective countries. For this reason, one may argue about whether studies that analyze only a relatively short period can capture these productivity- and growth-enhancing effects of ICT, especially in the context of cross-country regressions.

The trade-off between country and time coverage is doubtless based on the data available for ICT, especially for developing countries. By using the variable for ICT infrastructure, as composed in section 3.2, this trade-off can be overcome. We will use this variable to measure the impact of ICT on economic growth over the long-term and for a broad sample of countries at all stages of development. As the paper of Hardy (1980) suggests, the start of the economic effects of ICT can be dated back to the early 1980s. Therefore, we measure the contribution of ICT to per-capita economic growth in the 1980-2010 period, based on a well-cited cross-country linear regression model, as presented in the next section. This closes the gap in the literature by providing an investigation that examines whether a positive and significant relationship of ICT and long-term economic growth across countries can be found.

5.3 Empirical Framework

In this section we explain our empirical strategy and present the empirical framework and the data sources. This section is divided into three subsections. In subsection 5.3.1, we describe the estimation framework used, followed by a validation of the estimation model in subsection 5.3.2. In subsection 5.3.3, we address the potential problem of endogeneity due to reverse causality between GDP and ICT. Based on this, we motivate our pursued instrument variable approach.

5.3.1 Estimation Framework

The investigation approach in this chapter is to measure the contribution of ICT infrastructure to long-term economic growth within the framework of a cross-country growth analysis in the 1980-2010 period. For this purpose we include the variable for ICT infrastructure (as constructed in section 3.2) to the commonly used cross-country regression model of MRW (see MRW 1992). The MRW model augments the neoclassical model of Solow (1956) by considering human capital. By including the ICT infrastructure variable we, in turn, extend the MRW model.

Being a popular structural model for the evaluation of long-term growth across countries, the MRW model has been modified by several authors since its appearance in 1992. These modifications concerned either the model structure or the usage of different methods and approaches to solve the model. Extensions of the MRW model have been conducted by several authors, e.g. Knowles and Owen (1995) by adding health capital, Ram (2007) by including IQ measure or Aixalá and Fabro (2007) by institutional indicators. The objective of these modifications has often been to increase the explanatory power of the model.

The standard Solow model of growth is based on the aggregated production function of Cobb-Douglas type with constant returns to scale. MRW augment the model by adding human capital as further production input. The extended production function is of the form:

$$Y_{it} = A_{it} K_{it}^{\psi} H_{it}^{\eta} L_{it}^{1-\psi-\eta}, \tag{16}$$

where Y_{it} denotes the real output, K_{it} denotes the stock of physical capital, H_{it} represents the stock of human capital, L_{it} denotes the supply of labor, while A_{it} represents the technical progress of country i in time period t. Furthermore ψ and η measure the output elasticity with respect to physical capital and human capital, respectively. MRW assume constant exponential rates for labor and technology:

$$L_{it} = L_{i0} \cdot e^{n_i t}, \tag{17}$$

$$A_{it} = A_t = A_0 \cdot e^{gt}, \tag{18}$$

where n_i is the exogenous rate of growth of the labor force in country i and g is the exogenous rate of technology growth. The latter is assumed to be constant across countries. Thus, it can be derived that physical capital and human capital expressed in effective units of labor evolves as follows:

$$\dot{k}_{it} = s_{ki} y_{it} - (n_{it} + g_t + \delta_t)\, k_{it}, \tag{19}$$

$$\dot{h}_{it} = s_{hi} y_{it} - (n_{it} + g_t + \delta_t)\, h_{it}, \tag{20}$$

where the small letters – $k_{it} = \frac{K_{it}}{A_{it} L_{it}}$, $h_{it} = \frac{H_{it}}{A_{it} L_{it}}$ and $y_{it} = \frac{Y_{it}}{A_{it} L_{it}}$ – denote quantities per effective labor unit. s_{ki} and s_{hi} present the rate of accumulation of physical and human capital of country

i, respectively. Additionally, both types of capital depreciate at the same rate δ_i. The existence of diminishing returns to capital implies that $\psi + \eta < 1$. Under these initial conditions, the capital follows a convergence path to the steady state (k_i^*, h_i^*) given by the system of equations:

$$k_i^* = \left(\frac{s_{ki}^{1-\eta} s_{hi}^{\eta}}{n_i + g + \delta} \right)^{\frac{1}{1-\psi-\eta}}, \tag{21}$$

$$h_i^* = \left(\frac{s_{ki}^{\psi} s_{hi}^{1-\psi}}{n_i + g + \delta} \right)^{\frac{1}{1-\psi-\eta}}. \tag{22}$$

Substituting equation (21) and equation (22) into the production function and taking logs we could express the equilibrium level of income per capita in two alternative ways. Firstly, as a function of investments in human capital s_h:

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \ln A_0 + gt - \frac{\psi+\eta}{1-\psi-\eta} \ln\left(n_i + g + \delta\right)_i + \frac{\psi}{1-\psi-\eta} \ln\left(s_{ki}\right) + \frac{\eta}{1-\psi-\eta} \ln\left(s_{hi}\right). \tag{23}$$

Secondly, as a function of the human capital level h_i^*:

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \ln A_0 + gt - \frac{\psi}{1-\psi} \ln\left(n_i + g + \delta\right)_i + \frac{\psi}{1-\psi} \ln\left(s_{ki}\right) + \frac{\eta}{1-\psi} \ln\left(h_i^*\right). \tag{24}$$

For estimation the choice between equation (23) and equation (24) depends on "whether the available data on human capital correspond more closely to the rate of accumulation [...] or to the level of human capital" (MRW 1992, p. 418).

The short-run dynamics that is the convergence of income per effective labor to its steady-state level is given by:

$$\ln\left(y_{it}\right) - \ln\left(y_{i0}\right) = \theta \ln\left(y_i^*\right) - \theta \ln\left(y_{i0}\right), \tag{25}$$

where $\theta = \left(1 - e^{-\lambda_i t}\right)$, and λ_i measures the rate of convergence to the long-term equilibrium. Equation (25) implies that the change of income per effective labor is a function of the determinants of both the steady state y^* and the initial initial level y_{i0} of income per effective unit of labor. Substituting for the steady state expression y^* in equation (25) we get:[68]

$$\ln\left(y_{it}\right) - \ln\left(y_{i0}\right) = \frac{\theta\psi}{1-\psi-\eta} \ln\left(s_{ki}\right) + \frac{\theta\beta}{1-\psi-\eta} \ln\left(s_{hi}\right) - \frac{\theta\left(\psi+\eta\right)}{1-\psi-\eta} \ln\left(n_i + g + \delta\right)_i - \theta \ln\left(y_{i0}\right), \tag{26}$$

The growth equation (26) has been estimated by MRW in their examination for a cross section of countries in the period 1965-1980. According to Ram (2007), the regression model used by MRW (1992, p. 426, Table V) can be written as:

[68] See MRW (1992, pp. 422-423) for a detailed explanation of the procedure.

$$\ln(\tfrac{Y_{1985}}{Y_{1960}})_i = \alpha + \beta_1 \ln(Y_{1960})_i + \beta_2 \ln(n+g+\delta)_i + \beta_3 \ln(I/GDP)_i + \beta_4 \ln(\text{School})_i + u_i, \quad (27)$$

where $(Y_{1960})_i$ and $(Y_{1985})_i$ denote the average income of working-age persons from country i in the years 1960 and 1985. Furthermore, n_i is the growth rate of the working-age population, g the rate of technical change and δ the depreciation rate of physical capital. The value of $g+\delta$ is usually assumed to be 0.05 and constant across the countries (see e.g, MRW 1992, Knowles and Owen 1995). In the investigation of MRW (1992), $(I/GDP)_i$ denotes the average ratio of investment over the period 1960-1985 as proxy for physical capital investment (s_{ki}). $(\text{School})_i$ denotes the average percentage of the working-age population in secondary school over the period 1960-1985 as proxy for human capital investment (s_{hi}). Furthermore, u_i denotes the error term.

In order to serve our purposes, we modify the model equation (27) in three ways. The first modification concerns the observation period, which we change to 1980-2010. The second modification concerns the variable School, which is only available for an insufficient number of countries in the context of our research. Other authors, such as Bechetti and Adriani (2005), use the average schooling years as proxy for human capital investment. This indicator, however, does not take potential decreasing returns to years of schooling into account. For this reason, we use the indicator of human capital per worker as suggested by Hall and Jones (1999), which is constructed by the average years of schooling and an assumed rate of return to education. This variable will be formally described in section 5.4.

According to the first two modifications, the model equation can be written as:

$$\begin{aligned}\ln(\tfrac{Y_{2010}}{Y_{1980}})_i = \alpha + \beta_1 \ln(Y_{1980})_i + \beta_2 \ln(n+g+\delta)_i + \beta_3 \ln(I/GDP)_i \\ + \beta_4 \ln(HC)_i + u_i,\end{aligned} \quad (28)$$

where, analogously, $(Y_{1980})_i$ and $(Y_{2010})_i$ denote the average income of working-age persons of country i in the years 1980 and 2010, and $(HC)_i$ denotes the human capital per worker.[69]

As a third modification, we introduce the ICT infrastructure variable in the model of equation (28). As one of the first, Nonneman and Vanhoudt (1996) propose a further augmentation of the model by explicitly including the (endogenous) accumulation of technological know-how. They suggest including other types of capital (e.g., infrastructure, equipment, other physical capital, human capital, know-how) in order to increase the explanatory power of the model. They further suggest considering technological know-how (in the sense of blueprints for production processes and new products) as any other input in production. The ICT variable, as used in this dissertation, applies to several of the extension types named by Nonneman and Vanhoudt (1996). Thus, the ICT variable is a proxy for infrastructure, equipment (e.g. end-devices such as PCs) and know-how (due to its ability to reduce information asymmetries). For this reason, the addition of ICT is a meaningful extension of the MRW model. By including the ICT infrastructure variable to the model, the model equation is given by:

[69] At this point we are leaving open from which year or years the human capital variable will be calculated.

$$\ln(\tfrac{Y_{2010}}{Y_{1980}})_i = \alpha + \beta_1 \ln(Y_{1980})_i + \beta_2 \ln(n+g+\delta)_i + \beta_3 \ln(I/GDP)_i + \beta_4 \ln(HC)_i + \beta_5 \ln(\text{ICT})_i + u_i. \tag{29}$$

We use the model equation, as given in equation (29), in the context of our empirical analyses in this chapter to assess the relationship between ICT and economic growth. In order to obtain robust results, we will also include other known growth determinants as control variables in the model. This serves to examine whether ICT influences growth only under particular economic, financial, institutional and/or policy environments. We will explain the variable sources in section 5.4.

As previously mentioned, this analysis suggests an endogeneity problem due to reverse causality between GDP and ICT. To prevent the suspected endogeneity problem, we will apply an instrumental variable approach, as will be explained in subsection 5.3.3. However, a form of reverse causality can also be assumed from the explanatory variables of capital investment and human capital. For instance, Hanushek and Woessmann (2012) suggest that growth provides added resources that can be used to improve schools. Hence, this could lead to higher human capital. For that reason, we use the values of human capital and the investment ratio from the initial year 1980. This serves to avoid further endogeneity problems and to prevent biased results of the (IV) estimates.[70]

5.3.2 Validation of the Estimation Framework

Due to the modifications we have made – in particular by using initial values of investment ratio and human capital – our results differ from those of the original MRW regressions. These differences concern both the significance and the magnitude of the estimation coefficients. In general, there are two reasons for these differences. Firstly, we have operationalized the MRW regression model with several other variables. Some of these stem also from other data sources (such as the Penn World Table). Secondly, unlike the original paper of MRW (1992), a different period is examined. Therefore, it is also possible that the MRW model cannot be applied robustly for the period 1980-2010. In order to verify the applicability of the model under the changed conditions, we will gradually transfer the original MRW model (for the 1960-1985 period) into the 1980-2010 period with modified variable operationalization. We discuss the results subsequently.

To regress their growth model, MRW use data from Real National Accounts constructed by Summers and Heston (1988). They build three samples to investigate the estimation results. In the first sample, they use all countries with available data but exclude countries with oil as the dominant industry. They obtain a sample of 98 (non-oil producing) countries. The second sample of the 75 intermediate countries excluding all countries whose real income figures are based on little primary data and countries with a population of less than one million in 1960. The third sample contains 22 countries from the OECD with a population greater than one million in 1985.

[70] Since the per capita growth of 1980-2010 cannot affect human and physical capital of 1980, the problem of reverse causality is being avoided.

Given the data-set in MRW (1992, pp. 434-436), we reestimate the regression results of the three samples. Table 5.1 presents the regression results from the orginial MRW estimation (see MRW 1992, p. 426, table V). In comparison to their result table, the reconstructed results differ minimally. Based on the available data-set, we also include a column with the regression results from all 104 available countries. In all four cases the GDP of the initial year 1960 significantly explains the per-worker growth in the period 1960-1985. The negative sign indicates a catch-up process of poorer, less developed countries. Also, the term of $\ln(n+g+\delta)$ and the investment ratio have significant explanatory power for GDP growth. The explanatory variable School explains the per-worker income growth in three of four cases, except for the OECD sample. The models explain between 43.5% and 65.1% of the dependent variable variance. In the following we will focus on the fourth sample, which contains all available countries.[71]

Table 5.1: Regression Results of the Original MRW Model (reconstruction)

	Non-Oil	Intermediate	OECD	all
dependent variable	log difference GDP per working-age persons 1960-1985			
c	3.022	3.709	2.755	3.113
	(0.000)	(0.000)	(0.035)	(0.002)
$\ln(Y_{1960})$	-0.288	-0.366	-0.398	-0.297
	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(n+g+\delta)$	-0.506	-0.545	-0.863	-0.507
	(0.083)	(0.063)	(0.020)	(0.047)
$\ln(I/GDP)$	0.524	0.538	0.332	0.553
	(0.000)	(0.000)	(0.073)	(0.000)
$\ln(School)$	0.231	0.270	0.228	0.216
	(0.000)	(0.001)	(0.135)	(0.000)
N	98	75	22	104
$\bar{R}^2$	0.463	0.435	0.651	0.496

Note: Reported are the regression coefficients and the p-values. The investment and population growth rates are averaged over the period 1960-1985. Variable *School* denotes the average percentage of working-age population in secondary school for the period 1960-1985.

We now gradually transfer the initial MRW model to the 1980-2010 period and specify the variables we use. The estimation results of this stepwise modified MRW model are shown in table 5.2. In the first step (column (2)), we use variables from Penn World Table 8.0 (PWT) to replace the dependent variable – the logarithmic difference of the GDP per working-age persons – with the logarithmic difference of the GDP (series *RGDPO*)[72] per engaged persons (series *emp*), the initial GDP level per working-age persons of 1960 and the population growth n of the MRW data-set. The PWT contains national-accounts data of 167 countries. Up to 1950, data on GDP, capital, employment and population are available. The PWT data-set does not contain data per working-age persons. We also use data from the World Bank to replace the investment ratio by the gross capital formation as percentage of GDP. Merely 90 of the 104 countries in the MRW

[71] The above-mentioned argumentation of MRW to exclude oil-producing countries from the investigation seems plausible to us. Nevertheless, we will also include this group of countries in our analysis, because we want to investigate the impact of ICT on economic growth using the broadest possible sample of countries.

[72] The PWT 8.0 distinguishes between expenditure-side and output-side real GDP. In this paper, we use the output-side real GDP at chained PPPs.

data-set could be matched with the data-sets of the World Bank and PWT. Unfortunately, the number of engaged persons in 1960 is missing for several countries in the PWT data-set. For that reasons, the dependent variable can not be determined for 14 countries. As expected, the regression coefficients in column (2) differ from the coefficients in column (1).

In the next step, we use data from the Barro and Lee (2013) database in addition to the data of PWT 8.0. The database contains data of the country population by age groups. Through the use of these data from Barro and Lee, the dependent variable, the initial income and n can be calculated as in the MRW by working-age persons in the age between 15-64. The estimation results are shown in column (3). The coefficients of the initial income and human capital become smaller while the coefficient of the intercept rises. The $\bar{R}^2$ slightly decreases to a level of 0.49. The major difference to model (1) is that the term $\ln(n + g + \delta)$ no longer has a significant explanatory power for this model and also the following models. This seems to be due to the operationalization of the variables. In several tests we have found that the original variables of MRW from Summers and Heston (1988) differ considerably from other data of commonly used databases (as the databases of Barro and Lee or World Bank). Since $g+\delta$ is assumed to be 0.05, the values of $\ln(n + g + \delta)$ across countries differ solely by population growth. Since population growth is not a source of long-term growth according to general growth theory, the insignificance of the term in the model is not of any further importance to us.[73]

To mitigate wide fluctuations of the respective variables caused by economical, meteorological or political fluctuations as well as armed conflicts (e.g. civil wars) in several countries, we use the GDP per working-age persons values as average of the five preceding years. Thus, the dependent variable describes the income growth in working-age persons between the average values for 1956-1960 and the average values for 2006-2010. In column (4), accordingly, the initial income $Y_{Initial}$ describes the averaged GDP per working-age person for the years from 1956 to 1960. The regression results in column (4) are similar to these of column (3) with a higher of $\bar{R}^2$ of 0.529.

In the next step (column (5)), the observation period is changed to 1980-2010. As mentioned in the previous subsection, we replace the averaged percentage of working-age population in secondary school (as used in MRW 1992) by the human capital variable according to Hall and Jones (1999). Based on the total years of schooling, they calculate rates of return for different stages of education.[74] According to the variable School, the data for human capital are averaged over the 1980-2010 observation period in model (5). The human capital variable according to Hall and Jones significantly contributes to the model. The estimated coefficient is remarkably higher as the correspondent value of the initial MRW estimation in column (1). The number of observations increases to 114, because of the higher country coverage of the variables in this observation period. The value of $\bar{R}^2$ rises through the use of the new human capital variable to a value of 0.634.

As explained in the previous subsection, we use the human capital and investment values of the initial year 1980 in order to avoid potential endogeneity problems. The results of this step are shown in column (6). We also averaged the human capital variable by average the values

[73] Czernich et al. (2011), whose research is based on the MRW model, also find no significant influence of the growth rate of the workforce.

[74] The precise calculation is described in section 5.4.

from 1975 and 1980. The coefficient value of the human capital variable differ markedly to the previous regression. The estimates coefficient is more than twice as high as in the previous model (5) and almost more than 3.5 times as high as in the original MRW model in column (1). The estimate coefficient of the investment variable only changes slightly. The $\bar{R}^2$ decreases on a value of 0.578.

Table 5.2: Regression Results of the Model Validation

	(1)	(2)	(3)	(4)	(5)	(6)
dependent variable	log diff. GDP per working-age persons 1960-85				...1980-2010	
c	3.113	0.689	1.552	1.482	0.646	0.868
	(0.000)	(0.411)	(0.040)	(0.040)	(0.306)	(0.178)
$\ln(Y_{Initial})$	-0.297	-0.213	-0.184	-0.166	-0.208	-0.196
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
$\ln(n+g+\delta)$	-0.507	-0.663	-0.236	-0.245	-0.059	-0.064
	(0.047)	(0.036)	(0.261)	(0.220)	(0.782)	(0.768)
$\ln(I/GDP)$	0.553	0.513	0.554	0.544	0.452	0.435
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(HC)$	0.216	0.201	0.171	0.195	0.304	0.752
	(0.000)	(0.001)	(0.004)	(0.001)	(0.000)	(0.000)
N	104	90	90	90	114	114
$\bar{R}^2$	0.496	0.509	0.490	0.529	0.634	0.578

Note: Reported are the regression coefficients and the p-values in parentheses. Explanatory variable $Y_{Initial}$ denotes the GDP per worker value of 1960 in model (1)-(4) and the value of 1980 in the models (5)-(6). The operationalization of the variables is described in subsection 5.3.2.

We have now modified the original MRW model in five steps for the use of growth regression in the period 1980-2010. The variables of the original model are specified with current data sources and transferred to the observation period 1980-2010. We further use levels of the investment ratio and human capital from the initial year 1980. Furthermore, we substitute the enrollment rates used in MRW by the human capital variable according to Hall and Jones (1999). The applied data, their sources and descriptive statistics are explained in section 5.4.

The estimation results of model (6) differ from that of the initial model (1) in two points. First, the term $\ln(n+g+\delta)$ does not explain the growth rate of GDP per working-age person significantly. Second, the estimated coefficient of the human capital variable is substantially higher due to the use of the human capital variable. Despite the differences between the modified growth model and that of MRW, we find it appropriate for further use. The differences do not affect the basic pattern of results of growth regression. For this reason we will use the modified model in further analysis.

5.3.3 Instrumental Variable Approach

Adding the ICT variable – as constructed in section 3.2 – to the MRW model, raises a potential endogeneity problem. As can be seen from the literature review in section 2.3, the majority of researchers find empirical evidence that ICT has a significant and positive effect on GDP per capita. Since we have found in the previous chapter that real income per capita is one of the

major determinants of ICT infrastructure, it is plausible to assume that the income level also affects the stage of ICT. This can have two reasons. First, individuals in high-income countries may have a higher ability to pay for a personal computer or broadband services (Czernich et al. 2011). This can lead to an increased demand in high-income countries, further reinforcing the expansion of ICT and ICT infrastructure (see Röller and Waverman 2001). Second, the ICT infrastructure (especially telecommunication) often underlies regulation and sectoral policies of the national administration. This intervention depends on the level of economic development and can be confounded with the growth effects of the ICT. For this reason, the effect of regulation and sectoral policies can be confounded with the effect of ICT infrastructure (Czernich et al. 2011). Hence, we can suggest a causal link between ICT infrastructure and economic development, because both variables are jointly determined.

In OLS regressions, endogeneity can lead to biased and inconsistent estimates. In general, there are two approaches to prevent this. The first possibility is to use the ICT variable from the initial year 1980 or earlier. In this case, reverse causality can be ruled out, as income growth between 1980 and 2010 cannot influence the level of ICT infrastructure in 1980 (or earlier). As mentioned in section 3.2, the ICT variable used has only been only available since 2001 and thus solely for the last decade of our research period. Hence, the first approach can be excluded. The second possibility is to apply an IV approach. In an IV approach, the ICT variable is replaced by instrumental variables (IVs), which have to fulfill the two conditions of validity (the IVs must be uncorrelated with the error term) and relevance (the IVs must be highly correlated with the ICT infrastructure variable). If both conditions are met, the IV estimator uses the variation of the ICT variable which is explained by the variation in the instruments, in the framework of an ordinary OLS estimation. Appropriate IVs are difficult to find in the growth context. In the literature, Acemoglu et al. (2001) investigate the relationship between institutions and economic growth. In their study they use two IVs:

- the average years of European settler mortality,[75]
- the absolute value of latitude (distance from the equator).

The underlying idea of using latitude as an instrument variable originates from Hall and Jones (1999). They argue with the widely known fact that the distance to the equator is negatively correlated with income per capita. For this reason, latitude seems to be a suitable IV within the scope of this growth regression model, especially since the information about the distance of each country to the equator can easily be determined. The European settler mortality, however, is only available for 66 of 126 countries and is therefore not appropriate for the analysis of a broader sample of countries.[76]

In order to apply an IV approach in this examination, we further use an additional, ICT-related IV. The number of telephone lines per 100 people of the year 1980 seems to be an appropriate IV in this context. As pointed out in section 3.2, the telephone network represents a basic ICT

[75] The mortality rates of European settlers are provided per 1,000 mean strength in the 19th century. These data can be found in appendix table A2 of Acemoglu et al. (2001).

[76] The data files can be downloaded at http://economics.mit.edu/faculty/acemoglu/data/ajr2001.

technology that has provided the first connections since the beginning of networked ICT. Served by publicly available databases, the number of telephone lines is available back to 1960. This variable fulfills the previously requested property of relevance, since the number of telephone lines is highly correlated with the other ICT-related indicators (see table 3.1) as well as the constructed variable of ICT infrastructure.[77] The other required property of validity is given by the fact that the income growth between 1980 and 2010 cannot influence the number of telephone lines in 1980.[78] The number of telephone lines from earlier years, as an IV for ICT, is also used by Czernich et al. (2011), who applied this variable as an instrument for broadband infrastructure.

In the following, we will use the absolute value of latitude and number of telephone lines per 100 people in the year 1980 as IVs in our IV regressions. Based on our argumentation, we consider our instruments to be valid ex-ante. Our assertions will be reviewed in the IV regression analysis of subsection 5.5.2 by the application of empirical tests.

5.4 Data

In this chapter we use the indicator for ICT infrastructure, whose construction we have discussed in section 3.2. We construct the variable from the averaged values of the components for the years 2001-2005, to achieve a trade-off between ICT values being constructed as near as possible to the year 1980 and the robustness of an average value over several years.[79]

The data used to operationalize the variables of the MRW model, the control variables and IVs are taken from various sources. We will first focus on the data of the MRW model, followed by a description of the variables we use to control for other potential growth determinants in the growth regressions. Lastly, we mention the sources of the IVs. The descriptive statistics and correlation matrix of all variables are shown in table C1 and C2 of the appendix.

Data on the MRW model

The dependent variable – growth of GDP per working-age persons – is calculated as the logarithmic difference of the output-side real GDP at chained PPPs per working-age persons between 1980 and 2010. The data on GDP are taken from the PWT (version 8.0).[80] The database contains information on levels of income and output, covering 167 countries between 1950 and 2011. Since version 8.0 of PWT, the database also provides GDP data constructed from the output side rather than from the expenditure side, which is more suitable for use in the context of growth analyses.[81] Therefore, we use the output-side real GDP at chained PPPs (series *rgdpo* in the PWT). The number of working-age persons is obtained from the Barro-Lee database. The database contains educational attainment data and information about population as well

[77] The correlation of the ICT infrastructure variable with the number of telephone lines per 100 people from 1980 lies at 0.91.

[78] Moreover, the number of telephone lines is the only possible ICT-specific indicator available for such an early year.

[79] The PCA is conducted using the spectral decomposition approach, which examines the correlations between the single aspects of ICT infrastructure.

[80] A detailed documentation of the database is provided in Feenstra et al. (2013).

[81] A conceptual comparison of output-side and expenditure-side real GDP is given in Feenstra et al. (2009).

as average schooling years at primary, secondary and higher levels. The data are available for 5-year age groups of the adult population age 15 and over and cover 146 countries between 1950 and 2010. According to MRW (1992), we define working-age persons as population in the age between 15 and 64. The initial income Y_{1980} describes the GDP per working-age person of the initial year 1980. We average the values of GDP per working-age persons over the five preceding years. This means that the dependent variable describes the per working-age persons income growth between the years of 1976 to 1980 and 2006 to 2010. Accordingly, $\ln(Y_{1980})$ describes the natural logarithm of GDP per working-age persons, averaged for the years of 1976 to 1980. For the sake of simplicity we will denote the term 'working-age persons' and 'worker' synonymously.

The term $\ln(n+g+\delta)$ describes a sum of growth rates, where n is the growth rate of the working-age population, g is the rate of technical change and δ is the depreciation rate of physical capital. Following MRW (1992) $g+\delta$ is assumed to be fixed and equal to 0.05 for all countries.[82] The propensity to accumulate physical capital (variable I/GDP) is proxied by the gross capital formation as percentage of GDP from the World Bank.

As already pointed out in subsection 5.3.1 we use human capital as proxy for human capital investment. As suggested by Hall and Jones (1999), human capital per worker is constructed by the average years of schooling from Barro and Lee (2013) and an assumed rate of return to education, based on Mincer equation estimates around the world (Psacharopoulos 1994). The latter is represented by returns to education, which is 0.134 up to the fourth year of education, 0.101 from the fourth to the eighth year and 0.068 beyond the eighth year. According to Hall and Jones (1999) the human capital per worker HC_i in 1980 is formally measured by

$$HC_i = \exp\left(\phi\left(s_i\right)\right) \text{ with } s_i = \begin{cases} 0.134 \cdot s_i & \text{for } 0 \leq s_i \leq 4 \\ 0.134 \cdot 4 + 0.101\left(s_i - 4\right) & \text{for } 4 < s_i \leq 8 \\ 0.134 \cdot 4 + 0.101 \cdot 4 + 0.068\left(s_i - 8\right) & \text{for } s_i > 8, \end{cases} \tag{30}$$

where s_i denotes the average years of schooling in country i in 1980. By taking the natural logarithm, we use this measure as $\ln(HC)$.

Data on the control variables

To assess the relationship between ICT and economic growth we control for other known growth determinants. These control variables also examine whether the ICT influences growth only under particular economic, financial, institutional and policy environments (Levine and Renelt 1992). Research on the robustness of explanatory variables in cross-country economic growth regressions has been undertaken in various papers such as in Levine and Renelt (1992), Sala-i-Martin (1997), Sachs and Warner (1997), Fernandez et al. (2001), Sala-i-Martin et al. (2004),

[82] MRW (1992) chose this value of $g+\delta$ to match the available data. They argued that reasonable changes in this assumption would have little effect on the estimates.

Hoover and Perez (2004), Ley and Steel (2009).[83] By using different methods,[84] the authors find different variables to be robust. According to this, they partly identified different variables. Some of these variables are already included in the MRW equation. In four of the seven studies,[85] the initial GDP level is identified as a robust explanatory variable. Sachs and Warner (1997) also identify population growth as robust variable. Sala-i-Martin et al. (2004) finds the population growth to robustly explain cross-country growth differences. For these reasons we can assume that the economic environment is sufficiently explained by the variables of the MRW model.

As a control variable for financial environment, we use the private credit by deposit money banks and other financial institutions as percentage of the GDP in 1980. These data are from the International Financial Statistics (IFS) of the International Monetary Fund (IMF) and are publicly available at the World Bank database.[86] To control for the institutional environments we use the indicator of civil liberties for the year 1980. The indicator has been considered robust by Sala-i-Martin (1997) and Fernandez et al. (2001). Civil liberties contain the freedoms of expression and belief, associational and organizational rights, rule of law, and personal autonomy without interference from the state. The countries are classified between the values of 1 (for 'most free') and 7 (for 'least free'). The data are available for an average of 172 countries since 1972 from Freedom House.[87] As the control variable for the policy environment, we use government consumption. The data are available from the Penn World table (series gc).[88]

We also average the values of all control variables for the five preceding years to mitigate fluctuations in the data and transform them by taking their natural logarithm, respectively.[89]

Data on the IVs

As mentioned in the section above, we use the absolute value of latitude and number of telephone lines per 100 people from the year 1980 as IVs in our IV regressions. The data for the absolute value of latitude are provided by the database of Socrata Open Data.[90] The number of telephone lines per 100 people of the year 1980 are provided by the World Telecommunication/ICT Development Report of the International Telecommunication Union (ITU) and are available at World Bank database.[91] Both IVs are transformed by taken their natural logarithms.

[83] An overview of variables in cross-country growth regressions is shown in appendix B of Durlauf et al. (2005).

[84] Sala-i-Martin (1997) used the Bayesian Average of Classical Estimates (BACE). At a value of $CDF(0) > 0.93$ he found a variable as robust. Fernandez et al. (2001) made use of the Bayesian Model Averaging (BMA) method. Hoover and Peret (2004) applied the LSE approach and Sala-i-Martin's (1997) variant of the extreme-bounds methodology. Lay and Steel (2009) found a variable as robust, showing a marginal posterior inclusion probabilities of the covariates > 0.9.

[85] Sala-i-Martin (1997), Fernandez et al. (2001), Sala-i-Martin et al. (2004), Ley and Steel (2009).

[86] http://databank.worldbank.org/data.

[87] https://freedomhouse.org/.

[88] http://www.rug.nl/research/ggdc/data/pwt.

[89] The 5-year average of the Civil Liberties variable could not be calculated.

[90] https://opendata.socrata.com/.

[91] http://databank.worldbank.org/data.

5.5 Results

We now turn to the presentation of the regression results. In this section we measure the contribution of the ICT to economic growth within the ICT-augmented MRW framework, as described in subsection 5.3.1. This section is structured in two subsections. In the first subsection, the ICT-augmented MRW framework is regressed using an OLS estimator. On the assumption of potential reverse causality between per-capita GDP and ICT, we apply two different IV estimators in the second subsection. In a subsequent analysis, we test the solutions of the IV estimations for relevance and validity of the IVs. [92]

Each estimator is applied to five models. In model (1) we estimate the original MRW model, augmented by the variable for ICT infrastructure. In each of the models (2) to (4) we add a control variable to test our results on robustness. We further combine all of these control variables in model (5). The country coverage of the respective models can be found in table C3 of the appendix.

5.5.1 Ordinary Least-Squares Estimates

In this subsection we present the OLS regression results of the MRW model. We use an OLS estimator with heteroskedasticity-robust standard errors with the correction of MacKinnon and White (1985). Initially, the model is regressed without the ICT variable, as described in equation (28). This serves to analyze the effects of the added control variables and to ensure that these do not distort the result pattern of the MRW model. The regression results are shown in table 5.3.

Depending on the model, the regressions rely on a total of 95 to 114 observations. Model (1) of table 5.3 corresponds to model (6) of table 5.2. In model (2) the added control variable for the share of private credit has a significant and positive effect on the dependent variable. Also, the variable controlling for civil liberties has a significant and positive effect on the dependent variable in model (4). As expected, the variable has a negative sign, since a high degree of civil liberty is indicated by a low variable value. Furthermore, both variables explain the per-worker GDP growth significantly in model (5), where all control variables are combined. As shown in model (3), the control variable for governmental consumption has no significant effect on the model. By adding a control variable, the adjusted coefficient of determination $\bar{R}^2$ rises in most of the extended models (2) to (5). Hence, the significant control variables contribute to further explanation of the MRW growth model. The control variables do not bias the general result pattern of the MRW model.

In the next step we include the ICT variable in the MRW framework, as described in equation (29). This can be seen as an augmentation of the MRW model. The regression results of the augmented MRW model are shown in table 5.4. The results show that the ICT variable explains the per-worker income growth from 1980 to 2010 significantly and positively in almost all of

[92] All computations are programmed in R using the following packages: `lmtest` and `sandwich` (for the least squares regression with the computation of variance inflation factors and the heteroskadasticity-robust standard errors), `ivreg` (for the IV regression by two-stage least squares) and `riv` (for the robust IV estimator).

Table 5.3: Regression Results of the MRW Growth Model

	(1)	(2)	(3)	(4)	(5)
dependent variable		log diff. GDP per worker, 1980-2010			
c	0.868	0.654	0.853	1.112	0.953
	(0.204)	(0.318)	(0.220)	(0.095)	(0.153)
$\ln(Y_{1980})$	-0.196	-0.205	-0.200	-0.234	-0.223
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
$\ln(n+g+\delta)$	-0.064	-0.014	-0.063	-0.235	-0.043
	(0.790)	(0.955)	(0.797)	(0.317)	(0.852)
$\ln(I/GDP)$	0.435	0.450	0.435	0.446	0.476
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(HC)$	0.752	0.683	0.756	0.615	0.559
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(PrivateCredit)$		0.154			0.124
		(0.001)			(0.011)
$\ln(Gov.Consumption)$			-0.028		-0.091
			(0.779)		(0.149)
$\ln(CivilLiberties)$				-0.209	-0.171
				(0.001)	(0.016)
N	114	97	114	108	95
$\bar{R}^2$	0.578	0.644	0.574	0.624	0.659

Note: Reported are the regression coefficients and the p-values of the heteroskedasticity consistent covariance matrix in parentheses. Explanatory variable Y_{1980} denotes the GDP per worker in 1980.

the five models. Only in model (5) we cannot find a significant effect of ICT on the dependent variable at a 5% significance level, but at a level slightly above. Hence, the ICT variable has substantial explanatory power for the original model. By adding the ICT variable the values of the $\bar{R}^2$ increase up to 8.4% percentage points in comparison to the estimation results in table 5.3. We also observe that by including the ICT variable, the coefficient of human capital nearly halves. The regression coefficient value of investment also decreases, however to a lesser extent.

Since the natural logarithms are taken from both dependent and explanatory variables, the estimation coefficients can be interpreted as elasticities. In model (1), the 0.515 elasticity implies that a 10% higher level of ICT infrastructure at the beginning of the period corresponds to 5.15% higher GDP per worker on average in the 1980-2010 period. By including the control variables the elasticities in the models (2)-(5) decrease and range in the interval 0.387-0.510.

Thus, one can argue that the ICT infrastructure has a wide influence on economic growth. However, it can not be denied that there is a potential endogeneity between the ICT infrastructure and economic growth per capita. For this reason we repeat the measurement setting in the following subsection by using IV estimators.

5.5.2 Instrumental Variable Estimates

As already mentioned above, there is a potential endogeneity between ICT and GDP per capita in the form of a reverse causality. The suspected endogeneity leads to biased and inconsistent

Table 5.4: Regression Results of the MRW Growth Model with ICT

	(1)	(2)	(3)	(4)	(5)
dependent variable		log diff. GDP per worker, 1980-2010			
c	2.007	1.880	1.981	2.056	1.900
	(0.002)	(0.013)	(0.003)	(0.002)	(0.017)
$\ln(Y_{1980})$	-0.338	-0.295	-0.342	-0.330	-0.297
	(0.000)	(0.002)	(0.000)	(0.000)	(0.006)
$\ln(n+g+\delta)$	-0.096	0.070	-0.099	-0.099	0.054
	(0.640)	(0.746)	(0.632)	(0.629)	(0.796)
$\ln(I/GDP)$	0.342	0.385	0.340	0.371	0.405
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(HC)$	0.370	0.364	0.378	0.333	0.325
	(0.000)	(0.002)	(0.000)	(0.001)	(0.004)
$\ln(PrivateCredit)$		0.068			0.059
		(0.159)			(0.263)
$\ln(Gov.Consumption)$			-0.031		-0.092
			(0.596)		(0.145)
$\ln(CivilLiberties)$				-0.071	-0.095
				(0.344)	(0.286)
$\ln(ICT)$	0.515	0.439	0.510	0.479	0.387
	(0.000)	(0.008)	(0.000)	(0.008)	(0.062)
N	102	91	102	98	89
$\bar{R}^2$	0.661	0.670	0.658	0.662	0.672

Note: Reported are the regression coefficients and the p-values of the heteroskedasticity consistent covariance matrix in parentheses. Explanatory variable Y_{1980} denotes the GDP per worker in 1980.

estimates. To address this problem, we apply an IV approach. As explained in subsection 5.3.1, the ICT variable is instrumented by two IVs, we considered to fulfill the properties of validity and relevance ex-ante. We use the absolute value of latitude, as suggested in the literature and the number of telephone lines per 100 people in the year 1980 as ICT-specific IV.

IV estimators are often implemented using two-stage least-squares (2SLS). In the first stage of the 2SLS approach, the endogenous ICT variable is regressed to the chosen instruments and all exogenous variables. Since the instruments are exogenous (a condition we mentioned in subsection 5.3.3), this approximation of the endogenous variables will not correlate with the error term. In the second stage, the regression of interest is estimated as usual, but all endogenous explanatory variables are replaced by the approximate values from the first step (see Greene (2008) for more). In the following we will first use the Generalized Instrumental Variable Estimator (GIVE), which performs an IV regression by applying a 2SLS approach. This estimator is appropriate in our case, because the number of instruments (two, in our case) is higher than the number of endogenous regressors (one, in our case).

The regression results obtained by the application of the GIVE method are presented in table 5.5. At first sight, it is apparent that only minimal changes occur in comparison to the OLS estimator of the previous subsection. The general result pattern with regard to significance and algebraic sign remains unchanged. In comparison to the OLS estimator in table 5.4, it is merely

Table 5.5: Regression Results of the GIVE

	(1)	(2)	(3)	(4)	(5)
dependent variable		log diff. GDP per worker, 1980-2010			
c	2.278	2.112	2.270	2.331	2.159
	(0.001)	(0.014)	(0.001)	(0.001)	(0.016)
$\ln(Y_{1980})$	-0.369	-0.315	-0.374	-0.365	-0.323
	(0.000)	(0.003)	(0.000)	(0.000)	(0.007)
$\ln(n+g+\delta)$	-0.092	0.081	-0.095	-0.073	0.074
	(0.618)	(0.668)	(0.610)	(0.682)	(0.675)
$\ln(I/GDP)$	0.321	0.373	0.319	0.339	0.386
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(HC)$	0.281	0.312	0.284	0.253	0.273
	(0.028)	(0.024)	(0.026)	(0.032)	(0.026)
$\ln(PrivateCredit)$		0.052			0.043
		(0.253)			(0.371)
$\ln(Gov.Consumption)$			-0.022		-0.092
			(0.693)		(0.082)
$\ln(CivilLiberties)$				-0.023	-0.068
				(0.800)	(0.489)
$\ln(ICT)$	0.640	0.523	0.642	0.639	0.501
	(0.002)	(0.027)	(0.002)	(0.013)	(0.096)
N	102	91	102	98	89
$\bar{R}^2$	0.655	0.668	0.652	0.655	0.669

Note: Reported are the regression coefficients and the p-values in parentheses. Explanatory variable Y_{1980} denotes the GDP per worker in 1980.

remarkable that the coefficient values of human capital have decreased and the coefficient values of ICT have increased in all models of the GIVE. The coefficient values of ICT are now between 0.501-0.642, depending on the model. The ICT variable is significant in all of the five estimation models on a 10% level of significance. In addition, the level of the estimated coefficients as well as the $\bar{R}^2$ remain almost unchanged. Overall, the results confirm the significant impact of ICT on GDP per worker growth.

The comparison of OLS estimator and GIVE reveals only minimal difference in the estimation results. This can have two reasons. Firstly, there is possibly no endogeneity due to (potential) reverse causality between ICT and GDP. In this case, the results of the OLS estimator could be used to assess the significance and extent of ICT's impact on per worker growth. Secondly, the suspected endogeneity may also be present in the GIVE, since the two selected instruments are not appropriate for the IV regression. The relevance and validity of the instruments used as well as the presence of endogeneity can be tested. We use the following three tests:

1. the J-Test of overidentifying restrictions (also denoted as Sargan-Hansen test) to test the exogeneity of the used instruments,
2. the calculation of the F-statistics (of the first stage regression) to detect weak instruments,
3. the Hausman test for endogeneity.

Table 5.6: Tests on Instrumental Variables

	(1)	(2)	(3)	(4)	(5)
J-Test	1.427	1.443	1.605	0.856	1.502
	(0.232)	(0.230)	(0.205)	(0.355)	(0.220)
F-statistic (Staiger-Stock)	57.634	36.126	57.373	41.769	27.437
Hausman Test	1.979	0.470	2.154	1.725	0.565
	(0.163)	(0.495)	(0.146)	(0.192)	(0.454)

Note: Reported are the values of the respective test statistics. The p-values of the J-Test and the Hausman Test are in parentheses.

The test results for the five models can be found in table 5.6.

An IV regression model is referred to as overidentified if the number of instruments (two, in our case) is greater than the number of explanatory variables (one, in our case) that are potentially correlated with the disturbance term u. Since our model is overidentified we should test whether the chosen instruments are appropriately independent of the error process, because it allows to evaluate the validity of the instruments. For this purpose, the overidentified equation is estimated with the 2SLS approach and the obtained parameter estimates are used to determine the residuals in the original model. Subsequently, these residuals are regressed to all exogenous variables (model variables and instruments). If the instruments and the exogenous model variables are indeed exogenous, they should not be correlated with the residuals and the second step estimation should provide a low R^2. Under the null hypothesis that all instruments are exogenous, the test statistic nR^2 is asymptotically χ^2-distributed with $(m-k)$ degrees of freedom, where m is the number of instruments, k is the number of endogenous variables and n is the number of country observations. If the test statistic is greater than the corresponding critical value of the χ^2-distribution, the null hypothesis that all instruments are exogenous must be rejected. The rejection of the null hypothesis means that at least one of our instruments is invalid. As shown in table 5.6, the null hypothesis of exogenous instruments can be rejected in none of the five models. Hence, our instruments are valid, which is a condition of the IV approach we pointed out in subsection 5.3.3.

Another condition of our IV approach is the relevance of our instruments, i.e. that the excluded instruments are sufficiently correlated with the included potential endogenous regressor. If there is only a low correlation between the instruments and the endogenous regressors, we call instruments weak. Stock, Wright and Yogo (2002) examined the empirical and theoretical evidence that IV estimations with weak instruments may perform badly and even more poorly than OLS. As a rule of thumb Staiger and Stock (1997) propose that the F-statistic for (joint) significance of the instrument(s) in the first-stage regression should exceed a value of 10. As shown in table 5.6, the F-statistic is above that threshold in all five models. Hence, our instruments are relevant.

After we have tested that the used instruments fulfill the requirements of relevance and validity the question remains whether there is an endogeneity problem at all. To test for endogeneity, we use the Hausman test. Generally, the use of an instrumental variable regression should eliminate the bias and the inconsistency that occurs in the OLS regression due to (potential) endogeneity. Consequently, the OLS and the IV estimations should be different. For that

reason the Hausman test examines if the null hypothesis $H_0 : \hat{\boldsymbol{\beta}}_{OLS} = \hat{\boldsymbol{\beta}}_{GIVE}$ can be rejected. Under the null hypothesis, the OLS estimator is efficient and consistent, the IV estimator is only consistent. Under the alternative hypothesis, the OLS estimator, but not the IV estimator, becomes inconsistent. If the null hypothesis is correct, both estimators are consistent and the difference of the parameters estimated with the two methods is expected to be close to zero. The Hausman test statistic[93] is asymptotically $\chi^2(k)$-distributed under H_0. We reject the H_0 at level α if $H > \chi^2_\alpha(k)$. As shown in table 5.6 the null hypothesis can be rejected in none of the five models. Hence, there is no significant difference between these two estimators.

Since the instruments used are exogenous and relevant, this leads to the question of whether there is endogeneity due to reverse causality between per-worker GDP and ICT. There is no unambiguous answer to this question. Although endogeneity is not proven, it cannot be ruled out on the basis of the tests. As shown in table 5.6, the p-values of the Hausman test of models (1), (3) and (4) are below 0.2. In our opinion, these values are not sufficiently high to make a statement about the impact of ICT on per capita growth on the basis of the OLS estimator. To ensure the robustness of our results, we further use a robust IV estimator.

The GIVE is sensitive to the presence of outliers that can distort the estimated effect of a given regressor on the dependent variable. We use a robust IV estimator, that was initially proposed by Cohen-Freue and Zamar (2006). It uses a scatter S-estimator and replaces the classically estimated covariance matrices of the IV estimator by a variance-covariance matrix that is estimated based on the empirical influence function (see Lopuhaa 1989, Freue et al. 2013). As a result, the estimator is less sensitive to outliers in the used data set. The regression results obtained by the application of the robust IV estimator are presented in table 5.7. In comparison to the GIVE result (shown in table 5.5) the general result pattern with regard to significance and algebraic sign remains almost unchanged. Only the variable that controls for government consumption contributes negatively and significantly to the explanation of the models (3) and (5). Remarkable is the increase of the ICT estimation coefficients in comparison to the GIVE results, which are now between 0.652-0.786. As a result, the reduction of the outlier impact by using robust estimation methods leads to a higher assessment of the impact of ICT on per capita income growth.

The resulting patterns, in terms of significance and algebraic sign of the two additional IV estimators, are almost equal to the results of the OLS and GIVE estimators. As a consequence, they confirm the result that there appears to be no reverse causality between ICT and per capita income. This is remarkable against the background that per capita income was identified as a determinant of ICT in chapter 4 above.

For the quantitative assessment of the effects of ICT, the estimated coefficients of the OLS estimator can be used. These estimation coefficients can be interpreted as elasticities and the results of model (1) in table 5.4 implies that a 10% higher level of ICT infrastructure at the beginning of the period corresponds to 5.15% higher GDP per worker on average in the 1980-2010 period.

[93] For the calculation of the Hausman test statistics, see e.g. Greene (2008).

Table 5.7: Regression Results of the Robust IV Estimator

	(1)	(2)	(3)	(4)	(5)
dependent variable		log diff. GDP per worker, 1980-2010			
c	2.709	2.362	2.812	2.249	2.252
	(0.001)	(0.004)	(0.000)	(0.001)	(0.002)
$\ln(Y_{1980})$	-0.437	-0.383	-0.467	-0.389	-0.394
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(n+g+\delta)$	-0.133	-0.050	-0.115	-0.124	-0.009
	(0.373)	(0.758)	(0.435)	(0.415)	(0.954)
$\ln(I/GDP)$	0.324	0.346	0.314	0.342	0.329
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln(HC)$	0.247	0.262	0.250	0.249	0.242
	(0.005)	(0.008)	(0.003)	(0.005)	(0.007)
$\ln(PrivateCredit)$		0.047			0.027
		(0.261)			(0.495)
$\ln(Gov.Consumption)$			-0.121		-0.168
			(0.006)		(0.001)
$\ln(CivilLiberties)$				0.056	0.060
				(0.331)	(0.288)
$\ln(ICT)$	0.762	0.652	0.786	0.736	0.720
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	102	91	102	98	89

Note: Reported are the regression coefficients and the p-values in parentheses. Explanatory variable Y_{1980} denotes the GDP per worker in 1980.

Despite the substantial differences in the stage of ICT across countries – we found in the previous chapter 4 – it is remarkable that a positive and significant relationship to per-capita growth could be found globally. This indicates that ICT provides an explanation for long-term growth, regardless of the precise start and the pace of the digital revolution in the respective countries.

In addition to the previously conducted investigations, we have also examined the existence of non-linear effects of ICT. By adding a squared ICT infrastructure variable to the model, we have investigated the hypothesis of falling returns to scale. However, the hypothesis could not be confirmed on the basis of our regression results. Furthermore, we have investigated the existence of interaction effects between ICT and human capital. The underlying idea is that skilled and educated workers are necessary to use and program ICT systems, to automate processes and implement services usefully. In this context, the literature also considers the issue of the absorptive capacity of human capital (see, i.e. Niebel 2014), which means "the ability and effort of workers and managers to apply new technology" (Kneller 2005, p. 5). However, a statistically significant interaction effect between ICT and human capital could not be found. Furthermore, we examined interaction effects between ICT and the investments. Steinmueller (2001) states that (next to human capital) an access to equipment is necessary to make productive use of ICT. An interaction term of ICT and I/GDP has been added to the model, but does not show significant statistical explanatory power.

5.6 Summary

In this chapter, we have examined the contribution of ICT to long-term growth of per-worker income for a broad sample of countries at all development stages. Based on the augmented Solow model of Mankiw, Romer and Weil (1992), we include the constructed variable for ICT infrastructure to the growth regression framework. As a key finding of the OLS estimation it can be observed that the ICT variable significantly and positively explains the economic growth in the period 1980-2010. In comparisons to the original MRW model, the added ICT variable leads to a higher proportion of variance explained in all models. We can interpret the regression results that a 10% higher level of ICT infrastructure at the beginning of the period corresponds to 5.15% higher GDP per working-age persons in the 1980-2010 period.

Based on the assumption of potential reverse causality between the GDP an the ICT infrastructure variable we use two IV estimators. The results of these IV estimators confirm the significant and positive contribution of the ICT to economic growth. With respect to a set of different control variables and (IV-) estimators, this result is robust. The application of Hausman tests for the different models show that the suspected endogeneity could not be confirmed. If there is a possible endogeneity due to reverse causality between ICT and GDP, it does not seem to bias the basic pattern of results. Further investigation results are that neither the quadratic nor the interaction effects of ICT, human capital and investments have a significant influence on economic growth within the MRW model.

Overall, this indicates that ICT provides an explanation for long-term growth, regardless of the precise start and the pace of the digital revolution in the respective countries. Moreover, ICT provides a meaningful extension of the MRW model. Firstly, it increases the explanatory power of the model. Secondly, ICT resists the influence of control variables found in the literature to be robust explanatory variables in the context of cross-country growth regressions.

6 The Role of ICT in Productivity Growth

6.1 Motivation

In this chapter we examine the role of ICT in productivity growth. In the previous chapter 5 we have found that the ICT (infrastructure) significantly and positively explains long-term economic growth across countries. Beside the investigation of the global effects, the intention of the current chapter is to analyze the differences between countries with regard to the effect of ICT to productivity. This is motivated by the findings of the literature review in subsection 2.3.3 that previous research finds ICT investment to be associated with significant productivity gains for developed countries but not or to a lesser extend for developing countries (see e.g. Papaioannou and Dimelis 2007, Yousefi 2011). Nonetheless, developing countries have increased investments in ICT infrastructure (see subsection 3.3, figure 3.4). An important research objective is therefore to examine whether developing countries have been able to achieve significant productivity gains through investment in ICT.

Most of the macro data based literature uses growth accounting techniques as an analytical framework. The substantial disadvantages associated with this commonly used empirical approach (such as the assumption of constant returns to scale, perfect factor markets or disregarding changes in factor prices) and ICT capital as appropriate indicator, lead to doubts about the general validity of statements regarding the impact of ICT to productivity (see chapter 2 and 3). Furthermore, the review of the literature shows that cross-country studies which cover a sufficient number of countries at different development stages are less frequent.

In this chapter, we intend to overcome methodological disadvantages of commonly used empirical approaches and provide differentiated assessments of ICT contribution to productivity growth for a broad cross section of more than 120 countries at very different stages of development. The empirical approach followed in this chapter relies on an extension of the non-parametric Malmquist total factor productivity index. It utilizes the Multi-directional Efficiency Analysis (MEA) approach of Asmild et al. (2016a) to enable variable-specific analyses of productivity change across countries, respectively country groups. This allows us to analyze the differences in the input-specific productivity gains of a country's development stages and to make an explicit statement about the contribution of ICT to productivity. In contrast to most of the former studies (e.g. Colecchia and Schreyer 2002, Dewan and Kraemer 2000) we do not use investments in ICT as indicator for ICT. Instead, we use the non-monetary indicator for ICT infrastructure explained previously in section 3.2 and 3.3, generated for each year of our investigation period 2002-2012.

The analysis in this chapter proceeds by providing the current state of research and used methodology in section 6.2. On this basis, we motivate our research strategy. In section 6.3 we introduce the topic of measuring non-parametric as well as input-specific productivity changes and describe the used MEA approach. This is followed by the description of the database in section 6.4. The results of the MEA and ICT-specific productivity change are presented and discussed in section 6.5. Subsequently, we will extend the more descriptive analysis in section

6.6 by statistical regression estimates to explain the productivity differences in ICT between the different development groups. Section 6.7 summarizes this chapter.

6.2 ICT and Productivity in Developing and Developed Countries

A review of the literature shows that cross-country studies covering a sufficient number of countries at different development stages are less frequent (see e.g. Cardona et al. 2013). In theory, there are several reasons why the impact of ICT to productivity and growth might be different in developing rather than in developed countries. Steinmueller (2001, p. 194) states that "ICTs have the potential to support the development strategy of "leapfrogging", i.e. bypassing some of the processes of accumulation of human capabilities and fixed investment in order to narrow the gaps in productivity and output that separate industrialized and developing countries." Similarly, according to Yousefi (2011), in this context leapfrogging means that developing countries have the opportunity to skip several initial phases of ICT development as the way has already been paved by developed countries. Additional productivity gains of ICT could be triggered, as Stiroh (2002a, p. 43) suggests, by "ICT-related spillovers or network effects." Pilat (2005) suggests that ICT is able to lower transaction costs and to speed up the process of knowledge creation.

On the other side, Niebel (2014) suggests that it is also possible that developing countries could benefit less from ICT investments because they might not be well prepared to take advantage of this technology. One of these reasons might be the lack of absorptive capacities such as an appropriate level of human capital. In this context, absorptive capacities mean "the ability and effort of workers and managers to apply new technology" (Kneller 2005). Furthermore, Steinmueller (2001) deems access to equipment and know-how to be necessary for making productive use of ICT. Moreover, complementary technologies and sectors should be available in these countries and linked to ICT. Hence, it is not clear a priori whether the impact of ICT on productivity is larger in developing countries compared to developed countries. The impact seems to depend on a given set of conditions being internal and external to the economies and "vary considerably from one country to another depending on the stage of economic development" (Yousefi 2001, p. 586).

Several studies confirm investments in ICT to be associated significantly with productivity gains for developed countries. Nevertheless, the empirical evidence on the contribution of ICT (investments) to productivity and economic growth for emerging and especially developing countries is rather weak or ambiguous. The empirical literature at the macro level on ICT and productivity that differentiates between the development stage of the countries is quite sparse.

Dewan and Kraemer (2000) estimate a production function relating data on capital investment of ICT an non-ICT investments to GDP in panel data from 36 (22 developed and 14 developing) countries over the period of 1985-1993. They find a significant and positive effect of ICT capital on GDP growth for the developed, but not for the developing countries. Because the diffusion of ICT had just started by that time in developed countries, Niebel (2014) suggests that is was probably too early to see any significant economic effect in developing countries. Dewan and

Kraemer (2000), however, explain this finding in terms of potentially missing complementary factors in developing countries like human capital.

Pohjola (2002), using data on 42 countries for the period 1985-1999, does not find any significant relationship between ICT and GDP growth in the two subgroups of developing and developed countries. Papaioannou and Dimelis (2007) use an adjusted GMM estimator, based on Arellano and Bond (1991), on a sample of 22 developed and 20 developing countries, covering the 1993-2001 period. They find a positive and significant ICT growth effect in both subsamples, whereby they find the impact to be higher for developed countries.

Adopting the growth accounting framework, Yousefi (2011) uses panel-data from a total of 62 countries, covering the period 2000-2006. By adopting the World Bank income groups, he separates the countries into the four groups of low, lower middle, upper middle and high income. Yousefi finds a significant and positive contribution of ICT to economic growth in the groups of high and upper middle income, but not in the lower middle income group countries.

Dedrick et al. (2013) use data on ICT investment for 45 countries in the period 1994-2007 in the framework of a panel analysis. They find upper-income developing countries (most comparable to emerging countries) to have achieved positive and significant productivity gains from ICT investment. They provide evidence for the contribution of ICT to economic growth for both developing and developed countries, whereby the output elasticites of ICT in developed countries are slightly higher.

Niebel (2014) has the most recent and comprehensive data to date. Based on a sample of 59 developing, emerging and developed countries, he analyzes the impact of ICT capital on GDP growth by applying panel data regressions for the period 1995-2010. He finds statistically significant differences of the output elasticity of ICT for none of the three country subsamples.

Summing up the findings from the empirical literature at the macro level on ICT and productivity[94] it can be stated that differences seem to exist in the productivity gains between developed and developing countries. Moreover, some studies found investments in ICT only to be associated with significant productivity gains for developed countries but not for developing countries. Nonetheless, developing countries have increased investments in ICT (infrastructure) in the recent past (Dedrick et al. 2013). Hence, there is a need to research whether ICT investments pay off in greater productivity for developing countries and whether developing countries have been able to achieve significant productivity gains through investment in ICT.

Although there has been a lot of research to examine the contribution of ICT to productivity and economic growth during the last decades, the fundamental questions of the role and impact of ICT on productivity growth are not finally answered. The substantial critique of ICT capital as appropriate indicator as well as of both empirical approaches (growth accounting and estimation of production functions) leads to doubts about the general validity of statements about the impact of ICT to productivity. Furthermore, the review of the literature shows that cross-country studies are less frequent, but important for understanding the general effect of ICT as

[94] The collection of presented papers can be extended by the investigations of Jorgenson and Vu (2005), who analyze specific regional groups. Furthermore, the studies of Becchetti and Adriani (2005) as well as Vu (2011), examine for a broad set of developed and developing countries.

a technology. Because differences in the productivity gains between developing and developed countries have been found, there is a need to analyze the relation of ICT and productivity in the context of a country´s development stages.

This chapter contributes to the literature in two ways. Firstly, we intend to overcome the methodological disadvantages of the commonly used empirical approaches and provide differentiated assessments of ICT contribution to productivity growth across countries. In order to do so, we use an extension of the non-parametric Malmquist total factor productivity index, which utilizes the MEA approach to enable variable-specific analysis of productivity change. In contrast to growth accounting, MEA does not need equilibrium assumptions or information on factor prices. This method enables us to make considerations of input-specific efficiencies between countries and between country groups. In contrast to most of the former studies we do not use the capital measures as indicator for ICT. Instead, we use the ICT Infrastructure variable described in section 3.2 and 3.3 as proxy for ICT. Secondly, we aggregate the results of the variable-specific MEA approach according to the country's level of development. As in Yousefi (2011), we therefore use the World Bank Atlas method to classify the countries by income in the four categories of low, lower middle, upper middle and high income. This allows us to analyze the differences in input-specific productivity gains between development stages and make an explicit assumption about the contribution of ICT to productivity.

6.3 Measuring Non-Parametric and Input-Specific Productivity Change

In this work, we examine the ICT-specific productivity change in different countries. For this purpose, we consider a macroeconomic production model with the three inputs of physical capital, human capital and ICT, producing GDP as output. In general, the productivity of a Decision Making Unit (DMU), a country in our case, is measured as the ratio of its output to its input (Fried et al. 2008). In the case of multiple inputs, aggregation weights are needed to measure productivity as ratio of the output to the aggregate input. In growth accounting, these aggregation weights have to be specified, for example in the form of factor prices. However, the determination of ICT capital is quite problematic, as quality improvements in particular cannot always be adequately reflected (see section 3.2).

The non-parametric Data Envelopment Analysis (DEA), as proposed by Charnes et al. (1978), determines the aggregation weights endogenously by solving linear programs. The DEA describes a body of concepts and methodologies based on linear programming, where a production frontier of best practices (most efficient input-output combinations) is established as a convex envelopment. The DEA combines the estimation of the technology set with efficiency measurement related to this technology.

According to the representations in the literature (as e.g. in Färe et al. 1992), we consider a set of technologies for each period t. This technology set describes all (in t) feasible transformations of three inputs $\boldsymbol{x}^t \in \mathbb{R}^3_+$ into one output $y^t \in \mathbb{R}_+$:

$$T^t = \left\{ \left(\boldsymbol{x}^t, y^t \right) \in \mathbb{R}^4_+ : \boldsymbol{x}^t \geq \mathbf{0} \,\text{can produce}\, y^t \geq 0 \right\}. \tag{31}$$

Following the literature on production economics (Shephard 1970), we assume that this technology set satisfies the axioms of free disposability of inputs and output, no free-lunch and convexity.

The technology can be also represented by an input correspondence, where $L^t(y^t)$ contains all feasible input combinations in t to produce the output y^t:

$$L^t(y^t) = \left\{ \boldsymbol{x}^t \in \mathbb{R}^3_+ : \left(\boldsymbol{x}^t, y^t\right) \in T^t \right\}. \tag{32}$$

The technology set can be constructed in two ways. In most dynamic analyses the technology set for a period t ($t = 1, \ldots, T$) is constructed using only observations of period t. These so-called "contemporaneous" (see Shestalova 2003) technology sets form a frontier of efficient countries. A frontier comprises all input-output combinations that would leave the technology set $L^t(y^t)$ if inputs were be reduced or outputs were be increased by an arbitrary small amount, or both. In contrast to the contemporaneous frontier, the "sequential" (see e.g. Tulkens and Vanden Eeckaut 1995) frontier incorporates the information of previous periods in a reference production set. This is based on the assumption that all preceding technology sets are also feasible in a certain period. Formally, the sequential input set is given by

$$\widetilde{L^t}\left(y^t\right) = \text{convex}\left\{ L^{t_0}\left(y^{t_0}\right) \cup L^{t_0+1}\left(y^{t_0+1}\right) \cup \ldots \cup L^{t-1}\left(y^{t-1}\right) \cup L^t\left(y^t\right) \right\}. \tag{33}$$

Hence, a sequential input set in t is the convex union of all contemporaneous input sets from the period t_0 up to period t given y^t.

As in many applications, the production possibility set is unknown and has therefore to be estimated. In contrast to paramteric approaches, non-parametric approaches do not rely on a specific functional form. Satisfying the axiomatic assumptions of free disposablilty and convexity, the DEA estimate of the sequential input set can be defined as

$$\widehat{\widetilde{L^t}}\left(y^t\right) = \left\{ \boldsymbol{x}^t \in \mathbb{R}^3_+ : \boldsymbol{x}^t \geq \widetilde{\boldsymbol{X}^t}\widetilde{\boldsymbol{\lambda}^t}, y^t \leq \widetilde{\boldsymbol{y}^t}\widetilde{\boldsymbol{\lambda}^t}, \widetilde{\boldsymbol{\lambda}^t} \geq \boldsymbol{0} \right\}, \tag{34}$$

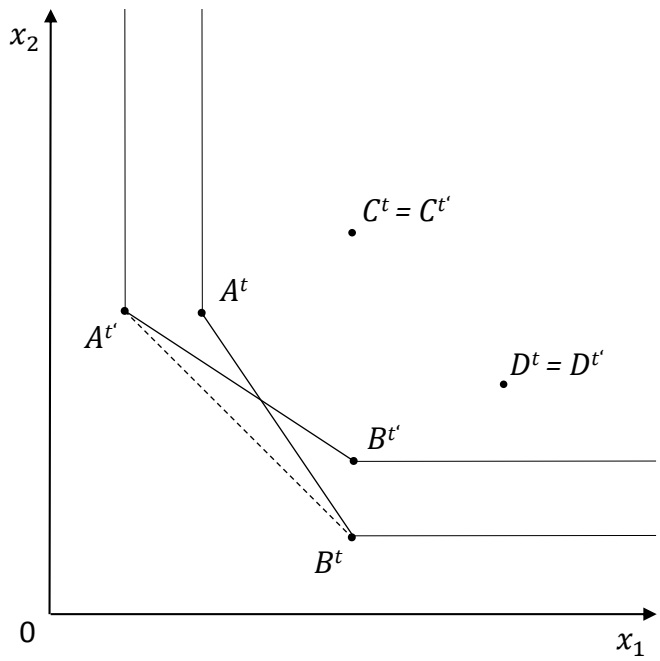
where $\widetilde{\boldsymbol{X}^t}$ represents the $3 \times (n \cdot v)$ matrix of input from n countries in $v = t_0, \ldots, t$ periods, $\widetilde{\boldsymbol{y}^t}$ denotes the $1 \times (n \cdot v)$ vector of output and $\widetilde{\boldsymbol{\lambda}^t}$ denotes the $(n \cdot v) \times 1$ vector of weight factors.[95] The inequalities imply that inefficient input-output combinations which use more input or less output are also part of the technology set (free disposability). It also prevents observations from being classified as best practice which are dominated by other observations.

To capture the idea behind the two different frontier types, consider figure 6.1, which illustrates a situation with two inputs and fixed output for the period t and $t' = t+1$. In this example, the frontier of the contemporaneous input sets is constructed by A^t, B^t and their convex combinations for period t and $A^{t'}$, $B^{t'}$ and their convex combinations for the period t'. The frontier of the sequential input set is given by the $A^{t'}$, B^t and their convex combinations. Following the axiom

[95] Assuming $\boldsymbol{\lambda}_t$ to be positive but otherwise unrestricted implies constant returns to scale. Adding the restriction $\boldsymbol{1}^T\boldsymbol{\lambda}_t = 1$ implies variable returns to scale (Banker et al. 1984).

of free-disposability, the input set is expanded by inefficient input-output combinations. This can be seen graphically from the horizontal and vertical extension of the Frontier, parallel to the x_1 and x_2 axes.

Figure 6.1: Example of Contemporaneous and Sequential Input Sets



The example in figure 6.1 shows that country A needs less of input x_1 in period t' compared to period t to produce the same amount of output. By contrast, country B in period t' needs more of input x_2 (at a constant amount of input x_1) to produce the same amount of output. That B in time t' is not able to produce the output with the same amount of input anymore can be interpreted as technical regress. Technical regress may be confusing and difficult to interpret, especially in a macroeconomic context. However, the phenomenon of declining productivity can be caused by the global economic downturn and external shocks. Shestalova (2003) shows that, for example, a recession is indicated as technical regress. In order to rule out technical regress, we use the sequential boundary of the technology set in the context of this work. This is based on the assumption that all preceding technologies that were feasible in the past are also feasible in the present. In a macroeconomic context, the sequential frontier has also been used by Henderson and Russell (2005).

In the given example, shown in figure 6.1, the countries C and D are located within the input set and hence can be regarded as inefficient because countries exist which produce the same amount of output with less input. The (in)efficiency of these countries is determined by their distance to a benchmark on the technological frontier. A range of measures have been employed to determine the efficiency of a country given a nonparamteric technology (see e.g. Cook and Seiford 2009 for an overview). Among the most frequently applied measures are Farrell efficiency

measures (Farrell 1957). These measures are radial, which means that all inputs are reduced equiproportionally until the frontier is reached.

A static example of efficiency measurement with two inputs is shown graphically in figure 6.2. The equiproportional reduction of the inputs for an inefficient country E is shown as line $\overline{0E}$. In the example, the frontier is reached in point S^{DEA}, which is also the benchmark for E. According to Farrell, the efficiency measure is the distance ratio of $\overline{0S^{DEA}}$ and $\overline{0E}$. This (radial) Farrell input efficiency states the percentage of input that can be reduced to produce the same amount of output. Input-oriented radial measures specify the efficiency of an observation as a one-dimensional value over the entire input set. They do not provide input-specific efficiency measures and are therefore not appropriate in the context of our investigation.

In contrast to radial measures, the approach of the Multi-directional Efficiency Analysis (MEA), proposed by Bogetoft and Hougaard (1999), is based on the (non-radial) directional distance functions (DDF). The underlying idea of DDFs is to determine potential improvements in an input-direction $\boldsymbol{d}_t \in \mathbb{R}^3_+$ and to measure the distance to the frontier in units of $\boldsymbol{d}_t$. This generally results in a different benchmark compared to the radial measurement. We define coordinates of the ideal reference point $\boldsymbol{d}_t^*$ related to a specific production process $\boldsymbol{x}^t \in \widetilde{L^t}\left(y^t\right)$ as

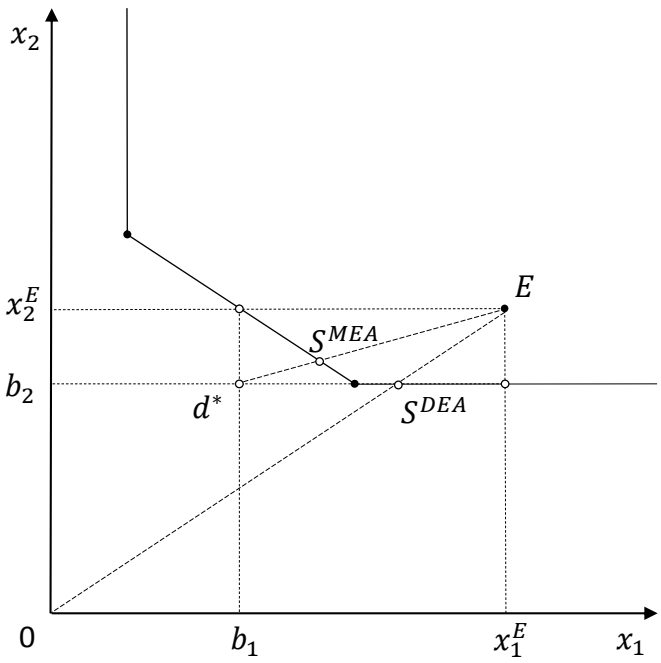
$$d_{i,t}^*\left(\boldsymbol{x}^t\right) = \min\left\{x_i^t \mid \left(x_1^t, \ldots, x_{i-1}^t, x_i^t, x_{i+1}^t, x_3^t\right) \in \widetilde{L^t}\left(y^t\right)\right\}, i = 1, 2, 3. \tag{35}$$

The coordinates of $\boldsymbol{d}_t^*$ related to a given observation are found by minimizing each input dimension separately over $\widetilde{L^t}\left(y^t\right)$, keeping the remaining (two) inputs fixed. The ideal reference point corresponds to the largest possible reduction in each input dimension taken separately. The benchmark selection S^{MEA} is given as the intersection between $\widetilde{L^t}\left(y^t\right)$ and the vector from the ideal point $\boldsymbol{d}_t^*$ to the observation, projecting the observation onto the frontier. In the example in figure 6.2, the coordinates of the ideal reference point $\boldsymbol{d}^*$ related to an observation E is given by the coordinates (b_1, b_2), revealing the reduction potential of each dimension. The selection of the benchmark point S^{MEA} on the frontier is given by the intersection between the sequential frontier and the vector from the ideal point $\boldsymbol{d}^*$ to observation E.

In the graphical example in figure 6.2, the direction to the reference point $\boldsymbol{d}^*$ and hence the benchmark point (S^{MEA}) on the frontier differ from that of the radial measure (S^{DEA}). Both, the ideal reference point as well as the reference direction are found specifically for each country and period. Furthermore, the position of $\boldsymbol{d}_t^*$ depends on the shape of the frontier in t.

Having explained the basic terms and the distance functions in the one-period example, we now proceed to consider the productivity change over time. There are several reasons for productivity changes of countries. On the one hand, countries are experiencing efficiency improvements and are able to produce output with fewer input factors, or to generate more output with the existing input factors. On the other hand, the frontier is driven by technological progress (also in other countries). In both cases, the relative position of countries to the frontier changes over time.

Figure 6.2: Illustration of the Benchmark Selection of MEA and DEA



To analyze productivity changes using distance functions, Caves et al. (1982) have proposed the Malmquist index.[96] Let $e^{t,t'}$ be an efficiency measure in period t against the technology in t', where $t < t'$. Let $e^{t,t'} = 1$, if a country in t is efficient against the technology in t' and $e^{t,t'} < 1$, else. According to Färe et al. (1994b), the Malmquist index is defined as

$$M^{t,t'} = \left[\frac{e^{t',t}}{e^{t,t}} \frac{e^{t',t'}}{e^{t,t'}} \right]^{1/2} . \tag{36}$$

The first ratio compares the efficiency of a certain country in t' against the technology in t with the efficiency of this country in t against the technology in t. If the country has improved its productivity from period t to t', $e^{t',t} > e^{t,t}$ and consequently the ratio $e^{t',t}/e^{t,t} > 1$. If this country has an efficiency measure of 60% in period t and 75% in period t', then its productivity has improved by the factor 1.25. One can interpret this as the country needing 20% $(= 1 - (1/1.25))$ less input to produce the same output or produces 25% more output with the same amount of input. Alternatively we could compare the efficiency of the country in t and t' against the technology in t', as given in the second ratio in equation (36). Since both options are possible, the Malmquist index is expressed as the geometric mean of the two.

As already mentioned, the productivity change of a country can be caused by a move relative to the frontier or a movement of the frontier itself. Färe et al. (1994b) proved the Malmquist index to be decomposed into technical efficiency change (the change of distance of a country to the frontier) and technical change (the frontier shift):

[96] The name refers to earlier works on index numbers by Malmquist (1953).

$$M^{t,t'} = EC^{t,t'} \cdot TC^{t,t'} \tag{37}$$

The efficiency change index (EC) measures the catching-up or falling-behind of a country relative to the present technology. If a country catches-up, $e^{t',t'} > e^{t,t}$ and the ratio is greater than 1:

$$EC^{t,t'} = \frac{e^{t',t'}}{e^{t,t}}. \tag{38}$$

The technical change index (TC) measures the frontier shift as

$$TC^{t,t'} = \left[\frac{e^{t,t}}{e^{t,t'}} \frac{e^{t',t}}{e^{t',t'}} \right]^{1/2}, \tag{39}$$

where the efficiency of a time-fixed production set is measured against changes in the technology. In the case of technical progress the ratio $e^{t,t}/e^{t,t'}$(respectively $e^{t',t}/e^{t',t'}$) is greater than 1. As in the Malmquist index, the TC is the geometric average of the two possible ratios.

The Malmquist index for a specific country greater than 1 indicates a productivity increase, a Malmquist index smaller than 1 indicates a productivity decline. These values can also be assessed quantitatively. A Malmquist index of 1.17 for a specific country indicates a productivity increase of 17% in the period of t to t', an index value of 0.84 indicates a productivity decline of 16%. The interpretation of both efficiency change and technical change is analogous.

For the measurement of input-specific productivity change, different approaches have been developed. Input-specific growth has its origins in measuring sub-vector efficiency. As one of the first, Färe et al. (1994a) estimates the technical efficiency for a subset of inputs rather than for the entire input vector. Input-specific productivity change measures are recent, only a few applications are reported in the literature, i.e. in Oude Lansink and Ondersteijn (2006), Mahlberg and Sahoo (2011), Skevas and Oude Lansink (2014).

In order to combine the input-specific MEA approach with the standard Malmquist index, Asmild et al. (2016a) propose the MEA-Malmquist. In MEA-Malmquist we calculate the input-specific efficiency scores for the observations $\left(\boldsymbol{x}_j^t, y_j^t\right)$ of country j $(j = 1, \ldots, n)$ in period t, benchmarked to the frontier in period t' as

$$\boldsymbol{e}^{t,t'} = \frac{\boldsymbol{x}^t - \boldsymbol{\beta}^*_{t,t'} \left(\left|\boldsymbol{x}^t - \boldsymbol{d}^*_{t,t}\right|\right)}{\boldsymbol{x}^t}, \tag{40}$$

where $\boldsymbol{\beta}^*_{t,t'}$ denotes the directional distance from the observation $\boldsymbol{x}^t$ to the frontier in t' in the direction of $\boldsymbol{d}^*_{t,t}$. The input-specific efficiency scores $\boldsymbol{e}^{t,t'}$ represents a $3 \times n$ matrix. The scores of equation (40) take values between 0 and 1, where a value of 1 for a certain country and certain input indicates that there is no improvement potential on the input in question. An input efficiency score of 0.4, for example, indicates that the country under analysis could reduce the certain input by 60 % in order to be as efficient as the MEA benchmark (selected by equation (42) below).

Similar to (the one-periodic) MEA, Asmild et al. (2016a) describe two linear programs.[97] The first program calculates the ideal reference point $\boldsymbol{d}^*_{t,t'}$ for the input-output combinations of each country in period t against the frontier of period t'. For each country j and each input $i = 1, 2, 3$ in turn, the following program has to be solved:

$$\begin{array}{lrcl} \min\limits_{d_i,\boldsymbol{\lambda}} & d_i & & \\ \text{s.t.} & \widetilde{\boldsymbol{x}_i^{t'}}\boldsymbol{\lambda} & \leq & d_i \\ & \widetilde{\boldsymbol{X}_{-i}^{t'}}\boldsymbol{\lambda} & \leq & \boldsymbol{x}^t_{-i,j} \\ & \widetilde{\boldsymbol{y}^{t'}}\boldsymbol{\lambda} & \geq & y^t_j \\ & \boldsymbol{\lambda} & \geq & \mathbf{0}, \end{array} \tag{41}$$

where $\widetilde{\boldsymbol{y}^{t'}}$ represents the output vector of the countries in period t'. $\widetilde{\boldsymbol{x}_i^{t'}}$ denotes the ith row of $\widetilde{\boldsymbol{X}^{t'}}$ whereas in $\widetilde{\boldsymbol{X}_{-i}^{t'}}$ the ith row of $\widetilde{\boldsymbol{X}^{t'}}$ is excluded. Accordingly, $\boldsymbol{x}^t_{-i,j}$ denotes the input vector of country j without input i.[98] The program finds the coordinates of $\boldsymbol{d}^*_{t,t'}$ related to a given observation $\left(\boldsymbol{x}^t_j, y^t_j\right)$ by minimizing each input dimension i separately over $\widetilde{L^{t'}}\left(y^{t'}\right)$ keeping the remaining inputs fixed. As a result, we obtain a $3 \times n$ matrix $\boldsymbol{d}^*_{t,t'}$ with the ideal reference points for all countries.

The directional distance or excess $\beta^*_{t,t'}$ is found by solving the following program:

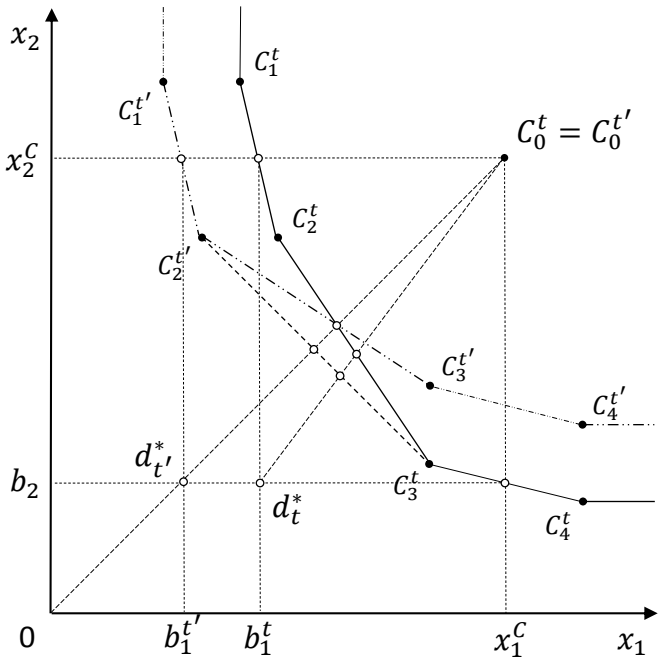
$$\begin{array}{lrcl} \max\limits_{\beta,\boldsymbol{\lambda}} & \beta & & \\ \text{s.t.} & \widetilde{\boldsymbol{X}^{t'}}\boldsymbol{\lambda} & \leq & \boldsymbol{x}^t_j - \beta\left(\left|\boldsymbol{x}^t_j - \boldsymbol{d}^*_{t,t}\right|\right) \\ & \widetilde{\boldsymbol{y}^{t'}}\boldsymbol{\lambda} & \geq & y^t_j \\ & \boldsymbol{\lambda} & \geq & \mathbf{0} \\ & \beta \text{ free of sign.} & & \end{array} \tag{42}$$

In the case where we benchmark observations of period t against the frontier of period t' it is possible that the coordinates of the ideal reference point (in at least one input direction) are greater than the coordinates of the observation point. Thus, we have to consider the absolute distance to the improvement potential by allowing β to be negative. As a result, we obtain the $1 \times n$ vector $\boldsymbol{\beta}^*_{t,t'}$ with the maximum proportion of improvement potentials. The vector takes values between -1 and 1, where a value of 0 implies that there is no improvement potential.

[97] We solve the linear problems with R using the `lpSolveAPI` package, which provides an interface to the R package `lpSolve`, a mixed integer linear programming solver.

[98] Since we use a sequential boundary of the technology set, we have to keep in mind that the length of $\widetilde{\boldsymbol{y}^{t'}}$ and number of columns $\widetilde{\boldsymbol{X}^{t'}}$ depend on the number of preceding periods.

Figure 6.3: Illustration of the MEA-Malmquist



A graphical example of the MEA-Malmquist is given in figure 6.3, where we illustrate a case with a single country C_0 using two inputs to produce a fixed level of output. For the sake of simplicity, the country is observed in two periods with unchanging input use. The production technology associated with the two time periods is represented by the efficient countries C_k^t and $C_k^{t'}(k = 1, 2, 3, 4)$. Since the country C_0 does not change its production between the time periods t and t', any productivity change measure M should result in a value of 1. By using the contemporaneous boundary of the technology set and radial measure, the EC index as well as the TC index would also result in a value of 1, because the benchmark point for C_0 on the frontier is the same in both time periods. By using the sequential boundary of the technology set, the frontier shifts from period t to t'. Using radial measure, the TC index would take a value of > 1 and the EC index would take a value of < 1, such that $EC \times TC = 1$. However, the input-specific examination with the MEA-Malmquist reveals some differences to the one-dimensional index of the DEA-Malmquist. In the input direction of x_1, MEA-Malmquist determines the benchmark points on the frontier as b_1^t in period t and $b_1^{t'}$ in period t'. Thus, the direction to the respective reference points $\boldsymbol{d}_t^*$ and $\boldsymbol{d}_{t'}^*$ is different, as well as their distances to C_0 and their benchmark points on the frontier. Similar to the radial measure, MEA-Malmquist determines in the input direction of x_1 a TC index of > 1, an EC index of < 1 and thus a M index of 1. In the input direction of x_2, however, the benchmark point on the frontier b_2 remains unchanged. The MEA-Malmquist does not determine an increase in technical change and a decrease in efficiency change in the input direction of x_2.[99] Hence, using the MEA-Malmquist in this particular example reveals that

[99] In this particular example, the shift of the reference point from period t to t' leads to a greater directional distance from C_0 to the frontier in period t'. This results in a marginal increase in the efficiency change and a decrease in the technical change for C_0 in the input direction of x_2, which however does not affect the change in productivity (i.e., $EC \times TC = 1$).

the (radially measured) increase in technical change and decrease in efficiency change of C_0 is attributable primarily to the frontier movement in the input direction of x_1.

In this work, we examine the ICT-specific productivity change in different countries over the period 2001-2012. We use the MEA-Malmquist under the use of sequential boundary of the technology set. We assume constant returns to scale. MEA-Malmquist can additionally be adjusted to account for variable returns to scale by adding the convexity constraint $\mathbf{1}^T\boldsymbol{\lambda} = 1$ to both linear programs, defined in equation (41) and equation (42) (Asmild et al. 2016b). Between countries there exist substantial differences in both absolute size of the economy as well as its level of development. For this reason it seems appropriate to allow for economies of scale and assume that the underlying technology is characterized by variable returns to scale. To test this assumption we use the return-to-scale test of Simar and Wilson (2002, 2011). It tests the null hypothesis of constant returns to scale versus the alternative hypothesis of variable returns to scale. For each time period analyzed in this work, the null hypothesis of constant returns to scale cannot be rejected. Thus, we have chosen to assume the underlying technology to have constant returns to scale.

As in the growth-accounting approach we use GDP as output of the production function as well as capital stock and labor as input factors, where the latter is represented by human capital instead of raw labor. We complement the production function by adding ICT as additional input factor. Without doubt it can be stated that ICT is also a part of the capital stock. Authors using the growth-accounting approach deal with this subject by splitting the capital stock or investment into an ICT and a non-ICT capital share. These studies are restricted to samples of developed countries where appropriate data are available. In our case, using a non-monetary measure of ICT, this procedure is not an option. The overlap of ICT and capital stock, however, seems to be a limitation we would accept due to the situation of available data on ICT for a broad cross-country sample.

6.4 Data

In the context of this chapter, we use the indicator for ICT infrastructure explained previously in section 3.2 and 3.3 (second version), generated for each year of our investigation period, 2002-2012.

Besides data on ICT, we require data on GDP, raw labor and physical as well as human capital. These remaining data are taken from the latest release of the Penn World Table (PWT, version 9.0).[100] The database contains information on levels of income and output, covering 182 countries between 1950 and 2014. Since version 8.0 of PWT, the database also provides GDP data constructed from the output side rather than from the expenditure side and, is, therefore, more suitable for productivity analyses.[101]

We use output-side real GDP at chained PPPs (series *rgdpo* in the PWT) as output variable for country j and period t, denoted as Y_{jt}. L_{jt} denotes the raw labor input, which is measured by

[100] A detailed documentation of the database is provided in Feenstra et al. (2013).
[101] A conceptual comparison of output-side and expenditure-side real GDP is given in Feenstra et al. (2009).

the number of workers (number of persons engaged) in the economy (series *emp* in the PWT). We use capital stocks as the physical capital input variable, denoted as K_{jt}. Until the latest version, the PWT unfortunately does not contain direct real capital (rk) measured at chained PPPs. As proposed by Krüger (2016) we fill this data gap by using the series *rkna* instead, which contains the capital stock at constant 2005 national prices. The capital data is real but not in PPPs, so we take the series *rgdpna*, containing the real GDP at constant 2005 national prices and multiply with the output-side real GDP at chained PPPs (series *rgdpo* in PWT). Thus, we compute real output as $rk = (rkna/rgdpna) \cdot rgdpo$.[102]

As already pointed out in section 6.3 the input factor of labor force is mapped by human capital instead of raw labor. As suggested by Hall and Jones (1999), human capital per worker is constructed by the average years of schooling from Barro and Lee (2013) and an assumed rate of return to education, based on the Mincer equation estimate around the world (Psacharopoulos 1994). The latter is represented by returns to education, which is 0.134 up to the fourth year of education, 0.101 from the fourth to the eighth year and 0.068 beyond the eighth year. According to Hall and Jones (1999), the human capital index HK_{jt} is formally measured by

$$HK_{jt} = h_{jt} \cdot L_{jt} = \exp\left(\phi\left(s_{jt}\right)\right) \cdot L_{jt} \; with s_{jt} = \begin{cases} 0.134 \cdot s_{jt} & \text{for } 0 \leq s_{jt} \leq 4 \\ 0.536 + 0.101\left(s_{jt} - 4\right) & \text{for } 4 < s_{jt} \leq 8 \\ 0.94 + 0.068\left(s_{jt} - 8\right) & \text{for } s_{jt} > 8, \end{cases} \quad (43)$$

where s_{jt} denotes the average years of schooling in country j and period t. This measure of human capital is also used by Henderson and Russell (2005). Since the release of PWT version 8.0, these data on human capital per person are available in the series *hc*.[103]

We additionally split the investigation period into three parts to test on differences in the result pattern over the course of time. For every variable of a respective country we form these subperiods by using the median value for three years. By doing this, we mitigate fluctuations of the respective variables caused by economical, meteorological or political fluctuations in the economic data of certain countries and years, respectively. Thus, we receive the periods $t_{1,2}$ from 2001-2003 to 2005-2007, $t_{2,3}$ from 2005-2007 to 2010-2012 and $t_{1,3}$ from 2001-2003 to 2010-2012. In so doing, we are also able to examine a possible effect of the global financial and economic crisis in 2007/08 on economic productivity. The first subperiod $t_{1,2}$ partly lies in the 'productive decade' between 1995-2005, when the ICT revolution started and the impact of computers could initially be seen in productivity statistics. We exclude the relatively small major oil-producing

[102] An alternative way to compute rk would be to use the capital stock at current PPPs (series *ck* in PWT), which is in PPPs but not real. By using the series *cgdpo*, containing output-side GDP at current PPPs, and the output-side real GDP at chained PPPs (*rgdpo*) as $rk = (ck/cgdpo) \cdot rgdpo$. It leads to a real capital stock series which is highly correlated (correlation coefficient in each year of the 2001-2012 period >0.99) to the variant we choose here.

[103] In this series both data from Barro and Lee (2013), Cohen and Leker (2014), as well as Cohen and Soto (2007) are used. The data from Barro and Lee are available only every 5 years, the data from Cohen, Soto and Leker only every 10 years. To obtain data for every year, the authors of PWT interpolate linearly between the observations. The procedure of linear interpolation between the data points is explained on the internet site of the Penn World Table (http://www.rug.nl/ggdc/docs/human_capital_in_pwt_90.pdf).

countries Bahrain, Brunei, Kuwait, Qatar, Saudi Arabia and the United Arab Emirates. We also exclude the countries which are merely large cities such as Hong Kong, Luxembourg, Macao and Singapore. Some descriptive statistics of the variables with regard to the periods can be found in table 6.1.

Table 6.1: Descriptive Statistics on Input and Output

Variables	Min.	1st. Qu.	Median	Mean	3rd. Qu.	Max.	Std dev
t_1: median of the years 2001-2003 (no. of obs = 127)							
GDP	1,649	15,242	45,911	431,169	271,439	133,09,916	1,390,116
Physical Capital	3,436	49,433	172,758	1,462,663	862,794	40,599,675	4,392,308
Human Capital	0.137	3.859	9.045	45.892	28.584	1,635.245	166.881
ICT	0.229	3.668	17.897	35.633	60.937	134.956	39.709
t_2: median of the years 2005-2007 (no. of obs = 127)							
GDP	2,141	19,666	77,229	538,818	347,604	15,083,465	1,659,019
Physical Capital	4,029	60,737	220,775	1,774,745	1,136,239	45,628,858	5,184,673
Human Capital	0.176	4.551	10.660	51.121	32.431	1746.023	180.578
ICT	0.796	21.645	53.297	63.813	110.364	149.534	46.935
t_3: median of the years 2010-2012 (no. of obs = 127)							
GDP	2,436	24,728	107,776	676,806	446,341	15,517,930	1,978,311
Physical Capital	4,689	89,584	280,066	2,349,254	1,566,900	49,279,025	6,747,052
Human Capital	0.261	4.981	11.764	55.543	36.832	1861.596	191.971
ICT	2.723	69.829	109.287	105.323	138.722	189.815	43.076

Reported are the descriptive statistics of complete input and output cases for MEA Malmquist from the respective periods t_1, t_2 and t_3. The summary statistic of t_2 and t_3 only contain data from the respective periods and do not include data from the previous periods (see description of the sequential frontier in section 6.3). GDP = output-side real GDP at chained PPPs (in mil. 2005 US$); Physical Capital = real capital measure at chained PPPs (in mil. 2005 US$); Human Capital = index of human capital according to Hall and Jones (1999); see section 6.4 for description of the ICT variable, gained as the first component of a Principal Component Analysis.

As a result, we obtain a dataset with complete observations for 127 countries. For the analysis, we group the results of the variable-specific MEA approach according to the country's level of development. We use the World Bank Atlas method to classify the countries with a population of more than 30,000 by income in four categories. For the (fiscal) year 2007, low-income economies are defined as those with a gross national income (GNI) per capita of $875 or less; lower middle-income economies are those with a GNI per capita between $876 and $3,465; upper middle-income economies are those with a GNI per capita between $3,466 and $10,725; high-income economies are those with a GNI per capita of $10,725 or more.[104] In our dataset, 34 countries belong to the group of low income countries, 36 countries are part of the group of countries with lower middle income, 25 countries can be assigned to countries with upper middle income and 32 are denoted as high income countries. The group affiliation of the countries is described in the appendix (table B3).

In section 6.6 we will explain the productivity changes in ICT. As explanatory variables we use human capital per worker (from PWT), employment in services (as % of total employment),

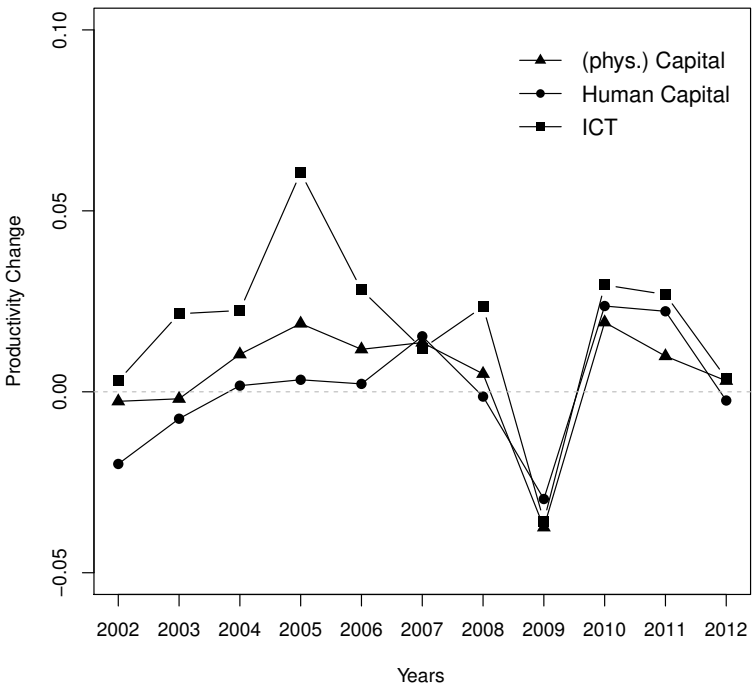
[104] Both current and historical classification by income can be downloaded from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

urban population (as % of total), surface of the countries (all from World Bank[105]) and the KOF Index of Globalization (available at the Swiss Economic Institute KOF of the ETH Zurich[106]). We average all variables over the years from 1995 to 2000 as far as possible and mostly transformed them by taking their natural logarithms.

6.5 MEA Malmquist Results

We now proceed to the analysis of the MEA Malmquist results. First, we examine the annual median productivity changes of the input factors from 2002-2012. We subsequently examine the productivity indices for the respective income groups over the periods $t_{1,2}$, $t_{2,3}$ and $t_{1,3}$, which also allow us to reveal possible effects of the global financial and economic crisis in 2007/08 on the productivity of the subsequent years. Finally, we focus on the productivity change of ICT in the respective income groups, revealing ICT's productivity to be higher in developed countries.

Figure 6.4: Development of the Input-Specific Productivity Indices



Note: Figure shows median productivity change of the MEA Malmquist indices of input factors over the years 2002-2012 for 100 countries with complete data coverage.

The annual median productivity changes of input factors, computed by equation (36), are shown in figure 6.4. The illustrated productivity change results from the respective input-specific Malmquist productivity index minus 1. Besides 2009, ICT shows the highest productivity index values of the input factors in most of the years. In 2009, the effect of the global financial and economic crisis is observable in form of a substantial productivity decline across all inputs. Even

[105] http://databank.worldbank.org/data.
[106] See Dreher (2006) for more information about the KOF Index of Globalization.

beyond 2009, however, the ICT-specific productivity changes are relatively volatile. In 2005, for example, a peak in the ICT curve with an average productivity increase of 7.57% is observable. This value is highly influenced by the productivity development of ICT in certain countries.

As described in section 6.4, we split the entire investigation period into subperiods. We use the median values of the input and output for three years and examine the periods $t_{1,2}$ from (the median of) 2001-2003 to (the median of) 2005-2007, $t_{2,3}$ from 2005-2007 to 2010-2012 and $t_{1,3}$ from 2001-2003 to 2010-2012. The first subperiod covers the 'productive decade' between 1995-2005 and the second subperiod allows the analysis of a possible effect of the global financial and economic crisis in 2007/08 on economic productivity.

Table 6.2: MEA Indicators of the Income Groups

	M			EC			TC		
	K	HK	ICT	K	HK	ICT	K	HK	ICT
Worldwide									
$t_{1,3}$	1.04	1.01	1.14	0.92	0.92	0.87	1.09	1.09	1.28
$t_{1,2}$	1.04	0.99	1.11	0.98	0.95	0.96	1.04	1.04	1.10
$t_{2,3}$	0.99	1.03	1.07	0.95	0.98	0.95	1.04	1.04	1.14
Low Income									
$t_{1,3}$	0.95	0.75	0.71	0.94	0.73	0.71	1.00	1.00	1.01
$t_{1,2}$	0.89	0.80	0.76	0.87	0.77	0.69	1.02	1.03	1.03
$t_{2,3}$	1.01	0.98	1.00	0.99	0.97	0.98	1.00	1.01	1.01
Lower Middle Income									
$t_{1,3}$	1.04	1.00	1.05	0.95	0.89	0.76	1.08	1.13	1.26
$t_{1,2}$	1.02	0.97	1.01	0.98	0.94	0.95	1.02	1.03	1.06
$t_{2,3}$	1.02	1.04	1.07	0.98	0.94	0.90	1.07	1.11	1.20
Upper Middle Income									
$t_{1,3}$	1.04	1.08	1.10	0.94	0.93	0.86	1.10	1.12	1.36
$t_{1,2}$	1.06	1.04	1.12	1.01	0.99	0.98	1.04	1.05	1.11
$t_{2,3}$	0.99	1.05	1.05	0.90	0.95	0.89	1.06	1.08	1.21
High Income									
$t_{1,3}$	1.08	1.04	1.45	0.91	0.96	1.00	1.16	1.09	1.60
$t_{1,2}$	1.07	1.00	1.25	0.97	0.94	0.98	1.08	1.06	1.34
$t_{2,3}$	0.95	1.04	1.15	0.92	1.00	1.02	1.06	1.02	1.17

Note: Reported are the group median MEA values in the respective time periods. M denotes the Malmquist productivity index, EC the efficiency change, TC the technical change of physical Capital (K), human capital (HK) or ICT. $t_{1,2}$ denotes the period from 2001-2003 to 2005-2007, $t_{2,3}$ from 2005-2007 to 2010-2012 and $t_{1,3}$ from 2001-2003 to 2010-2012.

Table 6.2 shows the median values of productivity changes both worldwide and the respective income groups. For each of the three investigation periods, the median of the Malmquist productivity index, as well as its decomposition into efficiency and technical change, are shown in the table for each of the input factors.

Worldwide, a productivity increase over the period $t_{1,3}$ is observable in each of the input factors. The median productivity growth for physical capital lies at 4%, for human capital at 1% and for ICT at 14%. The median of technical change is 9% for both physical and human capital and 28% for ICT. As a result, there is a positive frontier shift in all input directions, whereby

it is most pronounced in the direction of ICT. The median of all input factors indicates an efficiency decrease in the same period, 8% for both physical and human capital and 13% for ICT. The efficiency losses are largely attributable to economic developments in low-income countries. These efficiency losses are more moderate in the two higher income groups. In general, there have been noticeable efficiency losses across all input factors and both subperiods. The reason for this is twofold. Firstly, the indicated decrease in efficiency can be an effect of the implemented sequential frontier, capturing the effects of a technical regress (i.e. a recession) in the efficiency change component. Secondly, it is an indication that the technological frontier is driven by (few) countries with high technical change in the respective input-dimensions whereas other countries were not able to catch up.

Comparing the income groups it can be noted that low-income countries have lost productivity on average across all input factors. Whereas the productivity loss in the median is only 5% for physical capital, it is 25% for human capital and 29% for ICT. Thus, it can be noted that ICT is a rather inhibitory factor for the productivity change of low-income countries. There is only minor technical change in this group over the entire investigation period. The decline in productivity is mainly due to efficiency losses and primarily in $t_{1,2}$. The subperiod $t_{2,3}$ in this income group is characterized by a relatively constant, stagnating productivity.

In relation to the other input factors, the productivity development of ICT has increased over time in lower middle income countries. While the increase in $t_{1,2}$ is only 1% and thus behind the increase for physical capital, productivity increases by 7% in $t_{2,3}$ and thus stronger than for physical (2%) and human capital (4%). Over the entire period, ICT exceeds the other input factors. The increase in productivity can be explained by technical change. The productivity change for ICT is also the highest in countries with upper middle income. In $t_{1,3}$, productivity change for ICT is at 10%, whereas the productivity change for physical (4%) and human capital (8%) lies behind. The productivity increase for ICT, however, is considerably stronger in $t_{1,2}$ (12%) than in $t_{2,3}$ (5%). This can also be observed in the high income countries, where the productivity change for ICT in $t_{1,2}$ at 25% is considerably higher than in $t_{2,3}$ at 15%. However, ICT's productivity growth in this income group is remarkably higher than that for other input factors and thus has a positive impact on the productivity development of economies. In $t_{2,3}$ we can observe a productivity decrease in the direction of physical capital in the two upper income groups and a productivity increase in the direction of human capital.

In the two higher income groups we can observe that ICT's productivity growth has slowed down over time. In $t_{1,2}$, the average ICT-specific productivity change is 12% in the upper middle income countries and 25% in the high income countries. In $t_{2,3}$, the average ICT-specific productivity change is 5% in the upper middle income countries and 15% in the high income countries. In the group of low income countries, the ICT-specific productivity decreases in $t_{1,2}$ but ceases to decrease further in $t_{2,3}$. In the group of lower middle income countries, the average ICT-specific productivity change is 1% in $t_{1,2}$ and 7% in $t_{2,3}$. Hence, the productivity growth in this income group has accelerated over time. The fact that we can observe an increase in ICT-specific productivity growth in the lower middle income group and a slowed ICT-specific productivity growth in the two upper income groups over time, can be explained by two possible reasons. The first explanation is that the slowed ICT-specific productivity growth is due to

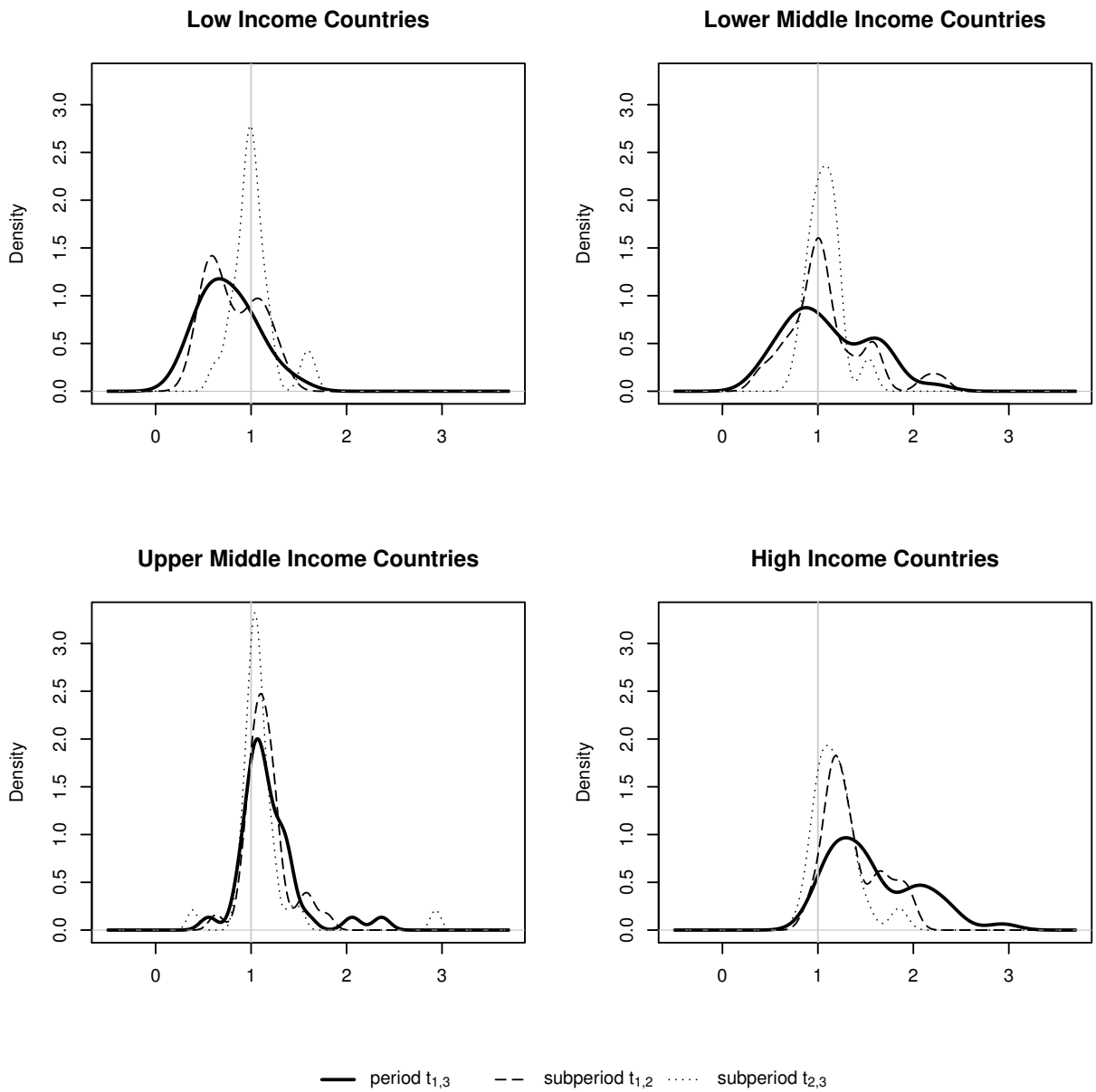
the financial and economic crisis, which had a stronger impact on the economies of developed countries (see e.g. Karshenas 2009). The second explanation is that productivity development on the frontier is slowly weakening and developing countries catch up. The latter is supported by the saturation and catch-up process of the worldwide ICT distribution, shown in section 3.3, but is not evident from the efficiency changes in table 6.2.

Taking all countries together, we find ICT to have the highest productivity index values as well as technical change of all other input factors. Table 6.2 shows that the median of ICT productivity changes increases with the income groups. In order to verify the general validity of the statement that the productivity has risen with increasing development stage, we will consider the distribution of the respective country values below.

Figure 6.5 shows the density plots of the MEA Malmquist productivity index for ICT by income group for the period $t_{1,3}$ as well as the subperiods $t_{1,2}$ and $t_{2,3}$. As indicated in table 6.2, ICT productivity changes rise with higher income groups. It can be seen that the densities shift to the right as the income group increases. For the period $t_{1,3}$ the peak of the curve of low income countries is below the value of 1 and indicates a productivity decrease in ICT for the majority of the countries concerned. For the two middle income groups, the peak is already around 1, for the lower middle income countries slightly below and slightly above for the upper middle income countries. Both distributions of the middle income groups have a positive skew, whereby few countries influence the mean value positively. Mostly apparent is the shift of high income countries. In this group, only a few countries have an ICT productivity decline. In general, it can be seen that the distribution of productivity indices in $t_{2,3}$ is relatively less dispersed across the income groups than in $t_{1,2}$. Figure 6.5 also shows that productivity changes in low and lower middle income countries in the second subperiod $t_{2,3}$ are more positive than in the first subperiod $t_{1,2}$. Conversely, figure 6.5 shows that productivity changes in upper middle and high income countries in the first subperiod $t_{1,2}$ tend to be higher than in the second subperiod $t_{2,3}$. The density plots of the ICT efficiency change values by income group are shown in table D1 in the appendix.

In this section we have shown that the worldwide productivity in the input-direction of ICT has increased by 14% and thus more than the other input factors. In contrast to the first subperiod $t_{1,2}$ (before the global financial and economic crisis and part of the 'productive decade'), productivity growth in ICT has slowed down in $t_{2,3}$, but continues to exceed the productivity growth for the other input factors. We find that developing countries benefit to a lesser extent from the productivity-enhancing effects of ICT, compared to developed countries, which is consistent with parts of the literature (Papaioannou and Dimelis 2007, Dedrick et al. 2013). A comparison is most likely to be possible with Yousefi (2011), due to both a similar investigation period (2002-2006) as well as a country classification by income according to the World Bank Atlas method. Yousefi (2011) also finds ICT to play a major role in the growth of high and upper middle income groups, but not so in the lower middle income group countries.

Figure 6.5: Density Plots of Productivity Change in ICT by Income Group



Note: Figure shows the density plots of ICT MEA Malmquist values (x-axis) according to the respective income groups and investigated periods. A Gaussian kernel is used. The smoothing bandwidth is estimated by the method of Sheather and Jones (1991).

This is remarkable as poorer countries could theoretically have been rapidly catching up in the past, which means that developing countries have the potential to skip several initial phases of ICT development as the way has already been paved by developed countries . As an example, the wired and mostly more expensive fixed-network technology compared to mobile communications technology can be mentioned in this respect. By skipping these technologies, the developing countries would at least have the potential to make greater progress in ICT productivity. Comparing the two subperiods, it is noticeable that ICT productivity growth has increased in the two lower income groups over time, while the productivity growth has slowed down in the two upper income groups.

6.6 Productivity Influencing Factors

In the preceeding section, we showed that developing countries benefit less from the productivity-enhancing effects of ICT than developed countries. This is consistent with most of the literature distinguishing between income groups (Papaioannou and Dimelis 2007, Yousefi 2011, Dedrick et al. 2013). However, it remains unclear why ICT-specific productivity changes depend on the income level of countries. For this reason, we examine explanatory factors for these productivity changes in the current section. In the following, we regress the productivity change on ICT in $t_{1,3}$ on various explanatory variables. The objective of this regression analysis is to find factors that provide explanations for the productivity differences between countries respectively income groups.

Table 6.3: Regression Results for ICT Productivity Change

Dep. Variable				$\ln M$			$\ln EC$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
c	-0.367	-0.536	-1.553	-0.592	-2.468	-3.150	-2.112
	(0.000)	(0.000)	(0.001)	(0.274)	(0.001)	(0.000)	(0.019)
LM	0.370	0.289	0.230	0.278	0.213	0.123	0.019
	(0.001)	(0.014)	(0.158)	(0.027)	(0.047)	(0.391)	(0.905)
UM	0.508	0.360	0.254	0.343	0.305	0.123	-0.140
	(0.000)	(0.002)	(0.148)	(0.018)	(0.006)	(0.459)	(0.478)
H	0.794	0.555	0.460	0.576	0.425	0.172	-0.147
	(0.000)	(0.000)	(0.020)	(0.001)	(0.001)	(0.367)	(0.524)
$\ln HumanCapital$		0.355				0.116	0.378
		(0.028)				(0.534)	(0.039)
$\ln ServiceEmploy$			0.361			0.268	0.106
			(0.012)			(0.068)	(0.546)
$\ln UrbanPop$				0.196			0.143
				(0.173)			(0.550)
$\ln Surface$				-0.033			0.014
				(0.127)			(0.378)
$\ln KOF$					0.571	0.494	0.131
					(0.004)	(0.027)	(0.534)
N	127	127	112	126	124	110	109
$\bar{R}^2$	0.392	0.413	0.436	0.418	0.469	0.501	0.227

Note: Reported are the regression coefficients and the p-values in parentheses. The latter are based on heteroskedasticity-robust standard errors with the correction of MacKinnon and White (1985). Dependent variable is the logarithmic ICT-specific Malmquist productivity index M (columns 1-6) and logarithmic ICT-specific Efficiency Change EC (column 7).

Since the sample of our regressions comprises both developed and developing countries, one would expect heteroskedasticity in the error variance. For this reason, we use the OLS estimator with heteroskedasticity-robust standard errors with the correction of MacKinnon and White (1985). The regression results of the OLS estimator are shown in table 6.3. Additionally, we check the robustness of the results by using the robust regression estimator by Koller and Stahel (2011),

which combines the advantage of a high breakdown point[107] with high estimation efficiency. The regression results of the robust estimator are shown in table 6.4.[108]

Table 6.4: Robust Regression Results for ICT Productivity Change

Dep. Variable				$\ln M$			$\ln EC$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
c	-0.353	-0.550	-1.541	-0.597	-3.012	-3.144	-1.902
	(0.000)	(0.000)	(0.000)	(0.066)	(0.000)	(0.000)	(0.009)
LM	0.381	0.281	0.199	0.285	0.220	0.117	-0.039
	(0.000)	(0.004)	(0.121)	(0.005)	(0.017)	(0.325)	(0.744)
UM	0.491	0.319	0.212	0.324	0.248	0.074	-0.209
	(0.000)	(0.003)	(0.160)	(0.006)	(0.012)	(0.605)	(0.165)
H	0.777	0.499	0.414	0.544	0.356	0.129	-0.224
	(0.000)	(0.000)	(0.015)	(0.000)	(0.003)	(0.461)	(0.211)
$\ln HumanCapital$		0.410				0.172	0.459
		(0.007)				(0.293)	(0.006)
$\ln ServiceEmploy$			0.369			0.219	0.055
			(0.006)			(0.079)	(0.678)
$\ln UrbanPop$				0.222			0.239
				(0.014)			(0.033)
$\ln Surface$				-0.040			0.013
				(0.015)			(0.420)
$\ln KOF$					0.711	0.535	0.035
					(0.000)	(0.001)	(0.828)
N	127	127	112	126	124	110	109
R^2	0.383	0.417	0.417	0.430	0.519	0.493	0.243

Note: Reported are the regression coefficients and the p-values in parentheses. Dependent variable is the logarithmic ICT-specific Malmquist productivity index M (columns 1-6) and logarithmic ICT-specific Efficiency Change EC (column 7). Renaud and Victoria-Feser (2010) explain the R^2 measure used in the case of the KS regressions.

In model (1) the dependent variable is regressed on three dummy variables which take the value unity if the observation is either a low middle (LM), upper middle (UM) or high income country (H) and zero otherwise. The dummy variables indicate the differences to the reference group of low income countries, represented by the intercept c. In the OLS regression, the intercept of -0.367 represents the mean value of the logarithmic Malmquist productivity index in the group of low income countries. Accordingly, the non-logarithmic mean index value in this group is 0.744. The logarithmic mean index value in the group of low middle income countries consequently is $-0.367 + 0.370 = 0.003$, which corresponds to a (non-logarithmic) index mean value of 1.004. The logarithmic mean Malmquist productivity index of the upper middle and high income countries are correspondingly at 0.141 and 0.427, which corresponds to non-logarithmic mean values of 1.151 and 1.533.[109] We find all dummy variables to significantly explain the productivity change of ICT and thus 39.2% of the variance. The coefficients for this model in the

[107] The breakdown point is defined as the smallest fraction of contaminated observations in the sample that can lead to an arbitrarily large deviation of the estimator.

[108] All computations are programmed in R using the following packages: `lmtest` and `sandwich` (for the least squares regression and the heteroskadasticity-robust standard errors), `robust` (for the robust regressions).

[109] These values are different from those in table 6.2, which are based on median values.

robust estimator differ only slightly from those of the OLS. Also with this estimator, all dummy variables significantly explain the productivity change of ICT and hence 38.3% of the variance.

In model (2) we add the level of human capital per worker as explanatory variable.[110] As pointed out in section 6.2, the lacking absorptive capacities as an appropriate level of human capital can be a possible reason why developing countries could benefit less from ICT (see Steinmueller 2001, Kneller 2005, Niebel 2014). The regression results show that the level of human capital explains the ICT-specific productivity change significantly and positively. However, the dummy variables retain their significant explanatory power. Since the natural logarithm is taken from both dependent and explanatory variables (with the exception of the dummy variables), the estimation coefficients can be interpreted as elasticities. In this model, the 0.355 elasticity implies that a 10% higher level of human capital at the beginning of the period corresponds to 3.55% higher productivity index on average in $t_{1,4}$. The robust estimator results in a higher elasticity for this model with 0.410. By adding the human capital variable the value of R^2 rises in both estimators (more than 2 percentage points in the OLS and more than 3 percentage points for the robust estimator).

We add the percentage of employment in services as explanatory variable in model (3) to control for the sectoral composition of the economies. As the underlying idea, the share of service (or manufacturing) sector is supposed to affect investment in ICT (see e.g. Caselli and Coleman 2001, Gust and Marquez 2004) and thus gains more experience in the productive use of ICT. The regression results show that the share of employment in the service sector significantly and positively explains the ICT productivity change. By adding the service employment variable, the significance of the dummy variable for the lower middle and upper middle income countries decreases substantially. The 0.361 elasticity in the OLS and 0.369 elasticity in the robust estimator implies that a 10% higher employment in service at the beginning of the period corresponds to 3.61% or 3.69% higher productivity index on average in $t_{1,4}$. Also in this model the value of R^2 rises, compared to model (1), for the OLS (+4.4 percentage points) as well as the robust estimator (+3.4 percentage points).

In model (4) we add the share of urban population and the surface area of the countries as explanatory variables to control for geographical cross-country differences. The underlying idea is that a higher share of urban population or smaller surface have a higher ICT productivity because of network economies and the experience of ICT-using firms which are mostly located in cities or in their area. The OLS regression results of the OLS estimator show that neither the share of urban population nor surface area explain ICT productivity differences on an appropriate level of significance. The robust regression estimator, however, found both regression coefficients to be significant. As expected, the negative sign of surface indicates that countries with larger area have significantly lower productivity changes with respect to ICT.

In order to control for the degree of globalization, we add the KOF Index of Globalization to model (5). The KOF index measures the economic, social and political dimensions of globalization. The underlying idea is that the knowledge of the productive use of ICT is being expanded internationally through the sale of ICT products, transnational value chains and the exchange of

[110] We use raw labor L_{jt} as denominator for HK_{jt} in order to obtain the level of human capital per worker.

employees. The regression results show that the KOF index significantly and positively explains the ICT productivity change. The estimated elasticities of the OLS and the robust estimator differ considerably. This prevents a statement about the quantitative effect of the globalization level. In both estimators, however, the variable contributes significantly and positively to the explanation of the model and increases the R^2 in contrast to model (1) by 7.7 percentage points (OLS) or 13.1 percentage points (robust estimator).

In model (6) all previously significant variables are used together. In both estimations, we find the percentage of employment in services and the KOF index to explain ICT productivity differences significantly and positively, while the regression coefficient of human capital per employee is not significant. Both estimators show that the dummy variables for the income groups no longer significantly contribute to the explanation of the model. By using all previously significant variables together in a model, the estimation coefficients change considerably. In comparison to models (2), (3) and (5), the estimation coefficients sink.[111]

In model (7) we also regress the component of ICT's efficiency change on the variables. In the OLS estimation, only the level of human capital per person explains the cross-country differences, while the dummy variables of the income groups have no significant explanatory power. The elasticity implies that a 10% higher level of human capital at the beginning of the period corresponds to 3.78% higher efficiency change index on average in $t_{1,4}$. The robust regression estimator also finds the regression coefficient of urban population share to be significant. The estimation coefficient of human capital at 0.459 is higher than that of the OLS estimator. The result is robust in the presence of the control variables for sectoral composition, geographical environment and globalization as well as under application of the robust regression estimator. Hence, we can conclude that human capital substantially explains the change in efficiency and thus the catch-up process.

We refrain from estimating a model for the ICT-specific technical change component. This component captures the frontier shift, which is in our case mainly driven by only a few countries. In a regression of the ICT-specific component the explanatory variables would thus only explain the ICT-specific technical progress in these few countries.

Besides the above-mentioned factors, it is possible that the institutional environment of the countries may also explain the productivity differences. In order to control for institutional environment, we add commonly-used variables to the regression model (1), such as indexes for property rights, civil liberties and rule of law. Neither for the explanation of the change in productivity nor the change in efficiency can we find a significant contribution of these variables.

Our interest in this section is to find explanatory factors for the productivity differences between countries respectively income groups. By using an OLS as well as a robust regression estimator we find the service employment share as representative for the sectoral composition and the KOF Index of Globalization to explain these productivity differences, in addition of the income group dummy variables. In the case of efficiency change, the level of human capital per person mainly

[111] One possible reason for the change in the coefficients may depend on the smaller country sample. Data for the model (6) are available for 110 countries, especially as data on the percentage of employment in services are not available for each of the 127 countries.

explains the cross-country differences. The robust regression estimator additionally reveals the urban population share to have a significant influence on efficiency change.

6.7 Summary

In this chapter, we have analyzed the differences between countries with regard to the effect of ICT to productivity. Therefore, an extension of the non-parametric method of Multi-directional Efficiency Analysis (MEA) has been used to provide input-specific efficiencies across countries respectively country groups. The MEA has been applied on a macroeconomic production function consisting of physical capital, human capital and ICT.

In the investigation period from 2001-2012 the worldwide productivity of ICT increased by 14% and thus by more than in the case of physical (4%) or human capital (1%), largely driven by technical change (frontier shift). We find that developing countries benefit to a lesser extent from the productivity-enhancing effects of ICT in comparison to developed countries, which confirms the statements from the recent literature. Comparing the two subperiods 2001-2003 to 2005-2007 (partly covering the 'productive decade') and 2005-2007 to 2010-2012 we observe an increase in ICT-specific productivity growth in the two lower income groups as well as a decrease in ICT-specific productivity growth in the two upper income groups.

Regression estimates find several factors to significantly explain the ICT productivity change. These are the service employment share as proxy for the sectoral composition and the KOF Index of Globalization. We also find the level of human capital per employee and the urban population share to significantly explain ICT efficiency changes. These factors provide a more appropriate explanation for ICT productivity changes than income levels.

7 Conclusion

In this dissertation we have examined the relationship between ICT, productivity and economic growth. We have conducted three empirical analyses that investigate this relationship globally for a broad range of countries at all stages of development. In chapter 4, we have investigated the determining factors of ICT (infrastructure) diffusion. The long-term contribution of ICT to economic growth has been analyzed in chapter 5. Chapter 6 has addressed the role of ICT in the productivity development. In this concluding chapter, we summarize the main findings of these studies, illustrate connections between them and highlight prospects for further research.

In our analyses, we have used a non-monetary proxy variable for ICT that is available for up to 178 countries at different levels of development for the 2001-2012 period. In chapter 3, we have constructed this ICT proxy variable from a PCA that has merged highly correlated penetration rates of ICT infrastructure to a single variable that comprises most of the information.

Our analysis of this ICT variable has revealed substantial cross-country differences in the stage of ICT infrastructure. High levels of ICT infrastructure can be found in North America and Europe as well as in countries which are merely cities (such as Hong Kong and Luxembourg) and small advanced countries (like Switzerland). By contrast, the South, East and West African countries are characterized by low levels of ICT infrastructure. In general it can be stated that countries with higher income levels also have a higher level of ICT in contrast to countries with lower income levels. This finding is in line with the branch of the literature concerning the 'global digital divide', a perception that describes the difference between developing and developed countries in terms of access to ICT services and technologies.

Since ICT is often suggested to be a determinant of macroeconomic growth what determines ICT infrastructure is of special interest for policy makers and, thus, explains the differences in its diffusion. In chapter 4 we have investigated economic and institutional determinants of ICT infrastructure. For this purpose, we have applied a variable selection method that originates from machine learning research. Based on a wide array of candidate variables, the Lasso method and several of its more advanced variants have selected the relevant variables, which explain ICT infrastructure in the 2002-2012 period. These selected variables have subsequently been used in common least squares regressions as well as in robust and semiparametric regressions to validate the results against the influence of outliers in the data and to uncover nonlinear effects of the explanatory variables, respectively.

The results show that real income per capita, electricity usage, urbanization, indicators of regulatory and institutional aspects as well as regional dummies are major determinants of ICT infrastructure. Jointly, these variables achieve a very high degree of explanatory power. The application of a semiparametric GAM estimator has revealed nonlinear effects for some explanatory variables, i.e. electricity usage. The bulk of the explanatory power, however, stems from the linear effects of the regressors. By splitting the sample period into two subperiods we have also been able to establish conditional convergence of the ICT infrastructure. This may be taken as evidence against the 'global digital divide', since ICT infrastructure converges to a country-specific long-term level as determined by the structural characteristics of that country.

Interestingly, human capital indicators have not been selected, although many of them have been included in the set of candidate variables. Thus, human capital differences across countries seem not to be directly related to differences in ICT infrastructure. At first glance, this seems counter-intuitive but it may be explained by the fact that many end devices are so easy to operate that not much formal education is actually needed for their usage. For the setup of the infrastructure only a few specialists are required, who may also be hired from abroad.

The findings regarding the electricity usage are particularly interesting from the perspective of growth economics. Comin and Hobijn (2004) highlight electricity production as an important prerequisite for the adoption of other technologies. Electricity is therefore a genuine general purpose technology in the sense of Bresnahan and Trajtenberg (1995), characterized by its pervasiveness and its role as a central precondition for other technologies. One sector which is particularly dependent on electricity is the entire ICT sector. These prerequisites for the deployment of ICT should therefore be the subject of further research.

In chapter 5 we have examined the research objective of whether there is a positive and significant relationship between ICT and long-term economic growth across countries. Since the digital revolution has taken place at varying speeds in different countries, we are particularly interested in whether the impact of ICT on economic growth has taken place in the long term. We therefore investigate the effect over a period of 30 years (1980-2010) since its first appearance in the scientific literature. This helps to fill a gap in the literature, providing an investigation which for the first time covers a period of more than 25 years and more than 95 countries at different stages of development.

The analysis is based on a commonly used cross-country linear growth regression model, which we have augmented by including the constructed variable of ICT infrastructure. From the consideration of ICT in the growth regression model, we have expected two insights. Firstly, we are interested in determining whether the ICT variable positively and significantly explains per capita growth during the investigation period. Secondly, we are interested in investigating whether the inclusion of ICT increases the proportion of variance explained in the growth regression model. We expect the latter in the context of ICT to be a (potential) GPT, which affects technological progress.

The results show that ICT (infrastructure) significantly and positively explains the economic growth of the observation period. In comparison to the original growth model, the added ICT variable leads to a higher proportion of explained variance. We find these results to be robust even under the influence of further variables that control for financial, institutional and policy environments. Since we have seen in chapter 4 that real income per capita is one of the major determinants of ICT infrastructure, it is plausible to suspect an endogeneity problem due to reversed causality between GDP per capita and ICT, which can lead to biased and inconsistent estimates in the context of OLS regressions. Based on the assumption of potential endogeneity we have applied two IV estimators. The estimation results reveal a similar pattern of findings as in the OLS regression and confirm the significant and positive contribution of the ICT to economic growth. The Hausman test has revealed that the regression results of the OLS and the

IV estimators do not differ significantly from each other, such that the endogeneity could not be confirmed.

Despite the substantial differences in the stage of ICT across countries, it is remarkable that a positive and significant relationship to per-capita growth could globally be found. This indicates that ICT provides an explanation for long-term growth, regardless of the precise start and the pace of the digital revolution in the respective countries. It is unlikely that the growth-enhancing effect of ICT is homogeneous across all countries. It can be assumed that the extent of the impact of ICT depends on characteristics that vary between the countries. For future research – aimed at the growth effect of ICT in recent years – the growth regression model could be conducted as a panel on a year-by-year basis. In this manner, the issue of unobserved (time-invariant) country heterogeneity could be addressed. Here, the work of Becchetti and Adriani (2005), who have provided a panel estimation based on an ICT-extended MRW model, could provide a valuable point of departure. We refrained from pursuing this possibility in the present analysis since the period of available data on the ICT variable is rather short.

Apart from the investigation of the global effects, we have analyzed the differences between countries with regard to the effect of ICT to productivity in chapter 6. This was motivated by a review of the previous research which has found ICT investment to be associated with significant productivity gains for developed countries but not or to a lesser extent for developing countries. Since developing countries have also increased investments in ICT in the past, we are particularly interested in the research question of whether developing countries have been able to achieve substantial productivity gains through ICT.

In order to overcome the methodological disadvantages of the commonly used empirical methodologies – in particular the growth accounting approach – we have used an extension of the non-parametric MEA approach. This approach provides input-specific analyses of the productivity change as well as its components efficiency change and technical change. We have applied the MEA to a macroeconomic production function which inputs physical capital, human capital and ICT to examine the role and contribution of ICT to the productivity change of more than 120 countries in the 2001-2012 period.

The MEA results show that the productivity of ICT has increased worldwide, whereby the ICT productivity change is even higher than that of physical and human capital and largely driven by technical change. We are particularly interested in discovering whether patterns can be found in the productivity development across different groups of countries. Classifying the countries by their income per capita into four categories, we find that developing countries benefit to a lesser extent from the productivity-enhancing effects of ICT in comparison to developed countries.

In order to explain the differences in the ICT productivity changes between countries, we have regressed these changes to a variety of explanatory variables for the 2001-2012 period. The regression results reveal that the service employment share – as proxy for the sectoral composition – and an index of globalization of the respective countries are able to explain these ICT productivity changes. Furthermore, the ICT efficiency changes can be explained significantly by the level of human capital per employee and the urban population share. These factors provide a more appropriate explanation for ICT productivity changes than income levels. The findings

regarding the significant effect of human capital is interesting in the context of the results from chapter 4. Although human capital differences across countries seem not to be directly related to differences in ICT infrastructure, a certain level of human capital is necessary to use ICT productively.

For future research we would recommend, in keeping with the last chapter, treating the MEA Malmquist results as a panel, but have refrained from doing so because of the short time-series of available ICT data. A suitable approach from our point of view is provided by the work of Du et al. (2018).

Alongside the further development of chapter 6, it would clearly be worth the effort to improve the construction of the dependent variable for ICT infrastructure. Since we have exploited the publicly available data sources to a considerable extent with a view to reaching a broad cross-country sample, this would require making use of information from commercial sources, which is available, e.g., from the International Telecommunication Union (ITU). Moreover, the country coverage or the set of available variables could be increased by trying to apply imputation methods for closing gaps in the available data. The work of Ilin and Raiko (2010) could provide a starting point for the exploration of an appropriate imputation method in the context of a principal components analysis.

Besides providing a contribution to the scientific literature, the findings of this dissertation can also serve (development) policy purposes. The main question of policy makers is whether policy interventions are needed to realize the maximum potential impact of ICT on the economy. The necessity to foster productivity growth gave rise to the launch of several policy initiatives that have focused exclusively on the (further) development of ICT infrastructure.[112] The results of this dissertation, however, have empirically shown that the development opportunities depend less on technical conditions (the existence of ICT infrastructure) than on the ability of its users to handle them productively. The question of a policy intervention consequently depends on the characteristics of the respective country. As a lesson from chapter 6, it should first be examined whether the sectoral composition (e.g. an ICT-using service sector) of the respective economy and its international orientation requires a (further) development of ICT infrastructure. Furthermore, complementary factors, such as an absorptive capacity of human capital, should exist in order to be able to use ICT productively. A policy intervention – regarding ICT infrastructure – should follow the country-specific analysis. Besides the (further) development of ICT infrastructure, this can also include the implementation of economic, technical and regulatory framework conditions. According to the findings of chapter 4 this could be, for example, the provision of electricity supply in the case of developing countries. Another possible form of policy intervention could be the further liberalization of the telecommunications sector with a view to reducing the cost of using the ICT infrastructure.

[112] In the U.S. the stimulus package "American Recovery and Reinvestment Act" from 2009 set $7.2 bn in grants to invest in broadband and wireless internet access (see e.g. Kruger 2011, Hauge and Prieger 2015). In their "Digital Agenda for Europe", the European Commission has declared that at least 50% of European households should subscribing to internet connections above 100 Mbps by 2020 (see European Commission 2010).

References

Abramovitz, M. (1956), Resource and Output Trends in the United States Since 1870, The American Economic Review 46, 5-23.

Acemoglu, D., Johnson, S., Robinson, J.A. (2001), The Colonial Origins of Comparative Development: An Empirical Investigation, American Economic Review 91, 1369-1401.

Acs, Z.J., de Groot, H.L., Nijkamp, P. (2013), The Emergence of the Knowledge Economy: A Regional Perspective, Springer Science & Business Media.

Aghion, P., Howitt, P. (2007), Capital, Innovation, and Growth Accounting, Oxford Review of Economic Policy 23, 79-93.

Aixalá, J., Fabro, G. (2007), A Model of Growth Augmented With Institutions, Economic Affairs 27, 71-74.

Akerlof, G.A. (1970), The Market For "Lemons": Quality Uncertainty and the Market Mechanism, The Quarterly Journal of Economics 84, 488-500.

Arellano, M., Bond, S. (1991), Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, The Review of Economic Studies 58, 277-297.

Asmild, M., Baležentis, T., Hougaard, J.L. (2016a), Multi-Directional Productivity Change: MEA-Malmquist, Journal of Productivity Analysis 46, 109-119.

Asmild, M., Baležentis, T., Hougaard, J.L. (2016b), Multi-Directional Program Efficiency: The Case of Lithuanian Family Farms, Journal of Productivity Analysis 45, 23-33.

Azman-Saini, W.N.W., Baharumshah, A.Z., Law, S.H. (2010), Foreign Direct Investment, Economic Freedom and Economic Growth: International Evidence, Economic Modelling 27, 1079-1089.

Bach, F.R. (2008), Bolasso: Model Consistent Lasso Estimation Through the Bootstrap, Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland.

Baily, M N. (2002), Distinguished Lecture on Economics in Government: The New Economy: Post Mortem or Second Wind?, Journal of Economic Perspectives 16, 3-22.

Baily, M.N., Lawrence, R.Z. (2001), Do We Have a New E-conomy?, American Economic Review 91, 308-312.

Baily, M.N., Gordon, R.J., Nordhaus, W.D., Romer, D. (1998), The Productivity Slowdown, Measurement Issues, and the Explosion of Computer Power, Brookings Papers on Economic Activity 1988, 347.

Bajari, P., Nekipelov, D., Ryan, S.P., Yang, M. (2015a), Demand Estimation with Machine Learning and Model Combination, NBER Working Paper 20955.

Bajari, P., Nekipelov, D., Ryan, S.P., Yang, M. (2015b), Machine Learning Methods for Demand Estimation, American Economic Review: Papers and Proceedings 105, 481-485.

Baliamoune-Lutz, M. (2003), An Analysis of the Determinants and Effects of ICT Diffusion in Developing Countries, Information Technology for Development 10, 151-169.

Banker, R.D., Charnes, A., Cooper, W.W. (1984), Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, Management Science 30, 1078-1092.

Barro, R.J., Lee, J.-W. (2000), International Data on Educational Attainment Updates and Implications, NBER Working Paper 7911.

Barro, R.J., Lee, J.-W. (2013), A New Data Set of Educational Attainment in the World, 1950-2010, Journal of Development Economics 104, 184-198.

Barro, R.J., Sala-i-Martin, X. (1991), Convergence Across States and Regions, Brookings Papers on Economic Activity, 107-182.

Barro, R.J., Sala-i-Martin, X. (1992), Convergence, Journal of Political Economy 100, 223-251.

Barro, R.J., Sala-i-Martin, X. (1995), Economic Growth, McGraw-Hill, Boston, Mass.

Bayo-Moriones, A., Lera-López, F. (2007), A Firm-Level Analysis of Determinants of ICT Adoption in Spain, Technovation 27, 352-366.

Becchetti, L., Adriani, F. (2005), Does the Digital Divide Matter? The Role of Information and Communication Technology in Cross-Country Level and Growth Estimates, Economics of Innovation and New Technology 14, 435-453.

Becchetti, L., Bedoya, D., Paganetto, L. (2003), ICT Investment, Productivity and Efficiency: Evidence at Firm Level Using a Stochastic Frontier Approach, Journal of Productivity Analysis 20, 143-167.

Behrman, J.R., Rosenzweig, M.R. (1994), Caveat Emptor: Cross-Country Data on Education and the Labour Force, Journal of Development Economics 44, 147-171.

Beilock, R., Dimitrova, D.V. (2003), An Exploratory Model of Inter-Country Internet Diffusion, Telecommunications Policy 27, 237-252.

Belloni, A., Chen, D., Chernozhukov, V., Hansen, C. (2012), Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain, Econometrica 80, 2369-2429.

Bertschek, I., Cerquera, D., Klein, G.J. (2013), More Bits – More Bucks? Measuring the Impact of Broadband Internet on Firm Performance, Information Economics and Policy 25, 190-203.

Bogetoft, P., Hougaard, J.L. (1999), Efficiency Evaluations Based on Potential (Non-Proportional) Improvements, Journal of Productivity Analysis 12, 233-247.

Bosworth, B.P., Triplett, J.E. (2007), The Early 21st Century US Productivity Expansion is Still in Services, International Productivity Monitor 14, 3-19.

Bresnahan, T.F., Trajtenberg, M. (1995), General Purpose Technologies: Engines of Growth?, Journal of Econometrics 65, 83-108.

Bresnahan, T.F., Brynjolfsson, E., 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.

Brynjolfsson, E. (1996), The Contribution of Information Technology to Consumer Welfare, Information Systems Research 7, 281-300.

Brynjolfsson, E., Hitt, L.M (1995), Information Technology as a Factor of Production: The Role of Differences Among Firms, Economics of Innovation and New Technology 3, 183-200.

Brynjolfsson, E., Hitt, L.M. (1996), Paradox Lost? Firm-level Evidence on the Returns to Information Systems Spending, Management Science 42, 541-558.

Brynjolfsson, E., Hitt, L.M. (1998), Beyond the Productivity Paradox, Communications of the ACM 41, 49-55.

Brynjolfsson, E., Hitt, L.M. (2000), Beyond Computation: Information Technology, Organizational Transformation and Business Performance, Journal of Economic Perspectives 14, 23-48.

Brynjolfsson, E., Hitt, L.M. (2003), Computing Productivity: Firm-level Evidence, Review of Economics and Statistics 85, 793-808.

Brynjolfsson, E., McAfee, A. (2012), Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy.

Brynjolfsson, E., Hitt, L.M., Yang, S. (2002), Intangible Assets: Computers and Organizational Capital, Brookings Papers on Economic Activity 2002, 137-181.

Cardona, M., Kretschmer, T., Strobel, T. (2013), ICT and Productivity: Conclusions From the Empirical Literature, Information Economics and Policy 25, 109-125.

Caselli, F., Coleman, W.J. (2001), Cross-Country Technology Diffusion: The Case of Computers, American Economic Review 91, 328-335.

Castells, M. (2011), The Rise of the Network Society: The Information Age: Economy, Society, and Culture (Vol. 1), John Wiley & Sons.

Caves, D.W., Christensen, L.R., Diewert, W.E. (1982), The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity, Econometrica: Journal of the Econometric Society, 1393-1414.

Charnes, A., Cooper, W.W., Rhodes, E. (1978), Measuring the Efficiency of Decision Making Units, European Journal of Operational Research 2, 429-444.

Chernozhukov, V., Hansen, C., Spindler, M. (2015), Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments, American Economic Review: Papers and Proceedings 105, 486-490.

Chinn, M.D., Fairlie, R.W. (2007), The Determinants of the Global Digital Divide: A Cross-Country Analysis of Computer and Internet Penetration, Oxford Economic Papers 59, 16-44.

Chinn, M.D., Fairlie, R.W. (2010), ICT Use in the Developing World: A Review of International Economics 18, 153-167.

Cohen, D., Leker, L. (2014), Health and Education: Another Look with the Proper Data, CEPR Discussion Paper No. DP9940.

Cohen, D., Soto, M. (2007), Growth and Human Capital: Good Data, Good Results, Journal of Economic Growth 12, 51-76.

Cohen-Freue, G.V., Zamar, R.H. (2006), A Robust Instrumental Variables Estimator, International Conference on Robust Statistics, Lisbon, Portugal.

Colecchia, A., Schreyer, P. (2002), ICT Investment and Economic Growth In the 1990s: Is the United States a Unique Case?: A Comparative Study of Nine OECD Countries, Review of Economic Dynamics 5, 408-442.

Comin, D., Hobijn, B. (2004), Cross-Country Technology Adoption: Making the Theories Face the Facts, Journal of Monetary Economics 51, 39-83.

Cook, W.D., Seiford, L.M. (2009), Data Envelopment Analysis (DEA) - Thirty Years On, European Journal of Operational Research 192, 1-17.

Corrado, C.A., van Ark, B. (2016), The Internet and Productivity, In: Bauer, J.M., Latzer, M. (Eds.), Handbook on the Economics of the Internet, 120-145.

Corrado, C.A., Lengermann, P., Bartelsman, E.J., Beaulieu, J.J. (2007), Sectoral Productivity in the United States: Recent Developments and the Role of IT, German Economic Review 8, 188-210.

Crenshaw, E.M., Robison, K.K. (2006), Globalization and the Digital Divide: The Roles of Structural Conduciveness and Global Connection in Internet Diffusion, Social Science Quarterly 87, 190-207.

Czernich, N., Falck, O., Kretschmer, T., and Woessmann, L. (2011), Broadband Infrastructure and Economic Growth, Economic Journal 552, 505-532.

Dasgupta, S., Lall, S., Wheeler, D. (2001), Policy Reform, Economic Growth, and the Digital Divide: An Econometric Analysis, Policy Research Working Paper, Vol. 256, World Bank Publications.

Dasgupta, S., Lall, S., Wheeler, D. (2005), Policy Reform, Economic Growth and the Digital Divide, Oxford Development Studies 33, 229-243.

Datta, A., Agarwal, S. (2004), Telecommunications and Economic Growth: A Panel Data Approach, Applied Economics 36, 1649-1654.

Davenport, T. H., Prusak, L. (2000), Working Knowledge: How Organizations Manage What They Know, Boston, MA, Harvard Business School Press.

Daveri, F. (2002), The New Economy in Europe, 1992-2001, Oxford Review of Economic Policy 18, 345-362.

Daveri, F. (2003), Information Technology and Productivity Growth Across Countries and Sectors (No. 227), IGIER, Bocconi University.

David, P., Wright, G. (1999), General Purpose Technologies and Surges in Productivity: Historical Reflections on the Future of the ICT Revolution (No. 31), Economics Group, Nuffield College, University of Oxford.

Davis, G.B., Hamilton, S. (1993), Managing Information: How Information Systems Impact Organizational Strategy, Business One Irwin, Homewood, Ill.

Dedrick, J., Gurbaxani, V., Kraemer, K.L. (2003), Information Technology and Economic Performance: A Critical Review of the Empirical Evidence, ACM Computing Surveys 35, 1-28.

Dedrick, J., Kraemer, K.L., Shih, E. (2013), Information Technology and Productivity in Developed and Developing Countries, Journal of Management Information Systems 30, 97-122.

Dewan, S., Kraemer, K.L. (2000), Information Technology and Productivity: Evidence From Country-Level Data, Management Science 46, 548-562.

Dimelis, S.P., Papaioannou, S.K. (2011), ICT Growth Effects at the Industry Level: A Comparison Between the US and the EU, Information Economics and Policy 23, 37-50.

Doornik, J.A., Hendry, D.F. (2015), Statistical Model Selection with "Big Data", Cogent Economics and Finance 3, 1045216.

Dreher, A. (2006), Does Globalization Affect Growth? Evidence from a New Index of Globalization, Applied Economics 38, 1091-1110.

Du, K., Worthington, A.C., Zelenyuk, V. (2018), Data Envelopment Analysis, Truncated Regression and Double-Bootstrap for Panel Data With Application to Chinese Banking, European Journal of Operational Research 265, 748-764.

Durlauf, S.N., Johnson, P.A., Temple, J.R. (2005), Growth Econometrics, Handbook of Economic Growth, Vol. 1, 555-677.

European Commission (2010), A Digital Agenda for Europe, COM(2010) 245 final/2, Publications Office of the European Union.

Färe, R., Grosskopf, S., Lindgren, B., Roos, P. (1992), Productivity Changes in Swedish Pharmacies 1980-1989: A Non-Parametric Malmquist Approach, Journal of Productivity Analysis 3, 85-101.

Färe, R., Grosskopf, S., Lovell, C.K. (1994a), Production Frontiers, Cambridge University Press.

Färe, R., Grosskopf, S., Lindgren, B., Roos, P. (1994b), Productivity Developments in Swedish Hospitals: A Malmquist Output Approach in Data Envelopment Analysis: Theory, Methodology and Applications (Eds.) A. Charnes, W. Cooper, A. Lewin and L. Seiford (1994), Kluwer Academic Publishers, Dordrecht.

Farrell, M.J. (1957), The Measurement of Productive Efficiency, Journal of the Royal Statistical Society 120, 253-290.

Feenstra, R.C., Heston, A., Timmer, M.P., Deng, H. (2009), Estimating Real Production and Expenditures Across Nations: A Proposal for Improving the Penn World Tables, Review of Economics and Statistics 91, 201-212.

Feenstra, R.C., Inklaar, R., Timmer, M.P. (2013), The Next Generation of the Penn World Table, available for download at www.ggdc.net/pwt.

Fernandez, C., Ley, E., Steel, M. (2001), Model Uncertainty in Cross-Country Growth Regressions, Journal of Applied Econometrics 16, 563-576.

Forman, C., Zeebroeck, N.V. (2012), From Wires to Partners: How the Internet has Fostered R&D Collaborations Within Firms, Management Science 58, 1549-1568.

Freue, G.V.C., Ortiz-Molina, H., Zamar, R.H. (2013), A Natural Robustication of the Ordinary Instrumental Variables Estimator, Biometrics 69, 641-650.

Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (2008), Efficiency and Productivity, In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency and Productivity Growth. , Oxford University Press, 3-91.

Gordon, R.J. (2000), Does the "New Economy" Measure Up to the Great Inventions of the Past?, Journal of Economic Perspectives 14, 49-74.

Gordon, R.J. (2002), Technology and Economic Performance in the American Economy (No. 8771), National Bureau of Economic Research.

Gordon, R.J. (2012), Is US Economic Growth Over? Faltering Innovation Confronts the Six Headwinds (No. 18315), National Bureau of Economic Research.

Gordon, R.J. (2016), The Rise and Fall of American Growth: The US Standard of Living Since the Civil War, Princeton University Press.

Greenana, N., Mairesse, J. (2000), Computers and Productivity in France: Some Evidence, Economics of Innovation and New Technology 9, 275-315.

Greene, W.H. (2008), Econometric Analysis, 6th Edition, Pearson Prentice Hall, Upper Saddle River.

Griliches, Z. (1994), Productivity, R&D, and the Data Constraint, The American Economic Review 84, 1-23.

Guerrieri, P., Luciani, M., Meliciani, V. (2011), The Determinants of Investment in Information and Communication Technologies, Economics of Innovation and New Technology 20, 387-403.

Guillén, M.F., Suárez, S.L. (2001), Developing the Internet: Entrepreneurship and Public Policy in Ireland, Singapore, Argentina, and Spain, Telecommunications Policy 25, 349-371.

Gust, C., Marquez, J. (2004), International Comparisons of Productivity Growth: The Role of Information Technology and Regulatory Practices, Labour Economics 11, 33-58.

Hall, R.E., Jones, C.I. (1999), Why Do Some Countries Produce So Much More Output Per Worker Than Others?, Quarterly Journal of Economics 114, 83-116.

Haller, S., Traistaru-Siedschlag, J. (2007), The Adoption of ICT: Firm-Level Evidence from Irish Manufacturing Industries, Papers WP204, Economic and Social Research Institute (ESRI).

Hanushek, E.A., Woessmann, L. (2012), Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation, Journal of Economic Growth 17, 267-321.

Hardy, A.P. (1980), The Role of the Telephone in Economic Development, Telecommunications Policy 4, 278-286.

Hargittai, E. (1999), Weaving the Western Web: Explaining Differences in Internet Connectivity Among OECD Countries, Telecommunications Policy 23, 701-718.

Hargittai, E. (2003), The Digital Divide and What to Do About It, New Economy Handbook, 821-839.

Hauge, J.A., Prieger, J.E. (2015), Evaluating the Impact of the American Recovery and Reinvestment Act's BTOP on Broadband Adoption, Applied Economics 47, 6553-6579.

Hausman, J.A. (1978), Specification Tests in Econometrics, Econometrica 46, 1251-1271.

Henderson, D.J., Russell, R.R. (2005), Human Capital and Convergence: A Production-Frontier Approach, International Economic Review 46, 1167-1205.

Hoerl, A.E., Kennard, R.W. (1970), Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics 12, 55-67.

Hoover, K.D., Perez, S.J. (2004), Truth and Robustness in Cross-Country Growth Regressions, Oxford Bulletin of Economics and Statistics 66, 765-798.

Ilin, A., Raiko, T. (2010), Practical Approaches to Principal Component Analysis in the Presence of Missing Values, Journal of Machine Learning Research 11, 1957-2000.

Inklaar, R., O'Mahony, M., Timmer, M.P. (2005), ICT and Europe's Productivity Performance: Industry Level Growth Account Comparisons with the United States, Review of Income and Wealth 51, 505-536.

Inklaar, R., Timmer, M.P., Van Ark, B. (2008), Market Services Productivity Across Europe and the US, Economic Policy 23, 140-194.

Jacobsen, K.F.L. (2003), Telecommunications - A Means to Economic Growth in Developing Countries?, Chr. Michelsen Institute.

Jorgenson, D.W. (2001), Information Technology and the US Economy, American Economic Review 91, 1-32.

Jorgenson, D.W. (2005), Accounting for Growth in the Information Age, In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, Vol. 1A, Elsevier B.V, Amsterdam, 743-815.

Jorgenson, D.W. (2009), The Economics of Productivity, Edward Elgar Publishing.

Jorgenson, D.W., Stiroh, K.J. (1995), Computers and Growth, Economics of Innovation and New Technology 3, 295-316.

Jorgenson, D.W., Stiroh, K.J. (1999), Information Technology and Growth, American Economic Review 89, 109-115.

Jorgenson, D.W., Stiroh, K.J. (2000), Raising the Speed Limit: US Economic Growth in the Information Age, Brookings Papers on Economic Activity 2000, 125-210.

Jorgenson, D.W., Timmer, M.P. (2011), Structural Change in Advanced Nations: A New Set of Stylised Facts, The Scandinavian Journal of Economics 113, 1-29.

Jorgenson, D.W., Vu, K.M. (2005), Information Technology and the World Economy, The Scandinavian Journal of Economics 107, 631-650.

Jorgenson, D.W., Vu, K.M. (2016), The ICT Revolution, World Economic Growth, and Policy Issues, Telecommunications Policy 40, 383-397.

Jorgenson, D.W., Ho, M.S., Stiroh, K.J. (2002), Projecting Productivity Growth: Lessons From the US Growth Resurgence (Vol. 20036), Washington, DC.

Jorgenson, D.W., Ho, M.S., Stiroh, K.J. (2005), Information Technology and the American Growth Resurgence, Cambridge MIT Press.

Jorgenson, D.W., Ho, M.S., Stiroh, K.J. (2008), A Retrospective Look at the US Productivity Growth Resurgence, Journal of Economic Perspectives 22, 3-24.

Jovanovic, B., Rousseau, P.L. (2005), General Purpose Technologies, In: Aghion, P., Durlauf, S.N. (Eds.), Handbook of Economic Growth, Vol. 1, Elsevier B.V, Amsterdam, 1181-1224.

Karshenas, M. (2009), The Impact of the Global Financial and Economic Crisis on the Least Developed Countries, United Nations Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States (UN-OHRLLS).

Kiiski, S., Pohjola, M. (2002), Cross-Country Diffusion of the Internet, Information Economics and Policy 14, 297-310.

Kleinberg, J., Ludwig, J., Mullainathan, S., Obermeyer, Z. (2015), Prediction Policy Problems, American Economic Review: Papers and Proceedings 105, 491-495.

Kneller, R. (2005), Frontier Technology, Absorptive Capacity and Distance, Oxford Bulletin of Economics and Statistics 67, 1-23.

Knowles, S., Owen, P.D. (1995), Health Capital and Cross-Country Variation in Income Per Capita in the Mankiw-Romer-Weil Model, Economics Letters 48, 99-106.

Koller, M., Stahel, W.A. (2011), Sharpening Wald-Type Inference in Robust Regression for Small Samples, Computational Statistics and Data Analysis 55, 2504-2515.

Koutroumpis, P. (2009), The Economic Impact of Broadband on Growth: A Simultaneous Approach, Telecommunications Policy 33, 471-485.

Kruger, L.G. (2011), Broadband Infrastructure Programs in the American Recovery and Reinvestment Act, International Journal of Computer Research 18, 173-193.

Krüger, J.J. (2016), Radar Scanning the World Production Frontier, Journal of Productivity Analysis, 46, 1-13.

Krüger, J.J., Rhiel, M. (2016), Determinants of ICT infrastructure: A Cross-Country Statistical Analysis (No. 228), Darmstadt Discussion Papers in Economics.

Lambsdorff, J.G. (2005), The Methodology of the 2005 Corruption Perceptions Index, Transparency International and University of Passau.

Lechman, E. (2015), ICT Diffusion in Developing Countries: Towards a New Concept of Technological Takeoff, Springer.

Leavitt, H.J., Whisler, T.L. (1958), Management in the 1980s, Harvard Business Review 36, 41-48.

Levine, R., Renelt, D. (1992), A Sensitivity Analysis of Cross-Country Growth Regressions, American Economic Review, 942-963.

Ley, E., Steel, M.F. (2009), On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression, Journal of Applied Econometrics 24, 651-674.

Lichtenberg, F.R. (1995), The Output Contributions of Computer Equipment and Personnel: A Firm-level Analysis, Economics of Innovation and New Technology 3, 201-218.

Lopuhaa, H.P. (1989), On the Relation Between S-estimators and M-estimators of Multivariate Location and Covariance, The Annals of Statistics, 1662-1683.

Loveman, G.W. (1994), An Assessment of the Productivity Impact on Information Technologies, In Allen, T.J. and Scott Morton, M.S. (Eds.), Information Technology and the Corporation of the 1990s, New York: Oxford University Press.

Macdonald, S., Anderson, P., Kimbel, D. (2000), Measurement or Management?: Revisiting the Productivity Paradox of Information Technology, Vierteljahrshefte zur Wirtschaftsforschung/ Quarterly Journal of Economic Research 69, 601-617.

MacKinnon, J.G., White, H. (1985), Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties, Journal of Econometrics 29, 305-325.

Madden, G., Savage, S.J. (1998), CEE Telecommunications Investment and Economic Growth, Information Economics and Policy 10, 173-195.

Mahlberg, B., Sahoo, B.K. (2011), Radial and Non-Radial Decompositions of Luenberger Productivity Indicator With an Illustrative Application, International Journal of Production Economics 131, 721-726.

Malmquist, S. (1953), Index Numbers and Indifference Surfaces, Trabajos de Estadstica y de Investigacon Operativa 4, 209-242.

Mankiw, N.G., Romer, D., Weil, D.N. (1992), A Contribution to the Empirics of Economic Growth, Quarterly Journal of Economics 107, 407-437.

Murphy, K.P. (2012), Machine Learning: A Probabilistic Perspective, Cambridge (Mass.): MIT Press.

Niebel, T. (2014), ICT and Economic Growth - Comparing Developing, Emerging and Developed Countries, ZEW Discussion Papers 14.

Nonneman, W., Vanhoudt, P. (1996), A Further Augmentation of the Solow Model and the Empirics of Economic Growth for OECD Countries, The Quarterly Journal of Economics 111, 943-953.

Norris, P. (2001), Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide, Communication, Society and Politics, Cambridge Univ. Press.

OECD (2000), Measuring the ICT Sector, OECD Publishing, Paris.

OECD (2001), Understanding the Digital Divide, OECD Publishing, Paris, OECD Digital Economy Papers, 38.

OECD (2002), Measuring the Information Economy, OECD Publishing, Paris.

OECD (2009), OECD Guide to Measuring the Information Society, OECD Publishing, Paris.

OECD (2011), OECD Guide to Measuring the Information Society 2011, OECD Publishing, Paris.

Oliner, S.D., Sichel, D.E. (1994), Computers and Output Growth Revisited: How Big Is the Puzzle?, Brookings Papers on Economic Activity 25, 273-334.

Oliner, S.D., Sichel, D.E. (2000), The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?, Journal of Economic Perspectives 14, 3-22.

Oliner, S.D., Sichel, D.E. (2002), Information Technology and Productivity: Where Are We Now And Where Are We Going?, Economic Review 87, 15-46.

O'Mahony, M., Vecchi, M. (2005), Quantifying the Impact of ICT Capital on Output Growth: A Heterogeneous Dynamic Panel Approach, Economica 72, 615-633.

Oude Lansink, A., Ondersteijn, C. (2006), Energy Productivity Growth in the Dutch Greenhouse Industry, American Journal of Agricultural Economics 88, 124-132.

Papaioannou, S.K., Dimelis, S.P. (2007), Information Technology as a Factor of Economic Development: Evidence From Developed and Developing Countries, Economics of Innovation and New Technology 16, 179-194.

Pilat, D. (2005), The ICT Productivity Paradox, OECD Economic Studies 2004, 37-65.

Pohjola, M. (2002), The New Economy: Facts, Impacts and Policies, Information Economics and Policy 14, 133-144.

Pohjola, M. (2003), The Adoption and Diffusion of Information and Communication Technology Across Countries: Patterns and Determinants, New Economy Handbook, 77-100.

Psacharopoulos, G. (1994), Returns to Investment in Education: A Global Update, World Development 22, 1325-1343.

Qian, W., Yang, Y. (2013), Model Selection via Standard Error Adjusted Adaptive Lasso, Annals of the Institute of Statistical Mathematics 65, 295-318.

Ram, R. (2007), IQ and Economic Growth: Further Augmentation of Mankiw-Romer-Weil Model, Economics Letters 94, 7-11.

Renaud, O., Victoria-Feser, M.-P. (2010), A Robust Coefficient of Determination for Regression, Journal of Statistical Planning and Inference 140, 1852-1862.

Röller, L.H., Waverman, L. (2001), Telecommunications Infrastructure and Economic Development: A Simultaneous Approach, American Economic Review 91, 909-923.

Romer, P. (1990), Endogenous Technological Change, Journal of Political Economy 98, 71-101.

Ross, I.M. (1998), The Invention of the Transistor, Proceedings of the IEEE 86, 7-28.

Sachs, J.D., Warner, A.M. (1997), Fundamental Sources of Long-Run Growth, American Economic Review 87, 184-188.

Sala-i Martin, X. (1997), I Just Ran Two Million Regressions, American Economic Review 87, 178-183.

Sala-i-Martin, X., Doppelhofer, G., Miller, R.I. (2004), Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach, American Economic Review 94, 813-835.

Schneider, U., Wagner, M. (2011), Catching Growth Determinants with the Adaptive Lasso, German Economic Review 13, 71-85.

Schreyer, P. (2000), The Contribution of Information and Communication Technology to Output Growth: A Study of the G7 Countries (No. 2000/2), OECD Publishing.

Schreyer, P. (2001), Information And Communication Technology And The Measurement Of Volume Output And Final Demand - A Five-Country Study, Economics of Innovation and New Technology 10, 339-376.

Sheather, S.J., Jones, M.C. (1991), A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation, Series B (Methodological) of the Journal of the Royal Statistical Society 53, 683-690.

Shephard, R.W. (1970), Theory of Cost and Production Functions, Princeton: Princeton University Press.

Shestalova, V. (2003), Sequential Malmquist Indices of Productivity Growth: An Application to OECD Industrial Activities, Journal of Productivity Analysis 19, 211-226.

Sichel, D.E. (2001), The Computer Revolution: An Economic Perspective, Brookings Institution Press.

Simar, L., Wilson, P.W. (2002), Non-Parametric Test of Returns to Scale. European Journal of Operational Research 139, 115-132.

Simar, L., Wilson, P.W. (2011), Inference By the m Out of n Bootstrap in Nonparametric Frontier Models, Journal of Productivity Analysis 36, 33-53.

Skevas, T., Oude Lansink, A. (2014), Reducing Pesticide Use and Pesticide Impact by Productivity Growth: The Case of Dutch Arable Farming, Journal of Agricultural Economics 65, 191-211.

Solow, R.M. (1956), A Contribution to the Theory of Economic Growth, The Quarterly Journal of Economics 70, 65-94.

Solow, R.M. (1957), Technical Change and the Aggregate Production Function, The Review of Economics and Statistics, 312-320.

Solow, R.M. (1987), New York Review of Books, July 12.

Staiger, D., Stock, J.H. (1997), Instrumental Variables Regression with Weak Instruments, Econometrica 65, 557-586.

Steinmueller, W.E. (2001), ICTs and the Possibilities for Leapfrogging by Developing Countries, International Labour Review 140, 193-210.

Stiroh, K.J. (2002a), Are ICT Spillovers Driving the New Economy?, Review of Income and Wealth 48, 33-57.

Stiroh, K.J. (2002b), Information Technology and the US Productivity Revival: What do the Industry Data Say?, American Economic Review 92, 1559-1576.

Stock, J.H., Wright, J.H., Yogo, M. (2002), A Survey of Weak Instruments and Weak Identication in Generalized Method of Moments, Journal of Business & Economic Statistics 20, 518-529.

Summers, R., Heston, A. (1988), A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 countries, 1950-1985, Review of Income and Wealth 34, 1-25.

Swedin, E.G., Ferro, D.L. (2005), Computers: The Life Story of a Technology, Greenwood Publishing Group.

Tibshirani, R. (1996), Regression Shrinkage and Selection via the Lasso, Journal of the Royal Statistical Society Series B 58, 267-288.

Tibshirani, R.J. (2013), The Lasso Problem and Uniqueness, Electronic Journal of Statistics 7, 1456-1490.

Timmer, M.P., van Ark, B. (2005), Does Information and Communication Technology Drive EU-US Productivity Growth Differentials?, Oxford Economic Papers 57, 693-716.

Timmer, M.P., Ypma, G., van Ark, B. (2003), IT in the European Union: Driving Productivity Divergence?, Research Memorandum GD-67, Groningen Growth and Development Centre.

Triplett, J.E., Bosworth, B.P. (2004), Productivity in the US Services Sector: New Sources of Economic Growth, Brookings Institution Press.

Triplett, J.E., Bosworth, B.P. (2006), 'Baumol's Disease'Has Been Cured: IT and Multifactor Productivity in US Service Industries, Edward Elgar Publishing.

Tulkens, H., Vanden Eeckaut, P. (1995), Non-Parametric Efficiency, Progress and Regress Measures for Panel Data: Methodological Aspects, European Journal of Operational Research 80, 474-499.

Turban, E., Leidner, D., McLean, E., Wetherbe, J. (2002), Information Technology for Management, 3rd Edition, John Wiley & Sons, New York.

United Nations (2004), Report on the International Telecommunication Union on Information and Communication Technologies Statistics. E/CN.3/2004/16. OECD, Paris. Available from: http://unstats.un.org/unsd/statcom/doc04/2004-16e.pdf. [Date accessed: March 13th , 2018].

Van Ark, B., Inklaar, R. (2005), Catching Up Or Getting Stuck?: Europe's Troubles to Exploit ICT's Productivity Potential, Research Memorandum GD-79, Groningen Growth and Development Centre.

Van Ark, B., O'Mahony, M., Timmer, M.P. (2008), The Productivity Gap between Europe and the United States: Trends and Causes, Journal of Economic Perspectives 22, 25-44.

Van Reenen, J., Bloom, N., Draca, M., Kretschmer, T., Sadun, R., Overman, H., Schankerman, M. (2010), The Economic Impact of ICT, Final Report for the European Commission.

Varian, H.R. (2014), Big Data: New Tricks for Econometrics, Journal of Economic Perspectives 28, 3-28.

Venturini, F. (2009), The Long-Run Impact of ICT, Empirical Economics 37, 497-515.

Vu, K. (2011), ICT as a Source of Economic Growth in the Information Age: Empirical Evidence from the 1996-2005 Period, Telecommunications Policy 35, 357-372.

Webster, F. (2014), Theories of the Information Society, Routledge, 4th edition.

Wolf, S. (2001), Determinants and Impact of ICT Use for African SMEs: Implications for Rural South Africa, In Center for Development Research (ZEF Bonn), Trade and Industrial Policy Strategies (TIPS), Annual Forum at Misty Hills, Mulderdrift.

Wood, S.N. (2001), mgcv: GAMs and Generalized Ridge Regression for R, R News 1, 20-25.

Wood, S.N. (2006), Generalized Additive Models: An Introduction with R, Boca Raton: Chapman & Hall/CRC.

Wunnava, P.V., Leiter, D.B. (2009), Determinants of Intercountry Internet Diffusion Rates, American Journal of Economics and Sociology 68, 413-426.

Yousefi, A. (2011), The Impact of Information and Communication Technology on Economic Growth: Evidence From Developed and Developing Countries, Economics of Innovation and New Technology 20, 581-596.

Zou, H. (2006), The Adaptive Lasso and Its Oracle Properties, Journal of the American Statistical Association 101, 1418-1429.

Zou, H., Hastie, T. (2005), Regularization and Variable Selection via the Elastic Net, Journal of the Royal Statistical Society Series B 67, 301-320.

Appendix A

Figure A1: Development of the Loadings from the First Component of the PCA

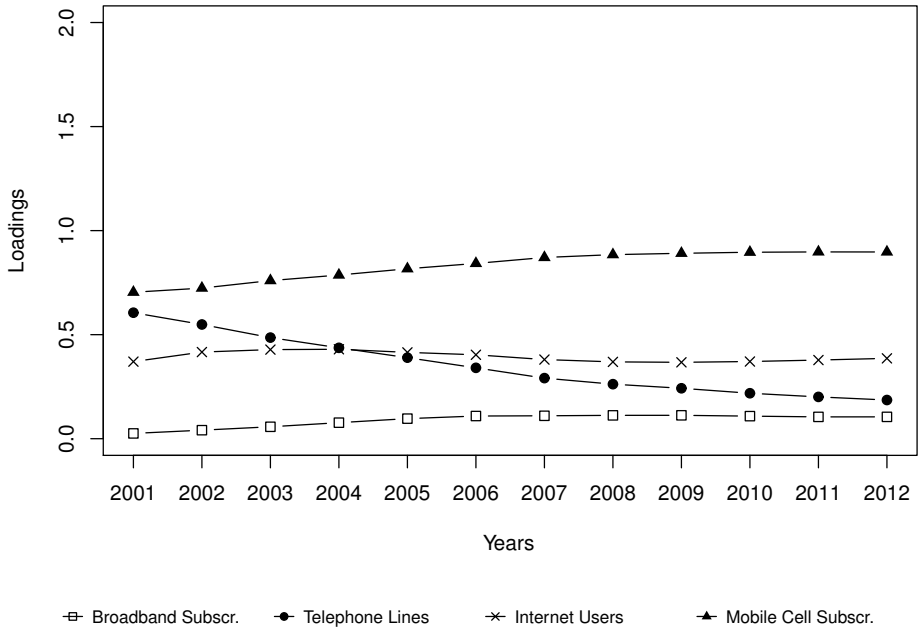


Table A1: Explaining Proportion of the First PCA Component

Year	Proportion	Year	Proportion
2001	0.942	2007	0.948
2002	0.937	2008	0.948
2003	0.950	2009	0.946
2004	0.951	2010	0.946
2005	0.948	2011	0.950
2006	0.947	2012	0.949

Note: Reported are the explaining proportion of the first PCA component to the total variance.

Figure A2: Map of the ICT Infrastructure Variable in Europe

Note: European distribution of the ICT infrastructure variable.
Countries with missing data represented by white color.

Appendix B

Table B1: ICT Sector Definition of the OECD (based on ISIC Rev. 4)

ICT manufacturing industries	
2610	Manufacture of electronic components and boards
2620	Manufacture of computers and peripheral equipment
2630	Manufacture of communication equipment
2640	Manufacture of consumer electronics
2680	Manufacture of magnetic and optical media
ICT trade industries	
4651	Wholesale of computers, computer peripheral equipment and software
4652	Wholesale of electronic and telecommunications equipment and parts
ICT services industries	
5820	Software publishing
6110	Wired telecommunications activities
6120	Wireless telecommunications activities
6130	Satellite telecommunications activities
6190	Other telecommunications activities
6201	Computer programming activities
6202	Computer consultancy and computer facilities management activities
6209	Other information technology and computer service activities
6311	Data processing, hosting and related activities
6312	Web portals
9511	Repair of computers and peripheral equipment
9512	Repair of communication equipment

Note: Reported are the ICT-related industry classes by the ISIC classification code Rev. 4 and name, respectively.

Figure B1: Density Plots of Dependent Variables

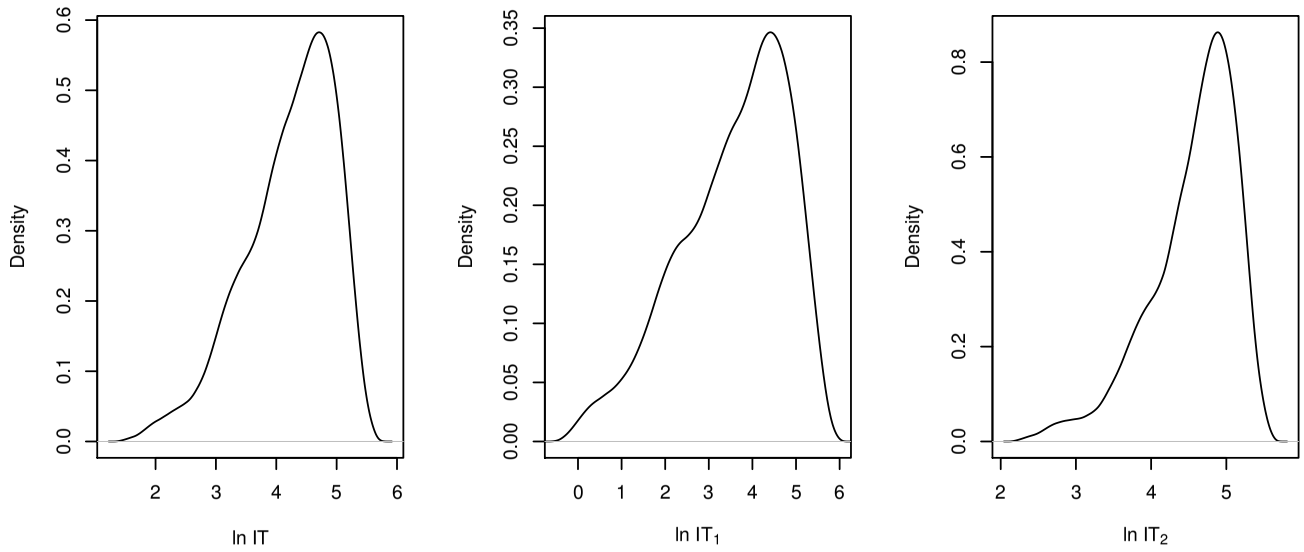


Table B2: Definition and Sources of Explanatory Variables

Variable	Source	Available at	Category	Literature references	Transform.		
					M	G	SD
Dummy variable for advanced countries	Barro-Lee	Barro-Lee	Economic Status and Structure				
Agriculture share in GDP	World Bank national accounts data, and OECD National Accounts data files.	World Bank	..	Caselli and Coleman (2001)	x	x	x
Share of gross capital formation (at current PPPs)	PWT	PWT	..		x	x	x
Developing countries	IMF	IMF	..				
Developed countries	IMF	IMF	..				
Number of persons engaged (in millions)	PWT	PWT	..		x	x	x
Foreign direct investments	International Monetary Fund, International Financial Statistics and Balance of Payments databases, World Bank, International Debt Statistics, and World Bank and OECD GDP estimates	World Bank	..	Crenshaw and Robinson (2006)	x	x	
Private credit by deposit money banks to GDP (%)	International Financial Statistics (IFS), International Monetary Fund (IMF)	World Bank	..		x	x	

Variable	Source	Available at	Category	Literature references	Transform.		
					M	G	SD
Colony Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Gross fixed capital formation (% of GDP and constant 2005 US$)	World Bank national accounts data, and OECD National Accounts data files	World Bank	..		x	x	
Gini coefficient (World Bank estimate)	World Bank, Development Research Group	World Bank	..	Wunnava and Leiter (2009)	x		
Gross national expenditure (% of GDP)	World Bank national accounts data, and OECD National Accounts data files	World Bank	..		x		
Capital stock	PWT	own calculation	..		x	x	x
Economic Globalization	Axel Dreher	QOG	..		x	x	
Index of Globalization	Axel Dreher	QOG	..		x	x	
Commercial banks and other lending (PPG + PNG) (NFL, current US$)	World Bank, International Debt Statistics	World Bank	..		x		
Manufacturing, value added (% of GDP)	World Bank national accounts data, and OECD National Accounts data files	World Bank	..		x	x	
Market Capitalization of listed Companies (% of GDP)	Standard & Poor's, Global Stock Markets Factbook and supplemental S&P data	World Bank	..		x		
Expenditure-side real GDP at chained PPPs (in mil. 2005US$) per person	PWT	PWT	..	most of found literature	x	x	x

Variable	Source	Available at	Category	Literature references	Transform.		
					M	G	SD
Colony Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Output-side real GDP at chained PPPs (in mil. 2005US$) per worker	PWT	PWT	..	most of found literature	x	x	x
employment of service sector (% of total)	International Labour Organization, Key Indicators of the Labour Market database	World Bank	..	Gust and Marquez (2004)	x		
Services, etc., value added (% of GDP)	World Bank national accounts data, and OECD National Accounts data files	World Bank	..		x	x	
Adult literacy rate	UNESCO Institute for Statistics	World Bank	Human Capital	Baliamoune-Lutz (2003)	x		
Percentage of graduates from tertiary education graduating from Engineering, Manufacturing and Construction programmes	UNESCO Institute for Statistics	World Bank	..		x		
Index of human capital per person	PWT	PWT	..		x	x	
Average Years of Tertirary Schooling Attained	Barro-Lee	Barro-Lee	..		x	x	
Average Years of Primary Schooling Attained	Barro-Lee	Barro-Lee	..		x	x	
Researchers in R&D (per million people)	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics	World Bank	..		x		

Variable	Source	Available at	Category	Literature references	Transform.		
					M	G	SD
Colony Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Research and development expenditure (% of GDP)	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics	World Bank	..		x		
Percentage of graduates from tertiary education graduating from Science programmes	UNESCO Institute for Statistics	World Bank	..		x		
Percentage of graduates from tertiary education graduating from Services programmes	UNESCO Institute for Statistics	World Bank	..		x		
Average Years of Secondary Schooling Attained	Barro-Lee	Barro-Lee	..		x	x	
tertiary education enrollment	UNESCO Institute for Statistics	World Bank	..	Crenshaw and Robinson (2006)	x	x	
Average Years of Schooling Attained	Barro-Lee	Barro-Lee	..		x	x	
English Speaking Population	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Fraction Speaking Foreign Language	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Business Freedom	Heritage Foundation	Heritage Foundation	Regulation		x		
Control of Corruption	World Bank (Worldwise Gov. Ind.)	World Bank	..		x		
CPIA business regulatory environment rating (1=low to 6=high)	World Bank Group, CPIA database	World Bank	..		x		
CPIA social protection rating (1=low to 6=high)	World Bank Group, CPIA database	World Bank	..		x		

Variable	Source	Available at	Category	Literature references	Transform.		
					M	G	SD
Colony Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Index of Economic Freedom	Heritage Foundation	Heritage Foundation	..	Baliamoune-Lutz (2003)	x		
Political Rights	Freedom House	QOG	..	Baliamoune-Lutz (2003), Wunnava and Leiter (2009)	x		
Financial Freedom	Heritage Foundation	Heritage Foundation	..		x		
Fiscal Freedom	Heritage Foundation	Heritage Foundation	..		x		
Freedom from Corruption	Heritage Foundation	Heritage Foundation	..		x		
Government Effectiveness	World Bank (Worldwise Gov. Ind.)	World Bank	..		x		
ICRG Indicator of Quality of Government	International Country Risk Guide/The PRS Group	QOG	..		x		
Investment Freedom	Heritage Foundation	Heritage Foundation	..		x		
Monetary Freedom	Heritage Foundation	Heritage Foundation	..		x		
Property Rights	Heritage Foundation	Heritage Foundation	..	Crenshaw and Robinson (2006), Caselli and Coleman (2001)	x		
Political Stability and Absence of Violence/Terrorism	World Bank (Worldwise Gov. Ind.)	World Bank	..		x		

Variable	Source	Available at	Category	Literature references	Transform.		
					M	G	SD
Colony Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Rule of Law	World Bank (Worldwise Gov. Ind.)	World Bank	..		x		
Regulatory Quality	World Bank (Worldwise Gov. Ind.)	World Bank	..	Chinn and Fairlie (2007)	x		
Corruption Perceptions Index	Transparency International	QOG	..		x		
Trade Freedom	Heritage Foundation	Heritage Foundation	..		x		
Voice and Accountability	World Bank (Worldwise Gov. Ind.)	World Bank	..		x		
Control of Corruption - Estimate	The Worldwide Governance Indicators	QOG	..		x		
Political Stability - Estimate	The Worldwide Governance Indicators	QOG	..		x		
Rule of Law - Estimate	The Worldwide Governance Indicators	QOG	..		x		
Civil Liberties	Freedom House	QOG	Demographic Factors	Baliamoune-Lutz (2003), Beilock and Dimitrova 2003), Wunnava and Leiter (2009)	x		
Population (in millions)	PWT	PWT	..		x	x	x
Dummy variable for East Asian and Pacific countries	Barro-Lee	Barro-Lee	Geographical Factors	Dasgupta et al. (2001)			
Countries of the European Union (Dummy)			..				
Dummy variable for countries in Europe and Central Asia	Barro-Lee	Barro-Lee	..				

Variable	Source	Available at	Category	Literature references	Transform.		
					M	G	SD
Colony Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Dummy variable for Latin American and Carribean countries	Barro-Lee	Barro-Lee	..	Dasgupta et al. (2001)			
	Barro-Lee	Barro-Lee	..				
Countries of the OECD (Dummy)			..				
Population of largest city as % of world urban total	United Nations, World Urbanization Prospects.	World Bank	..	Crenshaw and Robinson (2006)	x	x	
Dummy for South Asian countries	Barro-Lee	Barro-Lee	..				
Dummy for Sub-Saharan Africa	Barro-Lee	Barro-Lee	..	Dasgupta et al. (2001)			
Urban population (% of total)	United Nations, World Urbanization Prospects.	World Bank	..	Dasgupta et al. (2001)	x	x	
Air Distance to Big Cities	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Catholics as percentage of population in 1980	La Porta, Lóp ez-de-Silanes, Shleifer and Vishny	QOG	Miscallaneous		x		
Access to electricity (% of population)	World Bank, Sustainable Energy for all (SE4ALL) database from World Bank, Global Electrification database	World Bank	..		x		
Muslims as percentage of population in 1980	La Porta, Lóp ez-de-Silanes, Shleifer and Vishny	QOG	..				
Protestants as percentage of population in 1980	La Porta, Lóp ez-de-Silanes, Shleifer and Vishny	QOG	..		x		
British Colony Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				

Variable	Source	Available at	Category	Literature references	Transform.		
					M	G	SD
Colony Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Socialist Dummy	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Spanish Colony	Sala-i-Martin/Doppelhofer/Miller	SiMDM	..				
Dummy variable for former Spanish colonies	Barro (1999)	SiMDM	..				

Note: The table contains all data collected in the database. The abbreviation SiMDM stands for Sala-i-Martin, Doppelhofer and Miller. In column 5 we describe whether and where the variables are mentioned in the literature. The last three columns indelicate, whether the mean value (M), growth rate (G) and/or standard derivation of the growth rate (sd) has been calculated for a specific variable.

Table B3: List of Countries

Country code	IT data complete	in final dataset	Country code	IT data complete	in final dataset	Country code	IT data complete	in final dataset
ABW			GIN	x		NIC	x	
AFG	x		GMB	x		NLD	x	x
AGO	x		GNB	x		NOC		
ALB	x	x	GNQ	x		NOR	x	x
AND			GRC	x	x	NPL	x	x
ARB			GRD	x		NZL	x	x
ARE	x		GRL			OEC		
ARG	x	x	GTM	x	x	OED		
ARM	x	x	GUM			OMN	x	
ASM			GUY	x		OSS		
ATG	x		HIC			PAK	x	x
AUS	x	x	HKG	x		PAN	x	x
AUT	x	x	HND	x	x	PER	x	x
AZE	x		HPC			PHL	x	x
BDI	x	x	HRV	x	x	PLW		
BEL	x	x	HTI	x		PNG	x	
BEN	x	x	HUN	x	x	POL	x	x
BFA			IDN	x	x	PRI		
BGD	x	x	IMN			PRK		
BGR	x	x	IND	x	x	PRT	x	x
BHR	x	x	INX			PRY	x	x
BHS	x		IRL	x	x	PSE		
BIH	x		IRN	x	x	PSS		

Country code	IT data complete	in final dataset	Country code	IT data complete	in final dataset	Country code	IT data complete	in final dataset
BLR			IRQ	x		PYF		
BLZ	x		ISL	x	x	QAT	x	x
BMU			ISR	x	x	REU		
BOL	x	x	ITA	x	x	ROU	x	
BRA	x	x	JAM	x	x	RUS	x	x
BRB	x		JOR	x	x	RWA	x	x
BRN	x		JPN	x	x	SAS		
BTN	x		KAZ			SAU	x	x
BWA	x	x	KEN	x	x	SDN	x	
CAA			KGZ	x	x	SEN	x	x
CAF	x		KHM	x	x	SER		
CAN	x	x	KIR	x		SGP	x	x
CEA			KNA	x		SLB	x	
CEU			KOR	x	x	SLE		
CHE	x	x	KSV			SLV	x	x
CHI			KWT	x	x	SMR	x	
CHL	x	x	LAC			SOM	x	
CHN	x	x	LAO	x	x	SRB	x	
CIV	x	x	LBN	x		SSA		
CLA			LBR			SSD		
CME			LBY	x		SSF		
CMR	x	x	LCA	x		SST		
COD			LCN			STP	x	
COG	x	x	LDC			SUR	x	
COL	x	x	LIC			SVK	x	x

Country code	IT data complete	in final dataset	Country code	IT data complete	in final dataset	Country code	IT data complete	in final dataset
COM	x		LIE			SVN	x	x
CPV	x		LKA	x	x	SWE	x	x
CRI	x	x	LMC			SWZ	x	x
CSA			LMY			SXM		
CSS			LSO	x	x	SYC	x	
CUB	x		LTU	x	x	SYR	x	x
CUW			LUX	x	x	TCA		
CYM			LVA	x	x	TCD	x	
CYP	x	x	MAC			TGO	x	x
CZE	x	x	MAF			THA	x	x
DEU	x	x	MAR	x	x	TJK	x	
DJI	x		MCO			TKM	x	
DMA	x		MDA	x	x	TLS		
DNK	x	x	MDG	x		TON	x	
DOM	x	x	MDV	x		TTO	x	x
DZA	x		MEA			TUN	x	x
EAP			MEX	x	x	TUR	x	x
EAS			MHL	x		TUV	x	
ECA			MIC			TWN		
ECS			MKD	x		TZA	x	x
ECU	x	x	MLI	x	x	UGA	x	x
EGY	x	x	MLT			UKR	x	x
EMU			MMR	x		UMC		
ERI	x		MNA			URY	x	x
ESP	x	x	MNE			USA	x	x

Country code	IT data complete	in final dataset	Country code	IT data complete	in final dataset	Country code	IT data complete	in final dataset
EST	x	x	MNG	x	x	UZB	x	
ETH	x		MNP			VCT	x	
EUU			MOZ	x	x	VEN	x	x
FIN	x	x	MRT	x	x	VIR		
FJI	x		MUS	x	x	VNM	x	x
FRA	x	x	MWI	x	x	VUT	x	
FRO			MYS	x	x	WLD		
FSM	x		NAC			WSM	x	
GAB	x	x	NAM	x	x	YEM	x	
GBR	x	x	NCL			ZAF	x	x
GEO	x		NER	x	x	ZMB	x	x
GHA	x	x	NGA	x		ZWE	x	x

Note: The column "Country code" lists all available countries by ISO ALPHA-3 code, for which data is available in the database. In column "IT data complete" all countries with data for IT infrastructure are listed. The column "in final dataset" lists all countries, included in the final dataset.

Table B4: Descriptive Statistics

Variable	Description	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Advanced Economies	Dummy variable for advanced countries	0.000	0.000	0.000	0.212	0.000	1.000
Business Freedom	Business Freedom	40.000	55.000	70.000	67.030	73.000	100.000
Business Freedom (log)	Business Freedom	3.689	4.007	4.248	4.186	4.290	4.605
catholic_m	Catholics as percentage of population in 1980	0.000	0.800	18.700	34.600	76.400	96.900
CSH_m	Share of gross capital formation (at current PPPs)	0.051	0.139	0.194	0.199	0.249	0.522
CSH_m_log	Share of gross capital formation (at current PPPs)	-2.968	-1.974	-1.641	-1.700	-1.392	-0.650
CSH_sd	Share of gross capital formation (at current PPPs)	0.030	0.072	0.110	0.140	0.181	0.493
East Asia and the Pacific	Dummy variable for East Asian and Pacific countries	0.000	0.000	0.000	0.097	0.000	1.000
econ_freedom_m	Index of Economic Freedom	35.900	54.100	61.080	60.380	67.280	86.950
econ_freedom_m_log	Index of Economic Freedom	3.581	3.991	4.112	4.086	4.209	4.465
Elec_m	Access to electricity (% of population)	2.000	65.500	93.820	75.650	100.000	100.000
Elec_m_log	Access to electricity (% of population)	0.693	4.182	4.541	4.104	4.605	4.605
EMP_m	Number of persons engaged (in millions)	0.134	1.496	3.418	17.600	10.170	605.900
EMP_m_log	Number of persons engaged (in millions)	-2.007	0.403	1.229	1.325	2.320	6.407

Variable	Description	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
EMP_sd	Number of persons engaged (in millions)	0.005	0.015	0.019	0.021	0.023	0.066
EU	Countries of the European Union (Dummy)	0.000	0.000	0.000	0.195	0.000	1.000
Europe and Central Asia	Dummy variable for countries in Europe and Central Asia	0.000	0.000	0.000	0.142	0.000	1.000
fh_cl_m	Civil Liberties	1.000	2.190	3.810	3.608	4.810	6.952
fh_cl_m_log	Civil Liberties	0.000	0.784	1.338	1.144	1.571	1.939
fh_pr_m	Political Rights	1.000	1.333	3.429	3.400	4.905	6.905
fh_pr_m_log	Political Rights	0.000	0.288	1.232	1.026	1.590	1.932
financ_freedom_m	Financial Freedom	10.000	50.000	50.000	54.220	66.670	90.000
financ_freedom_m_log	Financial Freedom	2.303	3.912	3.912	3.916	4.200	4.500
fiscal_freedom_m	Fiscal Freedom	30.700	54.600	67.450	65.530	76.980	99.900
free_corrupt_m	Freedom from Corruption	10.000	27.830	46.000	45.190	57.170	95.000
free_corrupt_m_log	Freedom from Corruption	2.303	3.326	3.829	3.632	4.046	4.554
gfcf_m	Gross fixed capital formation (% of GDP and constant 2005 US$)	9.629	18.560	20.880	21.600	23.830	51.390
gfcf_m_log	Gross fixed capital formation (% of GDP and constant 2005 US$)	2.265	2.921	3.039	3.040	3.171	3.939
gneGDP_m	Gross national expenditure (% of GDP)	83.060	98.270	101.500	104.300	108.300	212.500
gneGDP_m_log	Gross national expenditure (% of GDP)	4.420	4.588	4.620	4.641	4.685	5.359

Variable	Description	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
HC_m	Index of human capital per person	1.126	1.836	2.212	2.242	2.686	3.429
HC_m_log	Index of human capital per person	0.118	0.607	0.794	0.774	0.988	1.232
hyr_m	Average Years of Tertirary Schooling Attained	0.006	0.082	0.236	0.280	0.400	1.210
hyr_m_log	Average Years of Tertirary Schooling Attained	-5.116	-2.501	-1.444	-1.739	-0.916	0.191
inv_freedom_m	Investment Freedom	10.000	50.000	63.330	58.190	70.000	90.000
inv_freedom_m_log	Investment Freedom	2.303	3.912	4.148	4.010	4.248	4.500
K_m	Capital stock at chained PPPs (in mil. 2005US$)	5,112.000	30,270.000	93,670.000	886,300.000	529,300.000	24,720,000.000
K_m_log	Capital stock at chained PPPs (in mil. 2005US$)	8.539	10.320	11.450	11.820	13.180	17.020
K_sd	Capital stock at chained PPPs (in mil. 2005US$)	0.014	0.028	0.040	0.048	0.057	0.169
Latin America and the Caribbean	Dummy variable for Latin American and Carribean countries	0.000	0.000	0.000	0.168	0.000	1.000
Middle East and North Africa		0.000	0.000	0.000	0.106	0.000	1.000
monet_freedom_m	Monetary Freedom	17.050	63.180	70.350	68.420	81.030	91.000
muslim_m	Muslims as percentage of population in 1980	0.000	0.000	1.000	18.520	16.200	99.400
OECD	Countries of the OECD (Dummy)	0.000	0.000	0.000	0.301	1.000	1.000
POP_m	Population (in millions)	0.252	3.551	8.680	40.850	24.780	1,101.000

Variable	Description	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
POP_m_log	Population (in millions)	-1.379	1.267	2.161	2.273	3.210	7.004
POP_sd	Population (in millions)	0.001	0.005	0.010	0.010	0.014	0.039
prop_rights_m	Property Rights	10.000	50.000	53.330	58.400	70.000	90.000
prop_rights_m_log	Property Rights	2.303	3.912	3.977	3.981	4.248	4.500
protestant_m	Property Rights	0.000	0.200	1.900	12.880	16.100	97.800
pyr_g	Average Years of Primary Schooling Attained	-0.356	0.060	0.229	0.254	0.423	1.093
pyr_m	Average Years of Primary Schooling Attained	0.734	3.230	4.304	4.222	5.268	8.542
pyr_m_log	Average Years of Primary Schooling Attained	-0.309	1.172	1.460	1.353	1.662	2.145
RGDPP_m	Expenditure-side real GDP at chained PPPs (in mil. 2005US$) per person	445.900	2,140.000	6,035.000	9,286.000	15,400.000	36,710.000
RGDPP_m_log	Expenditure-side real GDP at chained PPPs (in mil. 2005US$) per person	6.100	7.669	8.705	8.585	9.642	10.510
RGDPP_sd	Expenditure-side real GDP at chained PPPs (in mil. 2005US$) per person	0.016	0.033	0.046	0.053	0.061	0.256
RGDPW_m	Output-side real GDP at chained PPPs (in mil. 2005US$) per worker	913.000	6,337.000	16,730.000	22,140.000	35,910.000	78,080.000
RGDPW_m_log	Output-side real GDP at chained PPPs (in mil. 2005US$) per worker	6.817	8.754	9.725	9.530	10.490	11.270

Variable	Description	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
RGDPW_sd	Output-side real GDP at chained PPPs (in mil. 2005US$) per worker	0.013	0.031	0.048	0.056	0.065	0.283
South Asia	Dummy for South Asian countries	0.000	0.000	0.000	0.044	0.000	1.000
Sub-Saharan Africa	Dummy for Sub-Saharan Africa	0.000	0.000	0.000	0.230	0.000	1.000
syr_g	Average Years of Secondary Schooling Attained	-0.086	0.308	0.562	0.559	0.793	2.144
syr_m	Average Years of Secondary Schooling Attained	0.086	1.100	2.016	2.136	3.068	5.198
syr_m_log	Average Years of Secondary Schooling Attained	-2.453	0.095	0.701	0.519	1.121	1.648
trade_freedom_m	Trade Freedom	14.000	55.400	65.070	62.340	76.000	83.000
tyr_g	Average Years of Schooling Attained	-0.260	0.205	0.321	0.355	0.492	1.099
tyr_m	Average Years of Schooling Attained	0.906	4.758	6.600	6.638	8.600	12.310
tyr_m_log	Average Years of Schooling Attained	-0.099	1.560	1.887	1.783	2.152	2.510
UrbanPop_g	Urban population (% of total)	-0.079	0.023	0.095	0.161	0.242	1.487
UrbanPop_m	Urban population (% of total)	8.321	40.410	60.100	57.160	76.030	100.000
UrbanPop_m_log	Urban population (% of total)	2.119	3.699	4.096	3.924	4.331	4.605
wbgi_corcon_m	Control of Corruption - Estimate	-1.186	-0.615	-0.190	0.186	0.717	2.441
wbgi_pse_m	Political Stability - Estimate	-2.275	-0.583	-0.026	0.040	0.746	1.540

Variable	Description	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
wbgi_rle_m	Rule of Law - Estimate	-1.587	-0.606	-0.066	0.133	0.839	1.930
$\ln IT$	IT infrastructure (average over 2002-12)	1.972	3.797	4.322	4.218	4.828	5.166
$\ln IT_1$	IT infrastructure (average over 2002-06)	0.446	2.757	3.768	3.556	4.657	5.059
$\ln IT_2$	IT infrastructure (average over 2008-12)	2.597	4.265	4.654	4.557	4.975	5.254
$\Delta \ln IT$	IT infrastructure growth rate from 2002-06 to 2008-12	0.084	0.314	0.932	1.00	1.515	2.862

Note: The Suffix '_m' denotes that the variables is averaged over the years of 1980 to 2000. Accordingly '_m_log' denotes logarithm of the averaged value, '_g' the growth rate and '_sd' the standard derivation of the specific variable.

Appendix C

Table C1: Descriptive Statistics

Variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Growth	-1.148	-0.061	0.371	0.288	0.649	1.582
$\ln(Y_{1980})$	6.752	8.005	8.959	9.001	10.010	12.340
$\ln(n+g+\delta)$	-3.077	-2.787	-2.616	-2.670	-2.543	-2.294
$\ln(I/GDP)$	1.727	2.948	3.209	3.134	3.368	3.831
$\ln(HC)$	0.072	0.371	1.007	0.869	1.242	1.739
$\ln(PrivateCredit)$	0.915	2.771	3.190	3.057	3.560	4.758
$\ln(Gov.Consumption)$	-3.334	-2.081	-1.711	-1.744	-1.431	-0.330
$\ln(CivilLiberties)$	0.000	0.693	1.386	1.225	1.792	1.946
$\ln(ICT)$	0.002	0.306	0.791	0.926	1.460	2.152

Table C2: Correlation Matrix

	Growth	$\ln(Y_{1980})$	$\ln(n+g+d)$	$\ln(I/GDP)$	$\ln(HC)$	$\ln(PrivateCredit)$	$\ln(Gov.Consumption)$	$\ln(CivilLiberties)$	$\ln(ICT)$
Growth	1.000								
$\ln(Y_{1980})$	-0.008	1.000							
$\ln(n+g+\delta)$	0.198	0.064	1.000						
$\ln(I/GDP)$	0.278	0.401	0.263	1.000					
$\ln(HC)$	0.405	0.683	0.249	0.395	1.000				
$\ln(PrivateCredit)$	0.331	0.488	-0.003	0.469	0.504	1.000			
$\ln(Gov.Consumption)$	0.042	-0.233	0.027	0.074	-0.089	-0.151	1.000		
$\ln(CivilLiberties)$	-0.236	-0.553	-0.172	-0.102	-0.666	-0.497	0.092	1.000	
$\ln(ICT)$	0.364	0.786	0.185	0.352	0.825	0.629	-0.238	-0.730	1.000

Note: Pearson correlation coefficients are computed between each pair of variables using all complete pairs of observations on those variables. The correlation coefficients of 89 complete pairs are computed.

Table C3: List of Covered Contries

	(1)	(2)	(3)	(4)	(5)		(1)	(2)	(3)	(4)	(5)		(1)	(2)	(3)	(4)	(5)
ALB	x		x	x		GMB	x	x	x	x	x	NLD	x	x	x	x	x
ARG	x	x	x	x	x	GRC	x	x	x	x	x	NOR	x	x	x	x	x
AUS	x	x	x	x	x	GTM	x	x	x	x	x	NPL	x	x	x	x	x
AUT	x	x	x	x	x	HKG	x		x			NZL	x	x	x	x	x
BEL	x	x	x	x	x	HND	x	x	x	x	x	PAK	x	x	x	x	x
BEN	x	x	x	x	x	HUN	x		x	x		PAN	x	x	x	x	x
BGD	x	x	x	x	x	IDN	x	x	x	x	x	PER	x	x	x	x	x
BGR	x		x	x		IND	x	x	x	x	x	PHL	x	x	x	x	x
BHR	x	x	x	x	x	IRL	x	x	x	x	x	POL	x	x	x	x	x
BLZ	x	x	x			IRN	x	x	x	x	x	PRT	x	x	x	x	x
BOL	x	x	x	x	x	IRQ	x	x	x	x	x	PRY	x	x	x	x	x
BRA	x	x	x	x	x	ISL	x	x	x	x	x	QAT	x	x	x	x	x
BRB	x	x	x	x	x	ISR	x	x	x	x	x	RWA	x	x	x	x	x
BRN	x		x	x		ITA	x	x	x	x	x	SAU	x	x	x	x	x
CAF	x	x	x	x	x	JAM	x	x	x	x	x	SDN	x	x	x	x	x
CAN	x	x	x	x	x	JOR	x	x	x	x	x	SEN	x	x	x	x	x
CHE	x	x	x	x	x	JPN	x	x	x	x	x	SGP	x	x	x	x	x
CHL	x	x	x	x	x	KEN	x	x	x	x	x	SLV	x	x	x	x	x
CHN	x		x	x		KOR	x	x	x	x	x	SWE	x	x	x	x	x
CIV	x	x	x	x	x	KWT	x	x	x	x	x	SWZ	x	x	x	x	x
CMR	x	x	x	x	x	LAO	x		x	x		SYR	x	x	x	x	x
COL	x	x	x	x	x	LKA	x	x	x	x	x	TGO	x	x	x	x	x
CRI	x	x	x	x	x	LSO	x	x	x	x	x	THA	x	x	x	x	x
CYP	x	x	x	x	x	LUX	x	x	x	x	x	TTO	x	x	x	x	x
DEU	x	x	x			MAR	x	x	x	x	x	TUN	x	x	x	x	x
DNK	x	x	x	x	x	MEX	x	x	x	x	x	TUR	x	x	x	x	x
DOM	x	x	x	x	x	MLI	x	x	x	x	x	TZA	x	x	x	x	x
ECU	x	x	x	x	x	MOZ	x		x	x		UGA	x	x	x	x	x
ESP	x	x	x	x	x	MRT	x	x	x	x	x	URY	x	x	x	x	x
FIN	x	x	x	x	x	MUS	x	x	x	x	x	USA	x	x	x	x	x
FJI	x	x	x	x	x	MWI	x	x	x	x	x	VEN	x	x	x	x	x
FRA	x	x	x	x	x	MYS	x	x	x	x	x	ZAF	x	x	x	x	x
GBR	x	x	x	x	x	NAM	x		x			ZMB	x		x	x	
GHA	x	x	x	x	x	NER	x	x	x	x	x	ZWE	x		x	x	

Note: Reported are the countries by ISO ALPHA-3 codes of the total sample from the respective columns (1) of the regressions as well as their affiliation in the regressions of the columns (2)-(6).

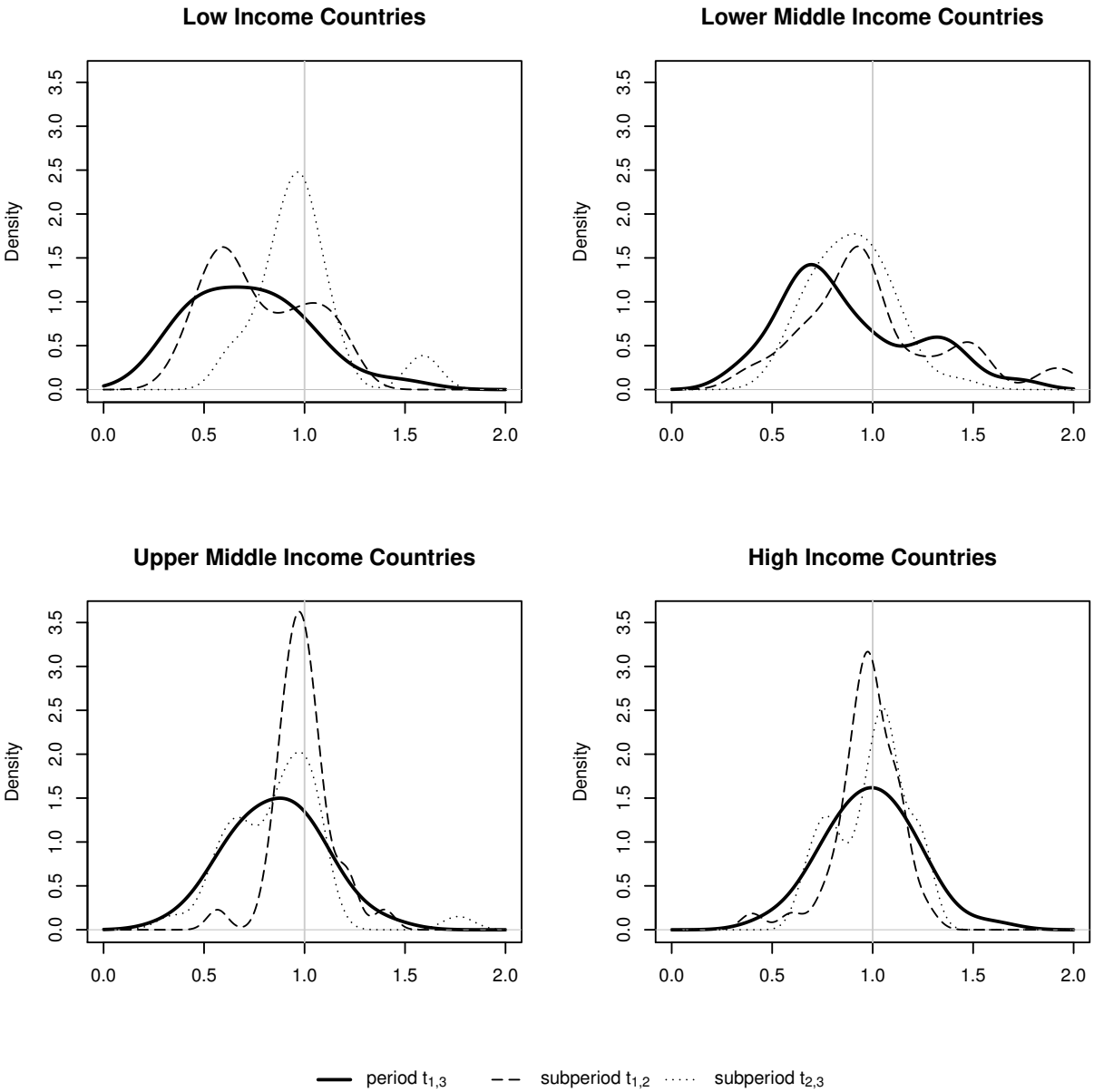
Appendix D

Table D1: List of Countries by Income Groups

Low Income	Lower Middle Income	Upper Middle Income	High Income
Bangladesh	Albania	Argentina	Australia
Benin	Algeria	Belize	Austria
Burkina Faso	Angola	Botswana	Barbados
Burundi	Armenia	Brazil	Belgium
Cambodia	Bolivia	Bulgaria	Canada
Central African Republic	Cameroon	Chile	Cyprus
Côte d'Ivoire	China	Costa Rica	Czech Republic
Ethiopia	Colombia	Croatia	Denmark
Gambia, The	Congo, Rep.	Fiji	Estonia
Ghana	Dominican Republic	Gabon	Finland
Haiti	Ecuador	Jamaica	France
Kenya	Egypt, Arab Rep.	Kazakhstan	Germany
Kyrgyz Republic	El Salvador	Latvia	Greece
Lao PDR	Guatemala	Lithuania	Hungary
Madagascar	Honduras	Malaysia	Iceland
Malawi	India	Mauritius	Ireland
Mali	Indonesia	Mexico	Israel
Mauritania	Iran, Islamic Rep.	Panama	Italy
Mozambique	Iraq	Poland	Japan
Myanmar	Jordan	Romania	Korea, Rep.
Nepal	Lesotho	Russian Federation	Malta
Niger	Maldives	South Africa	Netherlands
Nigeria	Moldova	Turkey	New Zealand
Pakistan	Morocco	Uruguay	Portugal
Rwanda	Namibia	Venezuela, RB	Slovak Republic
Senegal	Nicaragua		Slovenia
Tajikistan	Paraguay		Spain
Tanzania	Peru		Sweden
Togo	Philippines		Switzerland
Uganda	Sri Lanka		Trinidad and Tobago
Vietnam	Sudan		United Kingdom
Yemen, Rep.	Swaziland		United States
Zambia	Syrian Arab Republic		
Zimbabwe	Thailand		
	Tunisia		
	Ukraine		

Note: Reported are the countries used in this paper with a population of more than 30,000 by income group, defined by the World Bank for the fiscal year of 2007. Low-income economies are defined as those with a gross national income (GNI) per capita of $875 or less; lower middle-income economies are those with a GNI per capita between $876 and $3,465; upper middle-income economies are those with a GNI per capita between $3,466 and $10,725; high-income economies are those with a GNI per capita of $10,725 or more.

Figure D1: Density Plots of Efficiency Change in ICT by Income Group



Note: Figure shows the density plots of ICT MEA Malmquist efficiency change values (x-axis) according to the respective income groups and investigated periods. Used is a Gaussian kernel. The smoothing bandwidth is estimated by the method of Sheather and Jones (1991).