Emergence and Diffusion of Institutions and their Effect on Economic Growth

Manuscript committee: Prof.dr. J.L. van Zanden
Prof.dr. W.H.J. Hassink
Prof.dr. E. Stam
Prof.dr. O. Galor
Dr. J. Vanderstraeten

ISBN
Printed by

This book was typeset using LaTeX.

Emergence and Diffusion of Institutions and their Effect on Economic Growth

Het onstaan en de diffusie van instituties en effecten op economische groei
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op vrijdag 17 februari 2017 des middags te
2.30 uur

door

Krista Bruns
geboren op 7 oktober 1988, te Heemskerk

Promotoren:	Prof.dr. J.G.M. van Marrewijk
	Prof.dr. A. van Witteloostuijn
Copromotoren:	Dr. M.W.J.L. Sanders
	Dr. M.C. Schramm

Voor mijn ouders

Acknowledgements

"Going only part of the way is not the same as going the wrong way."

– Jostein Gaarder, *Sophie's World*

I found out the answer to Life, the Universe and Everything already at a young age, when I was reading the book "The Hitchhikers Guide to the Galaxy". It turns out to be 42. However, I also learnt from this piece of literature that you need to know exactly what the question is to understand what the answer means. Writing a dissertation, much like real life, is about finding out what you want to spend your energy on. A lot of time goes into reconsidering and reformulating your research questions. This is sometimes frustrating, but then you start accepting that this journey is actually the core of your research. This dissertation builds upon the question "Why are some people so rich and others so poor?" I think this question, in one form or another, is one of the most fundamental questions an economist can ask and it cries out for an answer. I do not want to claim that I have found the answer to this question or the solution to worldwide poverty, but I hope to have contributed to proving and disproving some of the possible hypotheses. Writing this dissertation has been an experience that I greatly enjoyed, having the freedom to fully dive into something. Yet it has also been a very challenging experience at times. I could not have done it without the scientific and moral support of many individuals.

First and foremost, I would like to thank my promotors and supervisors. I actually had a whole team on my side, all with their own strengths and expertise. Charles, you were mostly on the sideline, but your experience has proven very worthwhile in the final, stressful stages of writing this dissertation. Even from 9000 kilometers away you provided guidance and comments within a day. Arjen, you became so important during the last two years of my PhD. The trips to Tilburg always brought calmness and focus. When I was stuck, a few well-chosen words from your side gave

me back confidence and ideas for new solutions. Mark, we have known each other for almost ten years now. Although we are very different in the way we work and we sometimes had intense differences of opinion, I cannot forget your dedication and perseverance that helped to complete this project. You inspire me because you have very strong beliefs about how the world works and how it should work and are willing to fight for them. I also want to thank you for bringing me into contact with Oded and setting me on the trajectory for a research visit to one of the finest institutes of higher education in the world. Finally Marc, you were the light breeze in this sometimes heavy process. If I wanted to brainstorm about a topic, I could always come to you. From the top of your mind you would know at least ten papers that could help me progress.

I also want to express my gratitude to the members of my reading committee, prof. dr. J.L. van Zanden, prof. dr. W.H.J. Hassink, prof. dr. E. Stam, prof. dr. O. Galor and dr. J. Vanderstraeten, for their conscientious reading of the manuscript. Oded, you deserve a special mention here as you made me feel welcome at Brown University. Your research is very inspiring and I could not think of anyone better suited to have helped me with my modest first steps into modeling economic growth. Next, I want to thank my co-author Niels Bosma. Thank you for your contributions to the fourth chapter of this thesis. You are an expert on measuring entrepreneurship and your thesis has been on my desk during my PhD as an example of well-structured, good research. Utrecht University School of Economics (U.S.E.) is a young institution and I have witnessed it grow the last few years. I have experienced U.S.E. as a place where colleagues always want to help, chat about research ideas or give feedback during the research seminars. I want to express my gratitude to all of my colleagues for that. I have also greatly enjoyed teaching, particularly Macroeconomics, and I want to thank all the students and fellow teachers for making those three months every year a great adventure.

The PhD life can be lonesome, but I feel blessed to have gone through it together with the group of PhD candidates at U.S.E. It is never hard to find someone to have a coffee with, talk about research or complain to about grading exams. This way, colleagues have become friends. We ventured outside of the green gates during our many PhDinners, PhPubcrawls, bootcamps, karaoke nights and other adventures. The PhD community truly represents an overlapping-generations model. Saraï, Britta, Suzanne, Lu, Ioana, Emily, Emre and Malka, thank you for showing me how it is done, for reassuring me in times of distress and for helping me find my way in life after academia. Daan, Ian and An, as members of my own generation you

provided me with great support and more importantly, a wonderful sense of humor. As for the 'next generation': Emily, Jeroen, Dea, Tim, , Milena and Leydi, thank you for all the lunch talks and coffee breaks, and good luck in the coming months and years. I have also enjoyed the company of all my officemates. Maryam, you are such a pleasant influence, always bringing peace into the room. Werner, the structure you have accomplished on your desk and in life in general is something I can only dream of. Fortunately, I can still beat you in running. Zori and Franziska, I watched you go through many of the same struggles that I had already overcome, both in your own, opposite ways. Zori, thank you for always providing a listening ear and for helpful econometric tips. Franziska, thank you for supplying me with chocolate in times of need and for always being in a good mood.

The two individuals who have had to suffer most under my crises are my two paranymphs, Joyce and Swantje. Joyce, we basically grew up together as researchers, going from making songs about Hamiltonians to actually publishing a dissertation. You never settle for anything less than perfection, whether it concerns a dress or a mathematical equation. Swantje, you are the most calming influence in my sometimes chaotic life. I admire you for always keeping your patience and diplomacy and I think this will bring you very far in your personal life and career. Both of your strengths are exactly my weaknesses and this is why you make great paranymphs. Sandra, thank you for being my partner in crime during many conferences and summer schools. Doing research seems to come so natural to you and I look up to the way in which you have developed through your academic career. But besides that, you are also a great friend. During my PhD, I tried to take all the chances that were thrown at me. I spent time in Providence, where I met many amazing and smart people. I also enjoyed my time in the PhD Council and I want to thank all the fellow council members for the work we put in to make our professional environment better. Finally, my short PhD break at NWO gave me new inspiration to finish my own research projects and I want to thank all my colleagues for the great time I had there.

In addition to my friends from work, I have received support from my friends outside academia. Sports have always been a great outlet for me and therefore I found my best friends in times of sweat. Liefjes van Damesch '10 and MGD '11, you are the living proof that sometimes a 'benevolent planner' knows best. We were thrown together in a boat, but I could not have picked a better group of friends myself. Each of you is achieving great things in life, from rowing in the Olympics or the Boat Race to starting a company, from saving animal and human lives to telling hidden stories

of the refugee crisis. You never stop to inspire, surprise and motivate me. I also want to thank many others from Triton and Phoenix together with whom I have coached, rowed and run. What I like about running is that, as opposed to research, what you put in is what you get out, in a quite linear way. Running has become my (un)healthy addiction and fortunately I always found good company to share this with. I also want to thank my friends from UCU and high school for making these years so fun and carefree. Coosje and Francien, I loved having you as roommates and still as great friends. Marloes, Sandra, Maaike, Floor, Joyce and Daniëlle, although we do not always see each other on a regular base, we have been through so much together and I know you will always be there for me.

During the past few years, I became more and more aware of the importance of family. I could not have written this dissertation without your unconditional love and support. Oma Bruns, I am sad that you could not witness the completion of my PhD, but I am sure you are proud of your 'Grootdochter'. Opa and Oma Tillmans, you are my real-life superheroes and I never get bored of telling my friends about you. Opa, you are the best example of a lifetime of enjoying sports and you always have new amusing stories to tell. Oma, you radiate so much warmth that it is never quiet on the Hertainglaan, as everyone wants to spend their free afternoons with you. Vicky, we are true sisters, we quarrel over everything, but we sometimes have to admit that spending time together is pretty great. I hope to spend many more zussendagen with you. I am also very proud of the joy and ambition you found in your career and I think every child deserves a teacher like you. Sjeffies, I am the product of the two of you, so for all of my bad characteristics, but also for all of my achievements I have you to thank. Therefore I dedicate this dissertation to you. Mama, thanks for teaching me great work ethics by kicking me out of the house at age 14, not to come back until I found a summer job. Giving up is never an option, I learnt that from you. You always know better what I need than I do myself, and most of the times you are right. Papa, you and me are so much alike. When things get tough, our solution is to work harder, but we are also sentimental sobs. When I am in need, you drop everything to come help me. You do not want anything to stand in the way of my dreams and still help me to achieve them every day. Thank you for all your support.

Krista Bruns
Utrecht, February 2017

Contents

List of Tables

List of Figures

1

Introduction

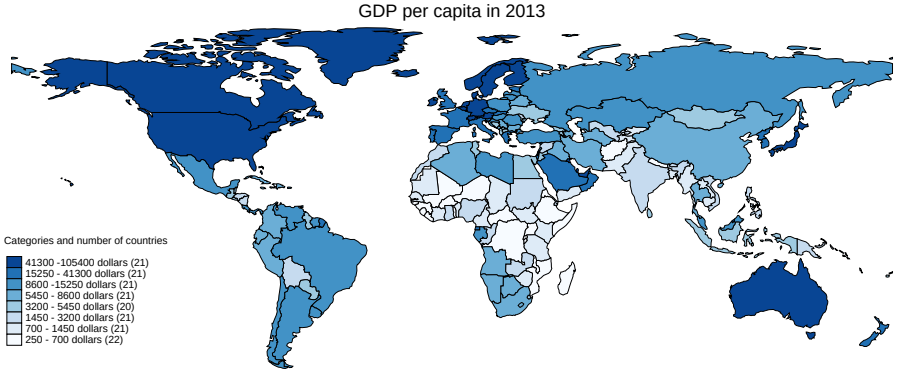
1.1 Motivation

In this dissertation I study the emergence, diffusion and impact of growth-enhancing institutions by building on different strands of literature in economic growth theory, New Institutional Economics and Unified Growth Theory. I introduce the dynamic interaction between geography, demographics and institutions that helps to explain the emergence and persistence of core-periphery clusters that drive economic growth. By now there is a large consensus that institutions contribute to differences in economic growth rates (Acemoglu et al., 2001; Greif, 2005; Acemoglu, 2009; Galor, 2011; Acemoglu and Robinson, 2012), but the mechanisms through which they developed in interaction with demography and geography have been relatively under-researched. In this dissertation I begin to open this 'black box'.

In economics, welfare and economic growth are expressed using the Gross Domestic Product (GDP) per capita.[1] Due to the principle of compounding, small differences in economic growth rates, persisting over several decades, lead to large differences in the level of GDP and thus in the standard of living between countries. This is precisely what was observed over the last two centuries. Figure 1.1.1 shows the pattern of inequality across the globe, using figures of annual GDP per capita in 2013. Each category contains one-eight of all the countries. This shows that equally many countries have a GDP per capita in the range of 250 to 700 dollars (constant prices) as in the range of 41,300 to 105,400 US dollars. In 2013, the

[1] The extent to which this is a good measure of welfare is discussed in the concluding chapter.

Figure 1.1.1: World inequality in GDP per capita

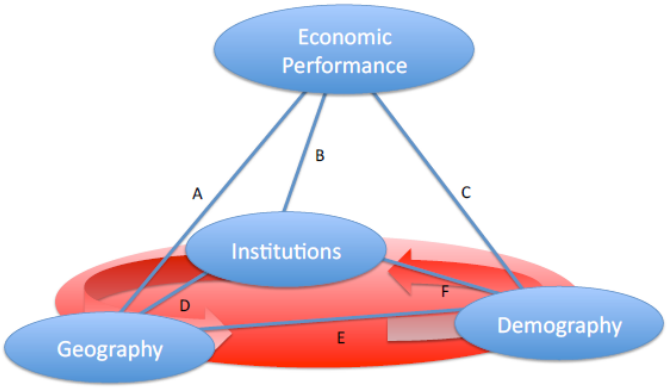


country with the lowest GDP per capita was Burundi, with income per inhabitant at 244 dollars per year, while Luxembourg was at the other end of the scale with 105,447 dollars. The ratio of the richest to the poorest country in 2013 is thus 432. Galor (2011) reports that when comparing supranational regions, the ratio of richest (Western Offshoots) to poorest (Africa) has increased from 3:1 in the year 1820 to 18:1 in the year 2000. Based on my calculations, this ratio is now approximately 22:1. It is clear that worldwide inequality has increased dramatically since the beginning of the nineteenth century.[2] This growing inequality is a consequence of the two centuries of sustained economic growth in the leading regions, since the rich get richer and the poor stay poor. The topic of growing inequality has regained the attention of the wider public through the works of Piketty (2014), who argues that the accumulation and unequal distribution of capital in the era of modern economic growth is the main driving force behind the growing gap between rich and poor.

Economists have addressed the question "Why are some people rich and others poor?" throughout the history of economic thought, going as far back as Smith (1776). The literature has since then developed several hypotheses on the fundamentals of growth and inequality. In figure 1.1.2, I cluster these hypotheses under the headings *Geography, Institutions* and *Demography*. The links between geography and economic performance (A) were explored by, e.g., Diamond (1997) and Sachs (2005). This literature argues that biogeographic (initial) conditions still determine to a large extent the economic fates of people, although the focus has shifted from direct

[2]This is in fact the case regardless of the exact variable taken (GDP per capita, wealth levels, broader concepts of welfare) and the level of aggregation that is studied.

Figure 1.1.2: The dynamic interaction between geography, people and institutions



effects of geography on welfare to indirect links. The links between demographics and economic growth (C) were explored as far back as Malthus (1798), and still features prominently in the work of, e.g., Galor and Weil (1999) and Galor and Moav (2007). In this literature, the interaction between population (growth) and economic development is key in understanding the transition from Malthusian Stagnation to Modern Economic Growth. Under Malthusian dynamics, all technology gains are translated into a larger population rather than into gains in per capita income. Therefore a demographic transition was needed for the onset of sustained economic growth. Finally, such authors as Coase (1960), North (1991) and Rodrik (2004) have emphasized the role of institutions as a third fundamental cause of economic growth and prosperity (link B). More recent empirical work by Acemoglu et al. (2001), Sokoloff and Engerman (2002) and Nunn (2008) supports this view.

In this dissertation, I take an integrative multidisciplinary approach, explicitly recognizing and analyzing the interactions between demographic characteristics of a society (population density, size and composition), its institutional dynamics and the (bio)geographic initial conditions and geography of the area it resides in (links D, E and F). I propose that prime institu-

tions such as property rights and contract enforcement emerge in response to geographic initial conditions, respectively a favorable climate for agriculture, in isolated societies. Subsequently these institutions co-evolve with demographic developments - higher population density requiring more sophisticated institutions enabling further specialization, higher income and population growth. Improvements in communication and transport technologies then cause an intensification of contact between societies in the form of communication, trade, migration and conquest, especially since the 16th century, which helps to spread these institutions. This historic interaction between population, institutions and geography helps to explain much of the modern geographic pattern of economic activity (Putterman and Weil, 2010).

An ongoing process of globalization, however, does not imply that in the future "the world is flat" (Florida, 2005; Friedman, 2005) and that all regions in the world will converge to the same set of institutions and corresponding levels of economic prosperity. Instead, the evidence in economic geography suggests that even in highly developed economies spatial specialization persists (Florida, 2005; Brakman and Marrewijk, 2008; Christopherson and Martin, 2008) and that core cities operate under a different set of 'rules of the game' than rural, peripheral areas. To help understand the distribution of economic activity at a more fundamental level, I analyze the emergence, diffusion and impact of institutions as reflected in spatial patterns of economic activity. This dissertation is structured in three research papers in three chapters. A first paper addresses how a demand for complex institutions may emerge endogenously. A second paper zooms in on the (historical) diffusion of such an institution through contact. A third paper links contemporary spatial patterns of economic performance to entrepreneurial activity and studies whether this relationship is moderated by differences in the quality of latent institutional 'ecosystems'. The chapters build on each other but can also be read as individual papers. Therefore some repetition may occur in the introduction and theory sections.

1.2 Natural selection in economic theory

A large part of this dissertation has an evolutionary perspective. The story of the development of institutions and the consequences for differences in welfare today is a story of evolution. The tracing of lineages reveals that all human populations descend from a single small group of individuals in Africa, and that as soon as this group split up into small bands migrat-

ing and adapting to different environments, they began to develop new social institutions and more complex forms of social organization. This new social behavior was not only a cultural but also a biological change, as modern humans developed more cognitive power than their ancestors as their brains grew in size. I feel that studying these endogenous, dynamic processes can only be done with the acknowledgement that evolutionary forces play a role and that competition is the selecting mechanism, whether the evolutionary process is cultural or biological or a combination of those two. However, I do not wish to fall victim to evolutionary or geographic determinism. The mechanisms that lead to differences in economic growth are influenced by geographic factors (including location, climate, the distribution of plant and animal species, soil type and topography) and non-geographic factors, subsumed under the term culture, history or institutions, and even by decisions of individual people. Many examples, such as the economic divide between North and South Korea, illustrate how places with the same biogeographical initial conditions can experience hugely different paths of development.
The aim of this dissertation is to analyze the global mechanisms through which the three main factors of interest —geography, demography and institutions —interact and determine economic outcomes. However, I also show where these global mechanisms still leave room for heterogeneous development, as a deterministic theory does not serve the purpose of describing the observed variety of outcomes.

In a sense this perspective does not differ from the original doctrine of Darwin, as Neo-Darwinism argues that this doctrine can explain the "varied and often puzzling facts of evolution: the different types and degrees of adaptation; the (geologically speaking) rapid transformation of some types side by side with the unchanged persistence of others; the coexistence of lower and higher forms; the succession of constantly more improved types; extinction; the facts of geographical distribution; the evolution of insect societies based on elaborate instincts." (Huxley, 2003, p.xv)

Another commonly heard criticism of evolutionary theories or perspectives in economics is that they are often used to describe how an optimal solution can be achieved, as if evolution or natural selection has an end goal in mind. As Aristotle said in his "Physicae Auscultationes" (lib 2., cap8, s.2.):

> *"The rain does not fall in order to make the corn grow, any more than it falls to spoil the farmer's corn when threshed out of doors, but what hinders the different parts of having this merely accidental*

relation in nature?"

This quote serves to illustrate that, although in the end, a collection of events may appear as if they happened or were made for the sake of something, this is due to adaptation and natural selection over random events. Natural selection does not imply conscious choice and it does not necessarily select the optimal solution. It merely means that those individuals, societies or systems that have an evolutionary advantage over others in the same circumstances, will be more likely to survive if evolutionary pressure is high. Evolutionary pressure is especially high in disruptive times like famine, war, climatic volatility and scarcity of resources. Taking an evolutionary perspective on economic theory requires some reflection on the assumptions that I make. One critical point is the definition of evolutionary fitness. In biological evolution, this concept is quite clear: the individuals or species with highest reproductive success have the highest evolutionary fitness. How does this concept extend to societies? I argue, based on historical evidence, that indeed in early stages of economic and societal development, societies that grew fastest in size also had an evolutionary advantage when selective pressure was high. These societies are able to free more resources for public goods and warfare through specialization. Translating this to macroeconomic concepts used today, it means that evolution 'selects' those societies that are able to produce the highest GDP levels. Societies with high GDP per capita levels also attract more immigrants and thereby grow even faster in size. The idea that institutions, their development and the conditions under which they survive, should be studied in an evolutionary manner is confirmed by Binmore (2011, p. 1):

> *"The moral rules that really govern our behavior consist of a mixture of instincts, customs, and conventions that are simultaneously more mundane and more complex than traditional scholarship is willing to credit. They are shaped largely by evolutionary forces —social as well as biological."*

There is some criticism on the use of evolutionary models in economics, as they are often interpreted as if a benevolent planner designs society in a way that makes everyone better of. In chapter 2 we elaborate on how evolutionary forces should be incorporated and interpreted in economic models of optimization.

1.3 Biogeographical conditions

In this dissertation, I argue for the importance of geographic or biogeographic initial conditions. What are those and why have they become so prominent in research on economic growth? The study of the determinants of economic growth is heavily complicated by the fact that most hypothesized factors suffer from endogeneity, in the form of simultaneous development of economic growth and its possible explanatory factors, making it hard to determine what causes what. Theories on economic growth have progressed from proximate causes —physical and human capital accumulation and trade (Solow, 1956; Mankiw et al., 1992; Becker et al., 1994; Frankel and Romer, 1999) to geography or genes as direct influences (Gallup et al., 1999; Sachs and Malaney, 2002), to deeper causes —luck, culture, institutions (Granato et al., 1996; Rodrik et al., 2004; Acemoglu et al., 2005; Tabellini, 2010). However, even with regards to the deeper causes, one can ask whether these are really the fundamental determinants of economic growth. What causes the fact that one society has developed institutions superior to those of other societies, allowing it to grow faster? To find an answer to that question one needs to go back to a time before differences in prosperity existed and study the only two concepts that can be truly 'exogenous': the initial biogeographic conditions and the distribution of genes in society that came about through the initial migration of humanity out of the African continent.

Diamond (1997) shows that biogeographical conditions, specifically climate (volatility), soil suitability for different crops and biodiversity, and the orientation of landmasses influenced the timing of the Neolithic transition and thereby the ability for some societies to gain a head start in the development of institutions. The genetic distribution does not refer to a specific genetic profile that would be in some sense superior, but to the genetic diversity that resulted as a consequence of the distance the population covered in the initial out of Africa migration. Ashraf and Galor (2013b) show that variation in migratory distance and thus genetic diversity has had a persistent hump-shaped effect on comparative economic development, indicating that there is an intermediate optimal level of diversity.[3]

These initial conditions provided more and less favorable conditions to an earlier agricultural transition for different societies. This agricultural transition in turn brought about many societal and institutional changes

[3]Diversity positively affects income through innovation but negatively as well, as it hinders the possibilities for peaceful cooperation.

that led to later economic development. This has led many authors to include geographical data, such as the distance to the equator or the annual rainfall, as explanatory factors for economic growth, aiming to capture the indirect rather than the direct effect of geography. Diamond (1997) put forward the hypothesis that the distance to the equator was an important predictor of standards of living today because it influenced the ease with which agricultural practices could be diffused, after agriculture independently emerged in several locations. Olsson and Hibbs (2005) and Ashraf and Michalopoulos (2015) show that biogeographical and climatic conditions indeed influence the location and timing of transition to sedentary agriculture, leading to complex social organization and eventually modern industrial production. Acemoglu et al. (2001) find that settler mortality rates in colonies around 1500 A.C. affected the type of institutions set up by Europeans and thereby explain the reversal of fortune observed in these colonized areas. We will build on this work by arguing that biogeographical conditions that allow an earlier Neolithic transition also provide the necessary conditions that enable the emergence of institutions that lower transaction costs and thereby stimulate exchange.

1.4 Institutions: setting the stage

What is meant by institutions? What kind of state of the world are we talking about when we discuss endogenous emergence of institutions 'ex genesis'? Can we distinguish which institutions are in general growth-enhancing? What were the first institutions that developed when societies started to grow and interact? Are there specific institutions without which it is impossible to progress to a period of sustained economic growth? These are all relevant questions requiring some reflection before continuing the quest into institutional development as a fundamental cause of economic growth. I elaborate further on the different types of institutions, the ways they arise and the mechanisms through which they influence society in the following chapters, but here I give a rough outline of the main issues regarding institutions. With regards to the first point, I do not hope to improve upon the definition put forward by North (1990), which has been used by scholars in the field of New Institutional Economics and also in other areas of Economics, Political Studies, Sociology and more. North (1990, p. 3) states that:

> *"Institutions are the rules of the game in a society, or more formally, are the humanly devised constraints that shape human inter-*

> *action. In consequence they structure incentives in human exchange, whether political, social or economic."*

The concepts of institutions, culture and social contracts are closely related. There are subtle differences and some authors do feel the necessity to make a clear distinction between them, but in this dissertation I am talking about the broad and complex system of all social norms, legal rules, cultural constraints and behavioral concepts that guide human behavior within a given setting. The social contract, in the words of David Hume, can be seen as a largely unrecognized consensus to coordinate on a particular equilibrium of the game of life that individuals play together (Hume, 1739). Culture, as defined by Inglehart (1997) refers to the *subjective* aspects of a society's institutions: "the beliefs, values, knowledge and skills that have been *internalized* by the people of a given society, complementing their external systems of coercion and exchange" (p. 15).

With regards to the second question, institution-free societies do not exist today and have never existed. Hobbes (1651) describes a violent state of nature in which men know no virtues and will never agree to any cooperation. Fukuyama (2011) refers to this as the Hobbesean fallacy: the idea that humans were primordially individualistic and at a certain stage made the rational calculation that cooperation was the best way for them to achieve their individual ends. But authors such as Brosnan (2006) and De Waal (2009) find that concepts of fairness, rejection of inequality and morality can also be found among other animal species. If one accepts the broad definition of institutions given above, animal bands too have certain norms and hierarchies that they live by, and the idea that institutions at some point developed out of pure anarchy is an erroneous thought. However, as I show in Chapter 2, it is possible to make a distinction between organic institutions that do not require active enforcement, such as the development of language and other costly forms of coordination, and more complex, designed institutions, requiring active design, collective investment and reinforcement. I also show in this chapter that there is not one set of institutions that will be growth-enhancing to all societies at all times. Institutions that were selected 'by evolution' under certain circumstances often remain in place even when the circumstances have changed. They may still be growth-enhancing in those societies as other processes and systems have adapted themselves to this set of institutions, and they are therefore very hard to replace, but this does not guarantee that the same set of institutions would stimulate growth when copied to a different society in other circumstances. One size does not fit all, as was often confirmed in

studies on developing countries (Pritchett and Woolcock, 2004; Rodríguez-Pose, 2013).

Giving a chronological ordering of the emergence of institutions is hard, as there is no evidence on this process and it seems to have a large "chicken and egg" connotation. It is argued that language and religion were (among) the first institutions to emerge in human bands (Searle, 1995). These are both examples of institutions that arise organically; they are established by human conventions without any promise and in the case of language, without the need to reinforce it, as not using the institution will give you large disadvantages.

The last question was whether any institutions form a sine qua non for sustained economic growth. Economic theory links growth to specialization, but this implicitly requires a low level of transaction costs. As societies grew in size, exchanges with non-kin members became more prevalent and cheating and shirking increased, leading to higher transaction costs. I thus argue that the first institutions that form a basis for economic growth were those that lowered transaction costs between strangers, thereby increasing the volume of sequential transactions between agents who do not share any kinship ties and cannot build on inter-group trust. Galor (2011) shows in his Unified Growth Theory model that population growth is the main driver for technological progress, as it has a positive feedback loop to technological growth, up to the point where population growth stagnates to free resources for a quality-quantity substitution in offspring (Becker and Lewis, 1974). From that point onwards, the positive feedback loop between human capital formation and technological progress takes over the role of population growth. I add to this proposition the idea that with population growth and higher population density come larger transaction costs. Hence, we define as growth-enhancing institutions any institutions that help to reduce transaction costs or prevent them from rising to prohibitive levels. These institutions are essential for the transition to sustained economic growth as described by Galor (2011).

Such institutions are a fundamental determinant of the extent to which a society can benefit from its factors of production and create sustained economic growth. A study of the mechanisms through which institutions come about and change, as well as the mechanisms through which they actually affect the relation between those factors of production and economic growth is thus highly relevant. The chapters in this dissertation build upon each other in different ways.

First, they have a chronological order. Chapter 2 can serve as an introduction to the rest of the dissertation, elaborating on institutional (change)

theory and the definitions of institutions. In this chapter, I start with the question how institutions have arisen 'ex genesis', i.e., what has led to pristine state development as Fukuyama (2011) describes it? I acknowledge that this pristine state formation has never been observed by anyone and truly institution-free societies are an illusion. Therefore I cannot attach a precise time period to this process, but I show that the transition from nomadic hunter-gatherer existence to an agricultural society can be a driving force for the emergence of economic and political institutions. Because of low population density and low mobility, contact between societies has been limited for the largest part of human history and societies were able to develop their institutional framework in isolation, to a large extent. Notably, there have always been wars between societies that temporarily disturbed or completely destroyed this formative process. However, as is shown throughout this dissertation, agricultural production, a sedentary lifestyle and the increasing technological change that comes with it leads to higher population pressure on the resources in a given area. Therefore, first contact within also between societies becomes more prevalent. This links the second chapter to the third chapter.

In the third chapter I study how different types of contact affect the diffusion of an institution. Once emerged, successful institutions quickly spread, as did technologies. However, the spread of most of these institutions left no paper trail or other form of physical evidence. The spread of language, religion, nuclear families as organizational units and early state formation are largely lost in the mists of time to us. Therefore, to illustrate the diffusion of institutions I turn in chapter three to the diffusion of female voting rights. The adoption of full suffrage for all women was well documented and can be related to data on inter-country contact. Specifically, I look at how war, trade, diplomatic exchanges, colonial ties and geographic distance with or to other prior adopters of full voting rights for women may lead to earlier adoption of this institution of a given country. Institutions are continuously in development, either due to endogenous changes within a country or through contact or shocks, but the path-dependent nature of institutional formation and the endogeneity of growth-promoting institutions and economic growth lead us to hypothesize that the current state of the world can be seen as a core-periphery pattern of institutional regimes or ecosystems. Like biological ecosystems, these institutional ecosystems may be found and disentangled at different levels of aggregation, and the boundaries may not always be clear.

The empirical analysis of institutional ecosystems is done looking specifically at the relationship between entrepreneurship and economic growth.

In the fourth chapter, I study the question why entrepreneurship does not promote economic growth equally in all regions. I argue that this may be due to the moderating effect of the institutional ecosystem, although the exact conditions or variables that contribute to such an ecosystem are still unclear and the disentanglement at the right level of aggregation remains an empirical challenge. In the final chapter I summarize my findings but also comment on the limitations and avenues for further research. Furthermore, I comment on the possible challenges for institutional frameworks as a consequence of recent social and political developments.

Besides this chronological sequence, the chapters also build upon each other with regards to the level of detail and the type of predictions that I can take away from each chapter. The process of endogenous state formation was not observed by anyone, nor is there any evidence of when this happened exactly in different parts of the world, so studying this process is automatically a theoretical exercise. The contribution of Chapter 2 is twofold; I propose a mechanism through which the Neolithic Transition, the transition from hunter-gatherer societies to agricultural societies, leads to a growing latent demand for institutions, and I evaluate the literature on this topic and sketch a flowchart describing why this demand did not translate into the actual formation of institutions or states in all societies. In Chapter 3, I can use historical data, and I use a duration analysis to predict the adoption of the institution of the extension of the franchise to women as a consequence of contact between societies, controlling for within-country factors and external shocks.[4] This shows us that the adoption of an institution cannot be studied without incorporating contact into the model, although general causal links are hard to establish. Contact and the adoption of the institution are likely endogenous and it is not clear to what extent the results on the diffusion of female voting rights can be extrapolated to other institutions, more or less embedded and diffusion at earlier or later periods in history still. What can be learnt from this chapter is that when controlling for within-country factors and external shocks, the effect of diffusion, especially through having many prior adopters as neighbours and through colonial dependencies, is important. Furthermore, I find that the channels and importance of diffusion are different for early and late adopters.

Finally, in the fourth chapter I use detailed regional contemporary data. This paper makes a contribution to the literature on entrepreneurship and

[4]The study of female suffrage is mostly a pragmatic choice. It would also be very insightful to study the diffusion of an institution like contract-enforcing institutions or the first prevalence of property rights, but as this is very loosely, if at all, documented we opt for a more visible example here.

institutions. Although most authors agree that entrepreneurship contributes to job creation and economic growth, this relation is not always found in empirical studies. It is often suggested that entrepreneurship can only lead to increases in welfare in an entrepreneurial culture or enhancing entrepreneurial ecosystem. I contribute to this literature by putting forward a methodology to disentangle entrepreneurial ecosystems and their moderating effects in a latent way rather than predefining the variables that feed into them and their exact scope.

1.5 Emergence

In Chapter 2, I build on Unified Growth Theory, as proposed by Galor (2011), to answer the question how institutions emerged in societies when these started growing in size and complexity. This literature models population dynamics and biogeography, focusing on the transition from stagnation to growth and the factors that have determined the pace and timing of this transition. Galor concludes that variations in prehistoric biogeography conditions have persistent effects on the formation of human capital, which has launched different convergence clubs. In this literature, however, the institutional context is (implicitly) assumed. My contribution is that I illustrate a general mechanism that leads to institutional formation, while also allowing for spatial variation in economic activity. Additionally, I give a survey of the literature on institutional emergence and institutional change, laying out the difference between the different types of institutions, designed and organic, enforced and unenforced, formal and informal, and the processes that lead to them. I conclude that modeling the process of institutional emergence requires a multi-layered approach and I show how this can be structured in a flowchart diagram.

I propose a theoretical model that shows how a society growing in size —because it is capturing productivity gains from the transition to agriculture —develops a latent demand for institutions. Adopting these institutions will give this society an evolutionary advantage and allow it to grow further. However, from game theory and principal-agent theory it is known that jointly cooperating to achieve socially optimal outcomes and enforcing the rules you agree upon can be very hard to achieve (Nash, 1950; Laffont and Martimort, 2009). This explains why a linear path of institutional development is hardly ever observed. Data on the emergence of institutions such as property rights and contract enforcement are by nature imperfect and hard to come by. Furthermore, the model cannot predict the exact tim-

ing of the emergence of demand for institutions, so I cannot test the model predictions empirically. Rather, the model illustrates how biogeographic initial conditions and the coevolution of population and institutions can produce heterogeneous patterns of institutional development across time and space and identify the key mechanisms responsible for such heterogeneity in outcomes. I illustrate the variety of possible outcomes with an array of historical examples. The model identifies and illustrates the key mechanisms in the lower part of Figure 1.1.2 (particularly link F), building on the mechanisms put forward in Unified Growth Theory (Galor, 2011).

1.6 Diffusion

In the third chapter, I address what happens to institutions in different societies as contact between those societies becomes more frequent and intense. At some point, societies growing in size meet the limits of their environment and the available resources. In hunter-gatherer societies, this problem could still be solved in isolation by moving to another area, as the world was sparsely populated. This was no longer the case in sedentary societies present in most parts of the world post-1500. They found other ways to meet their expanding needs, venturing out through trade, war, colonization or alliances with other societies. In Chapter 3 I study how the intensity and type of contact between societies influences the diffusion of institutions. Specifically, I analyze whether the extension of the franchise to include all women, having taken place in virtually all countries between 1893 and today, hinges on contact with prior adopters of this institution. This question is addressed in an empirical duration analysis. I predict the year of adopting full voting rights for women using different types of inter-country contact, while controlling for within-country determinants and external shocks. I also study whether the relative importance of those factors differs between early adopters and late adopters.

While the adoption of suffrage for all women may not be considered an essential growth-enhancing institution, it can serve to illustrate the mechanisms through which institutions diffuse, especially if these findings can be replicated using different institutions. In choosing a certain institution for the analysis, several considerations were important. First, the institution needs to have been established in a large part of the world. This makes cross-country comparisons most meaningful. Second, it is important that data is available on the different types of contact between nations for the timespan in which the institution was adopted. Finally, the adoption of

the institution of should be formally documented in some sort of legal document, in order to connect a precise adoption date to it and make cross-country comparisons possible. This means that the research is restricted to de jure, not necessarily de facto institutional change. This is necessary, however, because it is the only objective way of comparing institutional change across countries. In future research, the analysis might be extended to include more institutions, possibly including a more recent or more historical institution, or comparing the adoption of an economic and a political institution.
I find that the probability of adopting full voting rights for women is positively affected by i) the number of neighbours who adopted the institution and ii) being involved in a war, regardless of whether this is with prior adopters or not, and iii) diplomatic ties to prior adopters. Furthermore, I find that diffusion through contact is more prevalent for late adopters than for early adopters, for whom the adoption hinges more on external shocks such as wars.

1.7 Contemporary institutional patterns and economic activity

Another puzzle in the empirical growth literature is the fact that theoretical positive relations between inputs or factors and economic growth are not visible in the data, as the results are often weak and ambiguous. One could pin this on poor quality data, but I argue that it could also be due to the moderating role of the institutional framework. For my fourth study, I take into account the insights from the previous two chapters and take as given the institutional pattern today, consisting of different layers of institutional frameworks, some more growth-enhancing than others. This could be interpreted as a core-periphery pattern of institutions. The fourth chapter deals with the question whether this institutional pattern can be found in the data and can help explain why entrepreneurship is not found to contribute equally strong to economic growth in different empirical studies.

This core-periphery pattern can be found and is reproduced at different levels of aggregation; one may think of the global divide between developed and developing countries, but also within continents there are differences with regards to the institutional framework and how that drives growth, and the same can be said for individual countries. The divide may have something to do with the industrialization of certain areas whereas others remain mainly agricultural. Industrialization brings rapid techno-

logical change and this requires a different and continuously changing institutional framework. The core and the periphery also clearly differ in terms of population density, and from chapter 2 I take away that a higher population density brings more transactions with strangers, requiring institutions that can serve as a substitute for interpersonal trust in market exchange. Friedman (2005) argues that globalization and telecommunication have made the world flat, and your location is irrelevant for your productive capacities. However, many authors also show that this is not the case, especially at the regional level, where the intangible benefits of clustering in a certain location have proven to be important (Porter, 2000; Martin and Sunley, 2003; Rocha, 2004). High-growth enterprizes often cluster together in certain areas, while other regions may have higher total numbers of entrepreneurs who do not produce these impressive rates of economic progress. The benefits have been attributed to knowledge spillovers, but they may also be connected to concepts like an entrepreneurial culture or entrepreneurial ecosystem, and to certain sets of growth-enhancing institutions.

I connect to this literature by studying the relationship between entrepreneurship rates and economic growth. This relationship has not been widely studied, as data limitations and endogeneity provide serious problems, but when it has been studied results are not uniform. Some authors only find an effect for certain lags of time or for certain types of entrepreneurs (Fritsch and Schroeter, 2011), and Wennekers et al. (2005) find a U-shaped relationship between nascent entrepreneurship and the level of economic development.

I argue that the heterogeneous relationship between entrepreneurship and regional growth originates from a moderating effect of the region's institutional framework. Baumol (1990) already put forward the hypothesis that entrepreneurial talent self-selects in either productive, unproductive or destructive forms of entrepreneurship, depending on the conditions in the environment. I attempt to make this empirically visible in a study on European regions at the NUTS-2 level. However, institutional quality is a latent concept that cannot be captured directly in one or more regional-level variables. Any proposed proxy will be highly endogenous and can never truly capture this multidimensional concept, so I will allow for this moderating relationship to come out of the data, if it is present. I use a latent class analysis to see whether regions can be classified into multiple groups, dependent on the strength and possibly size of the relation between entrepreneurship and economic growth. A next step would then be to find variables that predict class membership and therefore help to

identify what constitutes this moderating entrepreneurial ecosystem or institutional framework. However, I find that the relationship between entrepreneurship rates and economic growth is very weak for this sample and that no heterogeneous classes can be identified. I do think that the methodology used in this study is promising for further studies. Latent Class Analysis can be used to capture the moderating effect of institutions, explaining possibly heterogeneous relationships between factors (like innovation, human capital and entrepreneurial activity) and economic growth in wider samples, if possible at lower levels of aggregation.

While the three research papers develop in distinct directions, they stem from the same belief that the interplay of biogeographical initial conditions and subsequent population dynamics have formed present-day institutional patterns. The study of the emergence, diffusion and moderating mechanisms of institutions is a relevant exercise because institutional patterns have a large impact on the distribution of income today, though this (causal) impact is challenging to reveal in the data.

2

The emergence of market-supporting institutions in early agricultural societies

Objectives

- To organize existing literature on institutional development and institutional change in a structural way
- To shed light on the process through which institutions emerge, focusing on demand and necessary and sufficient conditions for the establishment of institutions and on the conditions that determine their (in)stability
- To show in a mathematical model how population dynamics after the Neolithic Revolution led to a latent demand for designed, complex institutions
- To extract generalizations or testable predictions from this collection of theories and models to illustrate how some societies in human history were able to overcome obstacles on the way to sustainable institutional development while other societies failed
- To summarize in a flowchart diagram the different obstacles that societies needs to overcome to successfully establish and sustain institutions

2.1 Introduction

From experimental evidence (e.g. De Waal (2008)) we know that not only humans, but also animals can sustain cooperation or show altruistic behavior if that leads to future benefits for themselves or for their kin, even if the act itself reduces the donor's fitness at that moment. However, Stevens and Hauser (2004) and Hagen and Hammerstein (2006) show that reciprocity among non-human animals is constrained largely to interacting with relatives (kin selection) or with those who have cooperated before (reciprocal altruism). This way, groups of animals differ from modern human societies: when humans make a transaction, many people who have never met were involved in the full chain of actions, from providing the inputs to getting the final product or service. Yet strangers willingly cooperate without the need to threaten the other party with punishments if they defect. This is possible because humans have developed a complex system of institutions throughout history, designed and organic, enforced and unenforced, formal and informal, building on one another. This collection of institutions has immensely decreased transaction costs of deals between people all over the world and thereby improved economic performance. Differences in the institutional setting between countries may therefore also explain differences in economic performance. Economists have been studying institutions implicitly and explicitly since the beginnings of economic science. However, the role assigned to institutions in economic theory and the perspective on whether institutions can be changed, and if so, how, has greatly changed over the history of economic thought.

Before the 1980s, economists were mainly concerned with equilibria and took the institutional setting as given, allowing only for incremental changes in institutions (Hodgson, 2007). To name a few examples of this classical economics approach; to capture the gains from specialization in Adam Smith's *Wealth of Nations*, human cooperation needs to be sustainable and markets need to function (Smith, 1776). Smith proposed the idea that economic outcomes are not always the result of conscious decisions or design and that socially optimal outcomes and social order can emerge without direction. Ricardo et al. (1819) thought of social organization as something that had changed in the past, but had now reached maturity and would not change substantially or materially anymore. This view was shared by Clark, who argued that the only difference between primitive society, isolated man and the capitalist society is the degree of specialization. He believed that all three were bound by the same natural laws of human nature and that this leads (through the mechanisms of prices and markets)

to a well balanced universe (Clark, 1918).

The classical economists largely agree on the fact that institutions affect economic performance, but they focus mainly on equilibrium points within a given institutional setting, rather than conjecturing whether this setting is also subject to change. As Van Den Bergh and Stagl (2003, p. 290) state:

> *"Orthodox economics lacks a positive theory of the manner in which institutions actually change and have changed during economic history."*

Traditional (non-evolutionary) economics views institutions as constraints and focuses on transaction costs as the main focal point. Within this line of thinking, institutional change is framed as a control problem rather than an endogenous phenomenon. Individuals are assumed to be rational and able to solve complex decision problems within a predetermined institutional environment. In this chapter we depart from that assumption and argue that transaction costs are actually the main driver behind institutional change. The field of 'New Institutional Economics' (NIE), that merged in the 1980s, departed from admitting that institutions affect outcomes but taking them as given and started to talk explicitly about the concept of institutions and institutional change. Thus far, institutional economics was mainly regarded as belonging to the history of economic thought. Scholars in the field of NIE made two propositions: first, that institutions matter, and second, that they can be analyzed using the tools of economic theory Williamson (2000).

The first proposition was shared by most economists, but the second proposition made them stand out from the 'old institutionalists' as well as the 'classical' or 'mainstream' economists. North (1990, p. 3) gave the most important impetus to this field by defining institutions as the humanly devised rules of the game, the definition that is still used by most institutional economists today. Although widely used and appreciated in economic science, this definition is still only a starting point and far removed from making institutions empirically measurable or describing mechanisms for institutional change. Another important contribution to the field was made by Williamson (2000) who proposed a typology of the different degrees of embeddedness of institutions and showed what this entailed for the speed of institutional change.

In this chapter we take a step back and not only talk about how and when institutions can change, but also on how the institutions that separate human societies from groups of animals initially emerged. Our contribution is threefold. First, we review theoretical and empirical work on

the endogenous development of political and economical institutions in societies. Second, we propose that these institutions emerged after the Neolithic transition, when societies changed from a hunter-gatherer lifestyle to an agricultural and sedentary lifestyle. The Neolithic Transition brought about technological changes that affected the population size that could be sustained. It thereby enabled, but also demanded, a more elaborate institutional setting than was the case in the kinship-based hunter-gatherer societies. We propose a theoretical model that links increases in technology and therefore in the size of the society to a growing latent demand for institutions. Under evolutionary pressure, societies that are able to establish these institutions will be favored by natural selection. To the best of our knowledge, there is currently no theoretical model that links these technological changes and population dynamics to the emergence of institutions. Third, we use a flowchart to illustrate that institutional development is a complex trajectory and each stage may need different models and theories to describe it. While the mechanism that drives the growing demand for institutions is universal, the actual outcome of the process of institutional formation greatly differs between societies because of the different obstacles that need to be overcome to sustain cooperation. Modeling ex genesis institutional emergence is an immensely complex task, this chapter cannot claim to have fulfilled that task fully. Rather, we put forward the idea that modeling this development requires the three steps identified above. In this chapter we will not test any formal hypotheses, due to lack of data on this matter. Therefore it suffices to put forward this framework for future studies, rather than fully modeling the complex process of institutional development.

When we talk about 'ex genesis' emergence of institutions in early agricultural societies, we do not mean to claim that the preceding hunter gatherer societies did not have any institutions. Each society is embedded in an institutional setting. As Hodgson (2006) describes it; in the original, hypothetical 'state of nature' from which institutions are said to have emerged, a number of rules, institutions and norms have already (unavoidably) been assumed. However, as transactions in hunter-gatherer societies were less sequential and as most members in those society were tied based on kinship, the need for market-supporting and constraining institutions was smaller. Trust played a more important role in these transactions. The Neolithic Revolution is seen as an important turning point in the nature and breadth of institutions, making them more comparable to how institutions are regarded today. One could also distinguish between formal and informal institutions, but this is described by Hodgson (2006) as a false

dichotomy, because the 'formal' institutions always depend on inexplicit norms and unwritten rules in order to operate. Formal and informal institutions jointly develop throughout history, although we can argue that pre-Neolithic societies functioned primarily with informal institutions.

We thus take the Neolithic Revolution as a starting point to describe the process of institutional emergence and change. We argue that the transition from hunter-gatherer to agricultural societies is crucial in understanding how institutions are shaped today for several reasons. First, Hamilton (2008) describes that culture has ceremonial and technological aspects and that technological progress is what leads to institutional change, as the 'old' institutions are no longer suitable to deal with the new possibilities. Technology has been incrementally changing throughout the history of mankind and is the driving mechanism behind other changes in population, welfare and 'culture'. Naturally, technology is endogenous to economic development as well, but throughout this chapter we view technology as being moved only through exogenous shocks. The change from hunting and gathering to active food production brought about an acceleration in technological progress. Second, Galor (2011) explains how the Industrial Revolution triggered the 'take-off' to sustained economic growth by changing the returns to formal education, but that the processes that led up to the Industrial Revolution had their origin in the Neolithic Revolution. The timing of this revolution has a strong correlation to economic performance today. We build on this model and add to this explanation the development of institutions because of a growing population. We feel that institutions are implicitly assumed in Endogenous Growth Theory and that many of the processed required for the transition to the Modern Growth Regime would not be possible without the simultaneous development of certain institutions. Third, the technological changes brought about by the Neolithic Revolution not only asked for accompanying changes in institutions, but also enabled these changes by providing new ways of specialization and division of labour that led to a reserve supply of resources.

This chapter is structured in the following way. First we set the stage and describe the institutional setting in Pre-Neolithic societies in section 2.2. In section 2.3 we explain how the Neolithic Revolution caused a growing demand for institutions. We put forward a representative-agent model in which evolution selects on population size and in which the number of transactions with strangers in society is a driving force for the investment in institutions. Section 2.4 and 2.5 give an overview of the necessary and sufficient conditions that should be met in order for societies to succeed in

setting up these institutions[1] In section 2.6 we elaborate on the factors that contribute to the stability or collapse of these institutional structures. In section 2.7 we support the theoretical foundations with historical anecdotal evidence of societies that either successfully established designed institutions or failed to do so. Finally, in the concluding section 2.8 we structure our findings in a flowchart, showing each stage of institutional development, the hurdles that need to be overcome and the type of modeling that in our opinion is best suited to describe that stage.

Many opportunities still exist for future research to formalize institutional emergence and institutional change in an all-encompassing model. As Van Den Bergh and Stagl (2003) state, when explaining the dynamics of behavior and institutions, various feedback mechanisms need to be represented, resulting from multiple interacting layers of genes, individuals, groups and institutions, as well as norms and meta-norms. This leads to a multi-level and multilayered system. Hence, a simple evolutionary representative agent model is inappropriate to capture the entire process of institutional development. Our contribution lies in proposing one mechanism that could set in motion this whole chain of events and that makes hunter-gatherer tribes substantially different —in terms of their institutional structure —from the more complex societies that came afterwards. To gain a full understanding of the pattern of institutions across societies today, our model could be complemented by insights from behavioral economics and (cooperative) game theory in future research.

2.2 Pre-Neolithic institutions

To clarify what is meant by the emergence of institutions in isolated societies we need to explain what kind of institutional setting was present before. There is no such thing as an institution-free society, as human behavior is always constrained and guided by some code of conduct. The main difference between the institutional setting before the Neolithic Revolution and the type of institutions that emerged afterwards and that we are analyzing can be made along several lines: 'pragmatic' (designed) versus 'organic' (self-organizing) institutions (Menger, 1871; Greif, 2005); formal

[1] We are aware that the terms necessity and sufficiency have very strong implications in the field of logic that we cannot always meet when describing real world societies. By and large, all societies that have gone through the phases that we describe in these chapters will have developed some system of designed institutions, but the degree of complexity and the actual implementation may differ. We use these terms here for lack of a better way to distinguish between these two phases.

versus informal institutions; the institutional environment versus institutional arrangements; (Davis et al., 1971), agent intensive institutions versus self-sustaining or agent-insensitive institutions (Hodgson, 2006) and probably many more.

One of the most primitive, human-specific institutions is most likely spoken language, as it is self-sustaining and it does not rely on other institutions for its enforcement, while most other institutions can be found to rely to some extent on language for their enforcement (Searle, 1995). It is widely believed that language originated together with modern human behavior, but the details remain unclear. In any case it is difficult to imagine a society developing agriculture without the use of language, so this primal institution must have been present to some degree at the time of the Neolithic Revolution. Furthermore, foraging groups did have common norms and constraints on behavior, such as food sharing practices, within the group. However, as there is no way of freeing resources in foraging societies[2], these norms or institutions have to be self-sustaining. Self-sustaining implies that it is in every individual's best interest to stick to the rules, so no external enforcement is necessary. These institutions are *organic* rather *designed*; as it is clear that they serve everyone's best interest, they arise quite spontaneously without the need for some members to give directions. Greif (2005) explains that contract-enforcement institutions emerge organically in the early stages of market development. These initial institutions influence which additional, designed (intentionally created) institutions the economic agents demand. Social capital or trust in these groups is high, as they are all bound by close kinship-ties. This ensures strong reciprocity. Transactions in the hunter-gatherer setting are simple; there are no lagged exchanged or intermediate parties, due to limited storage possibilities, so transaction costs are low. In the words of Greif (2005, p. 747):

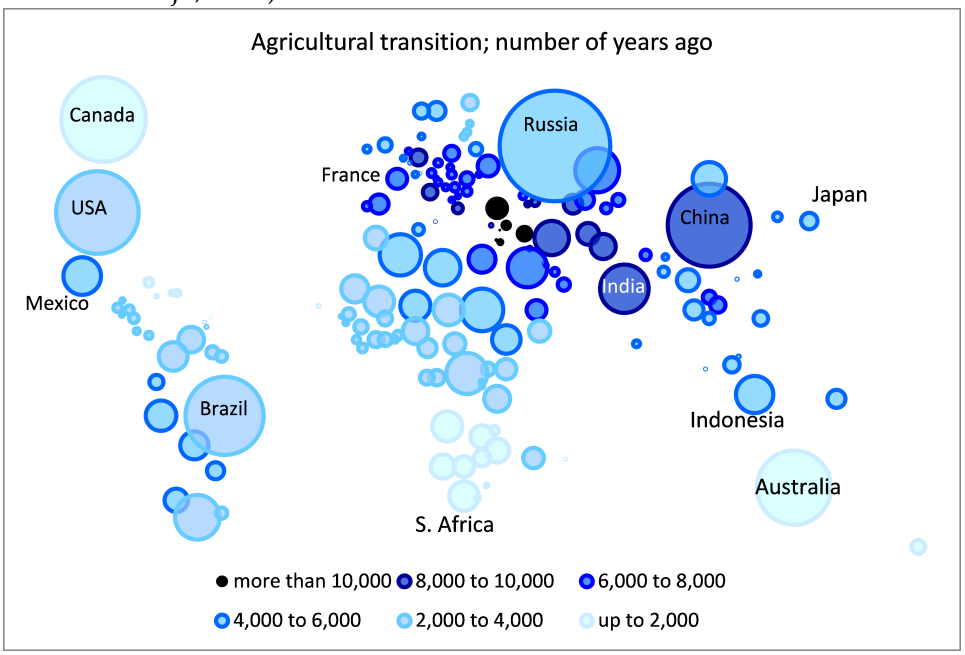
> *"In small social units, such as tribes, communities, and clans, the use of coercive power to capture another's property is likely to be effectively curtailed by the density of personal, social and economic ties and the relatively even distribution of coercive power. In larger social units, this is not the case, leading many, such as Hobbes, to argue that a state is required to protect rights and foster markets."*

Institutions are a way to economize on trust and to make transactions possible in 'mass, anonymous markets' (Shubik, 1999). As we will see in the

[2]Every able and fit member needed to contribute to the food-gathering process, and this food could also not be easily stored to distribute later on.

following sections, the Neolithic Revolution caused large changes in the way transactions were made and in the accompanying costs, leading to the changes in demand for more complex institutions and in some cases enabling the development of necessary and sufficient conditions to establish these designed institutions and, eventually, states. Societies that made the transition to agriculture early on in history thus had a head start when it comes to the development of market-supporting institutions, although this is by no means a linear trajectory. Both among the places where agriculture rose independently as well as in the rest of the world where it diffused gradually, there are considerable variations in the timing of the Neolithic Revolution (see Figure 2.2.1), shown to have persistent effects on outcomes today (Galor and Moav, 2002; Olsson and Hibbs, 2005).

Figure 2.2.1: Timing of Neolithic Revolution, number of years ago (source: Van Marrewijk, 2016)



In the next section we will argue that the Neolithic Revolution was the driver for a rising demand for more complex, non self-enforcing institutions.

2.3 The Neolithic Revolution and demand for institutions

2.3.1 Transaction costs

In his famous essay 'The Problem of Social Cost', Ronald Coase pointed out that the neoclassical model only holds under the restrictive assumption of zero transaction costs. As soon as one allows for positive transaction costs, institutions start to matter; but these are not present in the world of neoclassical model (Coase, 1960). Since Adam Smith economists have realized that gains from trade are the key to wealth and progress, but only in the last few decades it was understood that this exchange progress is also costly in itself. If transaction costs were at all considered before the 1970s, it was usually assumed that they were neutral or independent of the environment or society under study. North and Thomas (1973) place the notion of transactions costs at the heart of economic history and develop the idea that institutions evolve to economize on these transactions costs and in turn have an impact on economic outcomes. Transaction costs are the obstacles to realizing the possible gains from trade and therefore key to explaining the performance of economies. Transaction cost economics holds that all but the simplest transactions require some kind of mechanism to protect the transacting parties from the hazards associated with exchange outside of the circle of kin (Klein, 1998).

As North (1989) describes, in personal exchange where individuals engage in repeated dealings with each other or otherwise have large personal knowledge about the attributes and characteristics of the other party, transaction costs are very low. Cheating, shirking and opportunism are present in modern industrial organization theory but limited or absent in such a society with a dense social network of interaction, because the payoffs to this behavior are not high. In such a society, formal institutions do not exist and norms are not written down. North (1989) explains, however, that although transaction costs are low in this type of society, production costs are high due to a lack of specialization and division of labour. At the other end of the spectrum is a world of pure impersonal exchange. In this case, transaction costs are high because measuring the value of what is exchanged and enforcing the terms of exchange is difficult. This means that gains of engaging in shirking, opportunism and other forms of cheating are high. In order to prevent these activities elaborate institutional structures must be devised that constrain the participants and minimize these costs. This leads to the well-specified and well-enforced property rights,

contracts and monitoring systems that are omnipresent in modern Western societies. Naturally, differences among societies regarding initial levels of trust, the tendency to cheat or defect and the levels of transaction costs will still be present. Fisman and Khanna (1999) show that differences in trust do not rely only on historical determinants but also on the ease of information flow in societies.

The central idea in this section is that the Neolithic Revolution induced changes in technology and population that increased the transaction costs within societies. The Neolithic Revolution created a larger food surplus, thereby allowing for larger population densities within one community and leading to a more complex society, surpassing the kinship-based groups that were prevalent before. Society then finds itself in the position described by North (1989, p. 1324):

> *"The structural forms of human interaction that characterize societies are a combination of rules, enforcement features, and norms of behavior. Until we learn what these are, the costs of transacting are high. [...] The function of institutions is to provide certainty in human interaction, and this is accomplished by the inherent features of rules and norms."*

2.3.2 The stranger-effect model

The first stage in modeling the emergence of complex institutions is establishing a demand for these institutions. The Neolithic Revolution moves societies past the threshold from which point onwards transaction costs make investing in institutions that lower these costs worthwhile. In a representative agent model we show how the transition from hunting and gathering to agriculture as the main source of food production leads to larger societies and eventually a demand for designed institutions that require investment and reinforcement. We explain the underlying assumptions of this model and then reveal how evolution will favor those societies that are able to maintain a collective investment in institutions that lower transaction costs.

In hunter-gatherer communities, people live in small tribes, where members are bound by close connections (kinship) so the chances of interacting with strangers are near zero. At a certain point, societies make the transition to a more sedentary agricultural lifestyle. Although this process is referred to as the Neolithic Revolution, it was actually a slow and gradual process, especially since agricultural productivity was initially low and individual benefits of switching to agriculture were lower than the benefits of

remaining a hunter-gatherer. Skeletal evidence of early agricultural communities shows marks of poorer nutrition compared to forager communities in similar areas at similar dates(Armelagos and Cohen, 1984; Steckel and Rose, 2002). The paradox of why agriculture was then still adopted has been widely researched, and explanations include: other benefits besides nutritional value, a deterioration of the hunting environment, population growth outweighing the improvements in agricultural productivity or benefits at the group level (Weisdorf, 2009; Rowthorn and Seabright, 2010; Bowles, 2011). In any case, although initially many tribes combined hunting and cultivating food and although there were probably some societies who regressed back to a foraging lifestyle, the term Neolithic Revolution was still aptly chosen because the changes it brought about in the way people lived were undeniably dramatic and far-reaching. Bar-Yosef (1998) confirms that as hunter-gatherers became cultivators, social, technical and economic changes had to naturally follow.

Over time, through learning and doing, technological advancements built on each other and led to increases in agricultural productivity, allowing societies to build sedentary villages. This led to increases in population size and density for several reasons, one of the main reasons being that it was now possible for women to nurse more than two children at once. In a foraging community, this was impossible since the woman had to be highly mobile. Therefore population growth was virtually zero, as evidenced by stable tribe sizes (Birdsell, 1957). Population density fluctuated somewhat, linked to environmental factors such as rainfall, but no sustained population growth was observed. According to Blanton (1975), a band cannot grow in complexity (towards a chiefdom and eventually a state) unless it finds ways to produce and consume more energy. The demographic transition theory assumes that human population grows until the limits of its resources (Malthus, 1798; Galor, 2011).

The Neolithic Revolution brought about a large increase in the human population and its reproductive success, although the lifespan of most individuals was initially reduced (McFarland, 2015).[3] Livi-Bacci (2012) shows that while hunter-gatherer and, depending on the climatic conditions, had population densities between 0.1 and 1 person per square kilometer, agriculture permits population density to rise to 40 to 60 persons per square kilometer. This leads to the first assumption of our theoretical model:

[3]This was due to the aforementioned lower nutritional value of the first agricultural crops, but also because sedentary humans are more at risk to several enemies, human but also in the form of rats and mice and other disease-bearing creatures. Sedentary lifestyle increases the chance of cross-infection of these diseases.

Assumption 1. *As a community make the transition from a hunter-gatherer lifestyle to a sedentary agricultural lifestyle, its size and population density increases.*

Birdsell (1957) found the level of interaction between different foraging groups to be very low. Distance is the main isolating mechanism, hence the frequency of interacting with other groups is an inverse function of how removed or isolated the band or tribe is. But even for tribes living quite closely together intertribal marriages were lower than expected, showing evidence of cultural barriers. Therefore there was a high degree of inbreeding within the tribes and all members were closely related through kinship ties (Eller et al., 2009). We can consider Pre-Neolithic hunter-gatherer tribes as isolated societies. As a result of the increase in community size after the Neolithic Revolution, the average kinship distance between any two members in the community increased, until the point where kinship bonds become hard to trace.

Humans can also form bonds with non-kin in their community or village, such as friends, neighbours or fixed trading partners. This is in fact a main distinguishing feature between humans and most other animal species. The evolutionary growth in brain size is linked mainly to social, not intellectual development (Flinn et al., 2005). Maintaining the stability of these larger social networks puts a time demand on humans, therefore the evolution of large groups in the human lineage depends on the development of a more efficient method for time-sharing the processes of social bonding (Hill and Dunbar, 2003). Spoken language uniquely fulfills this requirement. However, the size that this network can obtain still has a limit, which may change over the course of time as other methods of communication are developed. Dunbar (1993) estimates this limit to be 150 individuals. It follows that as the community size surpasses this limit, one can no longer be familiar with all members of the community. The social network of the community may build on common lineage, but it is no longer comprised of direct relatives. An early institutions that led humans to start organizing themselves in bands, based on common lineage, while animals and humans living in tribal societies do not extend this familiarity to distant relatives, is religion. More specifically, it follows from the worship of dead ancestors (Fukuyama, 2011). This leads to our second assumption:

Assumption 2. *Due to the increase in community size, the chance of interacting with a 'stranger' increases once societies become sedentary.*[4]

[4]Stranger is defined here as someone not related to you through identifiable, direct kinship bonds, belonging to your social group.

We know that the extent of reciprocity is related to kinship distance in hunter-gatherer societies. Once tribes settle down, community size increases and this simple reciprocity breaks down (McFarland, 2015) and chances of cheating or shirking behavior are higher, increasing the average transaction costs between members of the community. At the same time, the possible gains from transactions are higher because different members of the community may specialize in different ways of producing food now. In our model, institutions are understood as the set of rules that facilitate the communication and transactions between 'strangers' (either bureaucracy or market), lowering the transaction costs. This leads to the following assumption:

Assumption 3. *As the chances of dealing with a stranger grow, the trust and reciprocity between members are lower and transaction costs increase, lowering consumption for given production levels.*

It is clear that at a certain point, transaction costs will become so high that they outweigh the benefits of making the transaction at all. While formal institutions can never fully substitute for social capital or trust (Putnam, 1993; Cook, 2001), they can function as a supplement, extending the trust that is present in the community to members outside of the direct social circle and thereby lowering transaction costs. We argue that the type of institutions that are needed to facilitate these market transactions go beyond the self-sustaining, organic institutions that were already present in hunter-gatherer societies and require collective action and active maintenance by members of the population.

The community needs to invest a share of its resources in institution-building and reinforcement. This can mean several things: they could free time from food-production to find and punish freeriders or they can vote to give power to a government or collective decision body that will set up the optimal level of institutions (and monitoring) for them, using the revenues of an income tax. The exact construction is not important here, but we assume that institution-development is always costly and requires an active and collective choice to invest in them. As Witt (1989, p. 156) states:

> *"[...] certain institutions cannot be expected to emerge in the way assumed by the alternative conjecture: the interest pursued by the individuals involved do not necessarily lead them to 'spontaneously' create or support an institution. In this case, for the institution to actually emerge, some kind of collective action would be required."*

As institution-building requires resources, we model this by requiring that a share of total production is collected to finance the initial investment and its maintenance.

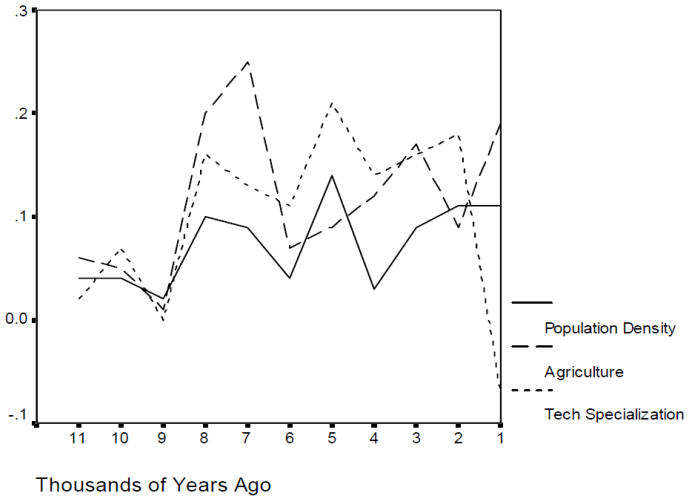
Assumption 4. *The community can reduce the negative effect of population size on transaction costs by collectively investing resources in setting up and maintaining institutions.*

Once established, these institutions lower the transaction costs of (market) interactions, leading to higher income from production. The community we consider still finds itself in a Malthusian trap, so these economic gains will be directly translated into larger population size, again raising the transaction costs and thereby the demand for institutions in a double feedback loop. Ashraf and Galor (2011) show significant positive effects of land productivity and the technological level on population density in 1500 A.D., 1000 A.D. and 1 A.D., while the effects of land productivity and technology on income per capita in these periods do not differ significantly from zero, when looking at long term effects. Eventually technological progress will induce a quantity-quality trade-off that leads to the demographic transition, this is explained in the Unified Growth Theory as described by Galor (2011). Our assumptions are supported by historical data, collected by Peregrine (2003).

Figure 2.3.1 reveals that since the Neolithic Revolution, agriculture, population density and technological specialization coevolve together. [5] There is no prime mover observable until 6000 years ago, when agriculture starts to lag behind population density and technological specialization. It is hard to identify causality and this is also not the main aim of this chapter. We take the inter-correlation between these three variables as given and show how the feedback loop between agricultural productivity and population density paves the way for but also necessitates the development of complex, designed institutions. We use these assumptions in a simple Malthusian model in which fertility depends on per capita income to show how collective investment in transaction costs reducing institutions emerges endogenously in the course of social evolution.

[5] The authors first-differenced these time-series data in order to de-trend the series.

Figure 2.3.1: Coevolution of population density, agriculture and technological specialization (source: Peregrine, 2003)



2.3.3 The intergenerational problem

Our model builds on Galor's model of a community in the Malthusian state (Galor, 2011) and describes an overlapping-generations economy in which activity extends over infinite discrete time. In every period, individuals produce a single homogeneous commodity, using land and labour as inputs. The supply of land is exogenous and fixed over time, but the supply of labour depends on the decisions made by households in the preceding period on the number of children to raise. Technology is exogenous in the model.
Households face the same maximization problem every period:

$$\underset{c_t, n_t}{\text{maximize: }} U_t = {n_t}^{\gamma} {c_t}^{1-\gamma}$$

$$\text{s.t.} (1-\tau_t) z_t = \rho n_t + c_t$$

Where U is utility which is derived from n, the number of offspring and c, consumption where $0 < \gamma < 1$ is the elasticity for offspring in utility.

The budget constraint states that after-tax disposable income per person, $(1-\tau_t)z_t$, is spent on consumption and child rearing, where we normalised the price of consumption to 1, ρ is the fixed cost of raising a child and $0 < \tau < 1$ is the tax rate.[6] It is straightforward to show that (see Appendix) households maximize their utility by setting the number of children and the level of consumption each period to:[7]

$$c_t* = (1-\tau_t)(1-\gamma)z_t \tag{2.3.1}$$

$$n_t* = \frac{\gamma}{\rho}(1-\tau_t)z_t \tag{2.3.2}$$

Optimal consumption depends positively on (after tax) income per capita, which individuals take as given. Finally, it depends negatively on γ, the exponential weight on children in utility. The optimal number of children also depends positively on (after tax) income per capita, negatively on the costs of raising children and positively on γ, the exponential weight on children in utility. Investing in institutions does not lead to higher utility at the household level as institutional quality is a pure public good. However, a household can consume more and have more children if after tax income is higher. This establishes the Malthusian link between (after tax per capita) income and population growth.

2.3.4 Production and population

We describe the emergence of institutions when a community is still isolated from contact, so the only way in which the community grows is through reproduction. Normalizing initial population to 1, population dynamics are then described by the following function:

$$L_{t+1} = n_t L_t \quad \text{where:} \quad L_0 = 1 \tag{2.3.3}$$

The community consumes a homogeneous commodity Y (Food), which is produced every period with inelastically supplied labour, L, and land ,X, using technology A with a Cobb-Douglas technology according to:

$$Y_t = (AX)^{\alpha} L_t^{1-\alpha} \tag{2.3.4}$$

[6] To be very precise, it is the share of disposable income collected and allocated to building and maintaining institutions in the community. We will refer to this as the tax rate from here on.

[7] Throughout the chapter, asterisks are used to denote optimal levels.

where $0 < \alpha < 1$ is the output elasticity of the land-technology complex. However, due to transaction costs this output is not fully available to households and the community is left with disposable income:

$$Z_t \equiv \theta_t Y_t = \theta_t (AX)^{\alpha} L_t^{1-\alpha} \tag{2.3.5}$$

where $0 < 1 - \theta < 1$ captures the share of output lost to transaction costs. The amount of land is assumed to be fixed and the level of technology is taken as a given.[8] What is added here compared to standard growth models is θ, which represents transaction costs that we will endogenize below.

2.3.5 Transaction costs and institutions

The community now produces Y_t and every household can spend disposable income $z_t \equiv Z_t / L_t$ on children and consumption, where $Z_t = \theta_t Y_t$. If a community, however, somehow manages to collect τZ_t to invest in institutions, the transaction costs can be driven down. Moreover, we already argued that transaction costs will rise in the size of the population. To reflect these assumptions we must look for a function for θ that meets the following conditions:

- $\theta_1(\tau_t Z_t, L_t) > 0$
- $\theta_{11}(\tau_t Z_t, L_t) < 0$
- $\theta_2(\tau_t Z_t, L_t) < 0$
- $\theta_{22}(\tau_t Z_t, L_t) > 0$
- $\theta(\tau_t Z_t, L_t) = 1$ for all $L_t \leq \beta$
- $\theta(\tau_t Z_t, L_t) \leq 1$ for all $L_t > \beta$

The first two conditions state that θ is positive and concave in the resources invested in institutions $(\tau_t Z_t)$. The second pair of restrictions implies that

[8] As technology is exogenous in our model, we do not specify how it evolves. But one could imagine that technology shocks are a by-product of production, through learning by doing and specialization. One simplification we have made here is that foraging and agriculture follow the same production function. This can be justified because the transition from foraging to agriculture was not abrupt; both methods of food production coexisted for many centuries and hunting is still a source of food production today. See also Ashraf and Michalopoulos (2015) for a much richer model that captures the transition.

θ is negative and convex in population size to capture the "stranger-effect". The final pair of conditions implies that up to a population level β there are no transaction costs. We can imagine this as the size limit to community as a kinship-based community or the limit of a personal network, which is estimated to be around 150 by Dunbar (1993). Above this threshold, some level of inefficiency will exist due to increased transaction costs, but this negative effect can be reduced by investing in institutions. The community thus produces $Y_t = (AX)^{\alpha} L_t^{1-\alpha}$ but can only consume $Z_t = \theta(AX)^{\alpha} L_t^{1-\alpha}$, and transaction costs depend positive convex on population size and negative concave on the collective investment in institutions. Finally we impose:

$$\lim \theta_t(\tau_t Z_t, L_t) = 0$$
$$L_t \to \infty$$

to ensure that transaction costs never reach 100% of total output for finite populations.

2.3.6 Social evolution

Although in our model we now proceed to compute the optimal tax rate τ, we do not in any way wish to imply that there is indeed some central planner or external power that would compute and enforce such a tax. Instead we argue that evolution ultimately favors societies that manage to sustain a collective investment in institutions that maximizes total population in steady state $L*$. Therefore, societies that manage to get τ_t close to its optimal level by whatever means will have higher disposable income, higher n_t and consequently higher population growth and a larger steady state population size. If we choose τ to maximize steady state population size we obtain:

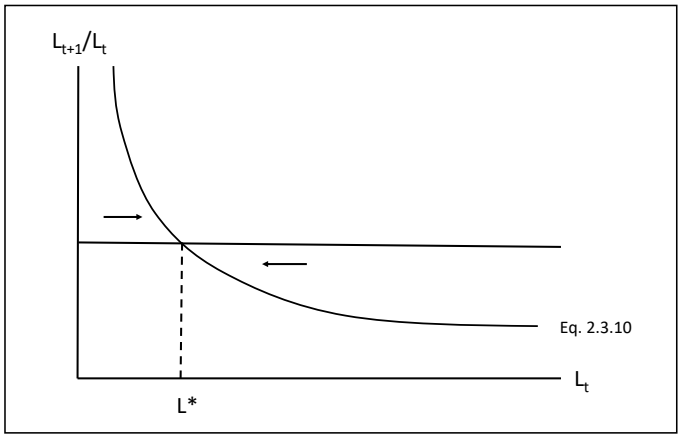
$$\tau^* = \underset{\tau_t}{\text{argmax:}} \bar{L}^*(\tau^*) \tag{2.3.6}$$

To obtain steady state population as a function of τ, we first plug in equation 2.3.2 into equation 2.3.3 to obtain:

$$L_{t+1} = \frac{\gamma}{\rho}(1-\tau_t) z_t L_t = \frac{\gamma}{\rho}(1-\tau_t) Z_t = \frac{\gamma}{\rho}(1-\tau_t)\theta_t (AX)^{\alpha} L_t^{1-\alpha} \tag{2.3.7}$$

We can now plot L_{t+1} as a function of L_t to investigate the dynamics in the model when we specify θ_t. A functional form that satisfies the conditions we have imposed would be:

Figure 2.3.2: Phase Diagram for Population



$$\theta_t = \left(\frac{\tau_t Z_t}{L_t}\right)^{\xi} = \left(\frac{\tau_t \theta_t (AX)^{\alpha} L_t^{1-\alpha}}{L_t}\right)^{\xi} \tag{2.3.8}$$

with $0 < \xi < 1$ such that after solving for θ_t and simplifying we obtain:

$$\theta_t = \left(\tau_t (AX)^{\alpha} L_t^{-\alpha}\right)^{\frac{\xi}{1-\xi}} \tag{2.3.9}$$

The equation driving population dynamics then simplifies to:

$$L_{t+1} = \frac{\gamma}{\rho}(1-\tau_t)\left(\tau_t (AX)^{\alpha} L_t^{-\alpha}\right)^{\frac{\xi}{1-\xi}} (AX)^{\alpha} L_t^{1-\alpha} \tag{2.3.10}$$

Dividing both sides by L_t gives an expression that is strictly negative and convex in L_t:

$$\frac{L_{t+1}}{L_t} = \frac{\gamma}{\rho}(1-\tau_t)\left(\tau_t (AX)^{\alpha} L_t^{-\alpha}\right)^{\frac{\xi}{1-\xi}} (AX)^{\alpha} L_t^{-\alpha} \tag{2.3.11}$$

Drawing this line in a graph with $\frac{L_{t+1}}{L_t}$ on the vertical and L_t on the horizontal axis gives us Figure 2.3.2 below in which we can analyse the dynamics in our model for any level of $0 < \tau < 1$.

The steady state is found where the downward sloping line intersects 1 on the vertical axis as where we have $L_{t+1} = L_t = L^*$. The same condition in equation 2.3.10 implies that steady state population size is given by:

$$L^* = AX\left(\frac{\gamma}{\rho}\right)^{\frac{1-\xi}{\alpha}}(1-\tau)^{\frac{1-\xi}{\alpha}}\tau^{\frac{\xi}{\alpha}} \tag{2.3.12}$$

which is a function of parameters, exogenous variables and our tax rate only. Now we can take the derivative of this expression with respect to τ and set it equal to zero to obtain the taxation level that maximises steady state population size. After some algebra (see Appendix) we obtain:

$$\tau^* = \xi \tag{2.3.13}$$

This result tells us that a community that grows beyond a population size of β should invest a share ξ in building institutions to reduce transaction costs. Obviously the optimal tax rate is bigger and in fact proportional and by our specification even identical to the elasticity of θ with respect to such investments. The value of ξ and the exact specification of θ are very context and time specific. In times of war and civil unrest the collective investments in lowering transaction costs can be very unproductive. One could even interpret such historical events as the collapse of the Roman empire as a shock to ξ, such that a breakdown of the link between investment and transaction costs causes societies to fall back to lower population size and collective action levels. We do not wish to overextend the implications of this very simple model though. For now it is sufficient that we have illustrated how population size, density and growth drive a (latent) demand for institutions in society.

2.3.7 Comparative statics

From equation 2.3.12 we can derive some interesting comparative statics. The steady state population is proportional to A and to X, reflecting the standard Malthusian dynamics that predict all technical progress and extension of the available land per capita will simply result in a larger population. We also see that L^* is positive in γ and negative in ρ as $(1-\xi)/\alpha$ is strictly positive. This makes sense because ρ is the cost of raising children and γ is the exponential weight on children in household utility. They thus should have opposing effects. The effect of α, the output elasticity of land and technology (i.e. not labour) is strictly negative, which is also intuitive. A lower output elasticity for labour (higher α) implies larger populations have a smaller positive feedback effect through higher income and children.

The most important conclusion from our model, however, is that for

societies that are able to successfully invest the optimal share of total output in institutions, steady state population will be higher than for those that cannot. Note here that societies can both over- and undershoot the optimum tax rate and suffer the consequences. In that sense societies can have too weak and too strong central taxation. Our model also implies that, once a community's population surpasses the threshold β, it needs to find a way to allocate resources to institution building to remain competitive. It is easy to show that when random but strictly positive shocks increase A (or X) over time, surpassing the population threshold is only a matter of time. Galor (2011) assumed such a gradual but random process inevitably triggers the quantity-quality trade-off and causes technological improvements to translate into per capita income gains rather than a growing population. I would propose, however that such a trigger can only arise once institutions have been brought to a level that the gains from investing in quality through e.g. education can actually be appropriated. As education favors specialists, this involves institutions that reduce transaction costs, such that specialization can actually occur. In that sense our model describes the dynamics prior to the demographic transition and explains how the assumed institutional preconditions endogenously come about.

2.3.8 Evolution as benevolent planner

Let us briefly return to the issue of the optimal tax rate as it may seem contradictory to some to use an optimization model when talking about evolutionary institutional change. As stated above, we do not argue that in reality a benevolent planner ever existed. However, carrying out this maximization we show that for the community as a whole, there are gains to be made from collecting taxes and investing these in setting up institutions that reduce transaction costs. These collective gains do not translate into household gains directly, so not all communities will be able to overcome the Prisoner's Dilemma type of coordination problems that arise, as will be discussed in the following sections. However, evolutionary forces do give an advantage to societies that are able to collect taxes close to the optimal level and thereby increase the total (economic) size of their community. Therefore these societies are more likely to survive and ultimately come to predominate.[9] Levine and Modica (2013) and Turchin et al. (2013)

[9] In an evolutionary sense, this means that random variation (mutations) introduces different societies with different tax rates embedded in their social contracts into an evolutionary selection environment, where the variations that are closest to their optimum are the ones that will win the evolutionary contest for dominance.

both show that a level of competition or war between societies is required in order to ensure that the community-level advantages of setting up costly 'ultrasocial institutions' are high enough to overcome the private incentives to refrain from doing so. Societies with greater control and coherence over their large populations will thus simply outcompete those that lack such traits. Levine and Modica (2013) show that when war as the main evolutionary force is taken into account, the long-run favors 'not a large population at the level of subsistence, nor yet institutions that maximize welfare or per capita output, but rather institutions that generate large amounts of free resources and directs these towards state power'. Of course we know from public sector economics that there is an optimal tax rate that maximises this flow of resources to the collective.

The free collective resources increase in per capita income and population, so evolution favors large rich societies. In our simple Malthusian models such societies are the ones with the largest steady state populations. How 'evolution' actually selects those societies is another subject that is beyond the scope of our study.[10] It could be that the victorious society imposes its culture on the defeated (by religious conversion, linguistic assimilation, or replacing literate elites of the defeated group) or the physical removal of households that operate under a different social contract (e.g., genocide, expulsion). As these institutions and traits are culturally rather than genetically transmitted, the latter is not essential and historically probably less important (Turchin et al., 2013). In chapter 3 we will return to the issue of the diffusion of growth enhancing institutions. For now it is clear that those societies with evolutionary advantages will increase in importance and therefore the institutions will gain a foothold. Silverberg et al. (1995) also explain that species with above-average fitness will expand in relative importance, and therefore evolutionary models of economic growth should take into account population dynamics.

Finally, our model predicts that countries that experienced the Neolithic Transition earlier in history will have had a head start in setting up the first set of designed institutions that helped reduce transaction costs and consequently caused early agricultural societies to grow in size and outcompete societies with suboptimal levels of investment in institutions. As such it serves to illustrate the mechanism through which the Neolithic Transition, through an increase in population size that was a consequence of technology improving food-production, increased transaction costs and led to a

[10] However, size of the population (and armies) clearly mattered in the Neolithic Revolution.

demand by the community for setting up institutions that proved essential and powerful drivers for future growth. We do not show how taxes are collected and public investment in institutions is made. We certainly do not wish claim that a benevolent government appeared out of nowhere and acted in the best interest of society. Even after having obtained or been given such authority, actual kings and emperors typically had no incentive to do so. In fact, we immediately concede that many societies, even today, fail to maintain the optimal level of investment in institutions, even though there are collective gains to be made if they would be able to do so. But such issues of political economy are less easily modeled and therefore best left for discussion in the next sections.

2.4 Necessary conditions for establishing designed institutions

We showed how the Neolithic Revolution increased demand for transaction-cost lowering institutions that require active investment and reinforcement. However, the emergence of this demand does not automatically lead to the successful implementation of these institutions. In this section we elaborate on the necessary conditions for the emergence of transaction-cost lowering institutions. The key variable that differentiates societies here is the extent to which specialization and division of labour can be achieved. Naturally, some degree of specialization already existed in Pre-Neolithic times, otherwise the costly transactions that we modeled in the previous section which led to the demand for institutions to alleviate these costs, would not be required. However, in order for institutions to emerge, this specialization needs to be taken to another level. Technology allows societies to store the food they cultivated, to a larger extent than their foraging counterparts did. In order to establish an authority that can collect taxes and execute the establishment and enforcement of institutions the protection of stored food reserves is required. A non-food-producing elite with (to some extent) the monopoly of violence is required. This leads to the beginning of what Diamond (1997) calls a *kleptocracy*. The transition to agriculture ensured that enough food could be harvested in some seasons to sustain the people in times when no food would grow at all. This led to a stability of social life, critical for its material enlargement (McFarland, 2015).

Geographical boundaries also play a role in determining whether a society can set up designed institutions or not. Specialization requires a certain extent of the market. If the area that is inhabited by a society only lends

itself for a few specific, labour-intensive crops, it is unlikely that specialization and division of labour will reach the level necessary to set up complex institutions, to feed bureaucrats and an army. Geographic isolation, for instance due to rivers, mountains or dense forests, may therefore prevent societies from setting up these institutions. Having contact with other communities has two reinforcing effects: it increases the demand for institutions, as dealing with members from another society means less trust and higher transaction costs, but these transactions also enable a larger market, increasing the possibilities for specialization and gains from trade. If these gains are realized, the society can again set up more institutions and achieve a higher hierarchy or complexity. The link between technological development and social stratification has been confirmed by many authors. Peregrine shows in his Atlas of Cultural Evolution that the mean values of technological specialization and social stratification, measured at 1000-year intervals, inhibit very similar linear trends with R-squared values of 0.960 and 0.935, as shown in Figure 2.4.1 and 2.4.2.

Figure 2.4.1: Mean of technological specialization by date (source: Peregrine, 2003)

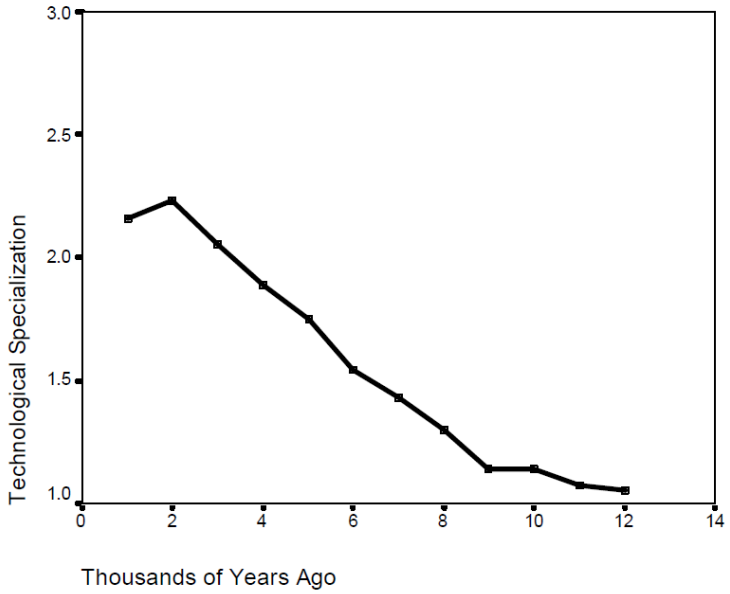
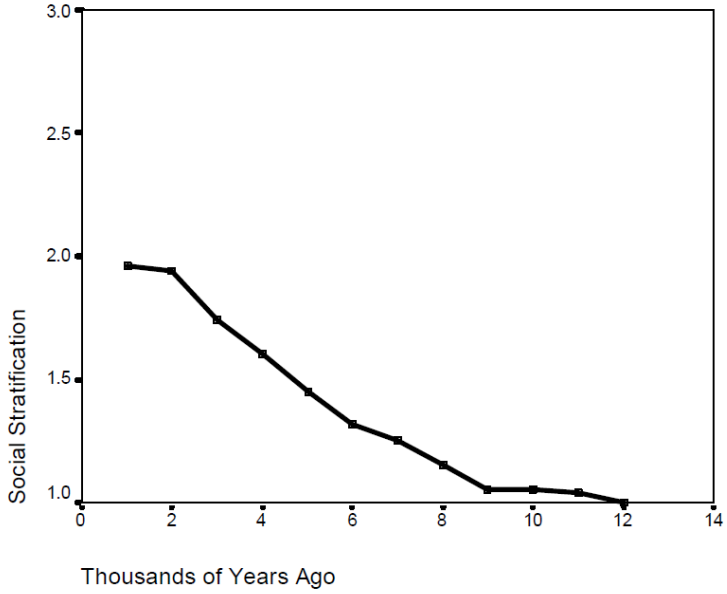


Figure 2.4.2: Mean of social stratification by date (source: Peregrine, 2003)



2.4.1 Specialization, bureaucracy and the nature of production

In the previous section we argued that a certain degree of specialization is a necessary condition for institutional development. Here we elaborate on how this specialization arose and which differences between early societies could be observed in this process of specialization and institutional or state development. The rise of civilization was first linked to specialization by the archeologist Childe (1936). He argued that the uneven and combined development of economic welfare and the extent of civilization depended on a surplus production over subsistence needs facilitated by agriculture, leading to a technical and social division of labour. An example of this is were the ruling classes in the Mesopotamia lowlands from 5500 B.C. onwards.

In these societies, part of the agricultural surplus was used to support full-time craft specialists, such as metalsmiths, providing these specialists with the supplies they needed. Soon this developed into an independent trading system between smiths and traders in ores, and the development of an unattached specialist. The specialists now had the possibility to innovate and extend their trade outside the control of bureaucracy. Food-producers and specialists were becoming more and more directly dependent on each other, leading to an increase in market transactions. Eventually another specialization emerged, a separation between direct producers and those who appropriated their goods and labour power, i.e. a distinction between manual and mental labour (Patterson, 2005). Naturally this came with the rise of a multitude of institutions, such as commodity and labour markets, the ability to buy on credit or provide loans and early accounting. This market development is described by e.g. Van Bavel (2010).

Although this sounds contradictory, this transition from a system where all exchange happens through lordly rulers and exploiters to a capitalist system with a producing class with direct possession of their means of subsistence, was what led to the development of more elaborate states. This is because the only objective of the elite, when positioned in a feudalist system, is to reproduce themselves by means of extracting surplus from the peasant producers (through coercion, taxes or unfree labour). With the existence of a productive class that produces for its own needs, more incentives for higher production and more exchange emerge, leading to higher surpluses, through which more specialization can occur and a larger non-food-producing class can be supported. This positive feedback loop creates the possibility but also the need for more and more complex states and institutions, as exchange becomes more sequential in nature, happens over

larger distances and in more valuable commodities. An obvious example is that to protect these commodities, a larger army or city watch is needed, leading to more hierarchy in the community.

Naturally, in some societies commodity production was less well developed and market exchange did not extend to all corners of everyday life. Marx (1857) distinguished seven modes of production: primitive communism, ancient, Asiatic, Germanic, feudal, capitalist and socialist. The succession of modes of production is not a chronological succession, nor should it be considered as an evolutionary process of one mode of production into another, but rather as a 'commentary on steps away from the original kinship-based community —that is, steps in the historical development of private property' (Patterson, 2005, p. 319).

In some societies, the balance of power was such that the elite could maintain their power for centuries and no significant exchange system could develop. Furthermore, hunter-gatherer communities could not create and store surplus production, while agricultural communities could. Depending on the type of agricultural production and the frequency with which crops can be harvested, storage and surpluses may be more or less predominant when comparing different agricultural communities. Differences in biogeographical conditions also led to the existence of two types of hunter-gatherer communities, 'upper' and 'lower' hunter-gatherers, as Testart et al. (1982) show. In some places, food was present all year round but scarce, leading to a lifestyle based on the immediate use of food resources, and a flexible and nomadic economy as can be seen in Bushmen and the Australian Aborigine. In other places, natural food resources were plentiful but seasonal, so they needed to be gathered *en masse* while available and then preserved and stored into a year-round food staple, leading to more sedentary foragers such as the Northwest Coast and California Indians.

A feature that hinders the development of more elaborate institutions is the specialization of the community in the production of goods that are essentially nontradable, or only tradable within a short time span and distance. Some crops lend themselves better for storage and transportation than others, and some societies were so isolated that trade with other communities, producing different commodities, was not possible at all. Looking at Madagascar, Stifel and Minten (2008) found a strong inverse relationship between agricultural productivity and isolation, condoning the Madagascans to fulltime production for own subsistence rather than creating a surplus that could be traded at a profit, leading to sustained poverty.

In conclusion, the development of technology for food-storage, specialization and the extent of exchange go hand in hand. When these three

aspects prevail, the necessary conditions for institutional development are met, but when either of those is underdeveloped or hindered by bureaucracy through an inert feudal system, the society is stuck with an underinvestment and underdevelopment of designed and more complex institutions.

2.5 Sufficient conditions for establishing designed institutions

In this section we elaborate on the factors that determine which societies meet the sufficient conditions for institutional setup and which societies do not[11]. This depends on the tools for collective action society is able to find, in the form of signals or screening tools identifying cooperating types and selection of community members, or in the form of diffusion agents or strong reciprocators. We give an outline of the problems that may prevent a society from actually establishing the institutions and ensuring their propagation, making use of Game Theory models, and describe each of the three aforementioned tools for collective action in detail.

Several problems may complicate the effective establishment of an institution. The first is the problem of coordination: how is the common will of the community actually directed towards a concrete initiative and how do members choose from possible options? The second problem is that of freeriding: even though it is clear that the establishment of the institution can only be achieved through collective action and that it is in the best interest of the group to invest in this, the individual has incentives to defect as the institution is a public good so no member can be excluded from its use once established. The final problem is something that has been extensively described in the field of agency theory as the principal-agent problem (Laffont and Martimort, 2009). Once society has appointed a non-food-producing elite to collect taxes and act on their behalf, what prevents this elite from acting in their own best interest rather than that of the whole population? Each of these problems has been extensively described in economic literature. However, history shows us societies have been able to overcome these problems and establish institutions. Therefore different so-

[11] As was explained before, in mathematics and logic, a sufficient condition means that the development is inevitable. We may not want to go that far here, as exceptions can always be found in social sciences and it is more a matter of the degree of development. For lack of a better word, we use this terminology here because it helps to clarify the distinction between the different types of hurdles in the two phases.

lutions to these coordination problems have been proposed.

Economists have been puzzled since decades by the question why sustained cooperation has been possible at all, while the individual incentives to defect are always present. As Leser and Menger (1883, p. 146) state:

> *"How can it be that institutions which serve the common welfare and are extremely significant for its development come into being without a common will directed toward establishing them?"*

Witt (1989) distinguishes three cases when it comes to the adoption or emergence of new institutions: in the first case, adopting the institution yields a net benefit to every agent from the very beginning and the further the institution propagates the more attractive it becomes, through increasing reputation or positive scale effects. This is the case for self-enforcing institutions like language. The second case is an institution with initially high net benefits but which decline with the number of adopters. In this case, the institution will be adopted but may be maintained after it is no longer optimal anymore. An example of this is the feudal system as discussed in the previous section: initially this enables some buildup of hierarchy and a non-producing class, but this class has no incentives to increase the extent of transactions and specialization and the accompanying market-supporting institutions. The final case is an institution for which the individual net benefit from adoption is increasing with the number of adopters but the initial benefit is below the threshold such that the institution cannot gain a foothold in the population. This leads to unstable propagation of the institution. An example of this is a state-centralized currency. Witt points to the existence of diffusion agents (organizers, leaders, agitators, political entrepreneurs) who gather a critical mass to ensure that the collective adoption will come about. We go deeper into the coordination problems and proposed solutions related to each type of institution.

First we focus on those institutions that seem to arise spontaneously and do not need be enforced. An example is the fact that early societies started to follow shared rules of linguistic communication. This is a case of a pure coordination games (Hodgson, 2009), with the payoffs as shown in table 2.1.

In game theory, this game is often explained using the decision to start driving left or right in a country. The players in table 2.1 are indifferent between both options, as long as they pick the same. How do members of society decide which option this will be? Members of society find a focal point, building on some common experience, historical or cultural, that is

Table 2.1: Coordination game —Driving left or right?

Player 1	Player 2		
		Left	Right
	Left	10,10	0,0
	Right	0,0	10,10

outside of the rules of the game (Boyer and Orléan, 1992; Klein, 1998; Binmore, 1994). If no focal points exist, someone could bring one into existence by making a 'dramatic suggestion' (Chwe, 2013). As explained before, before the Neolithic Revolution, societies were not institution-free, so the type of organic self-enforcing institutions that would arise depended on common experience. The designed, non self-enforcing institutions then build on the organic institutional structure that is already present. The 'dramatic suggestion' might be made by 'diffusion agents' who press for the execution of an institution in a certain direction (Witt, 1989).

However, in many cases some collective, costly action is needed to set up an institution or to enforce the members of society to follow the new rules. Although every member of society may agree that the establishment of and continuous investment in this institution is the optimal outcome for society as a whole, individual members would rather forgo this costly action and freeride on the efforts of others. The payoffs than look as in table 2.2.

Table 2.2: Public Good game —Prisoner's dilemma

Player 1	Player 2		
		Cooperate	Defect
	Cooperate	1,1	-1,2
	Defect	2,-1	0,0

The game is illustrated with two players, but the problem at hand is an n-person Prisoner's Dilemma. The unique Nash equilibrium is that all players defect and get zero payoffs. However, we know that public goods and institutions have been established throughout history. What can explain this? A first important factor is the presence of strong reciprocators. Strong reciprocators are individuals in society with a predisposition to cooperate with others and to punish those who violate the norms of cooperation, at personal cost, even when it is implausible to expect that these costs will be repaid either by others or at a later date (Gintis et al., 2003). Why would in-

dividuals want to punish even when they are not cheated themselves and when it decreases their own fitness? Several reasons include i) the desire to rinse the community from cheating types, as past cheating reveals bad types (Ghosh and Ray, 1996; Kranton, 1996), ii) reputation; the expectation that others will sanction an individual can be enough to motivate punishment (Greif, 1989; Fafchamps et al., 2004) or iii) an internalized sense of fairness that is shared among members of the social structure (Fehr and Fischbacher, 2004). These sanctions are not necessarily economic; they can also be social. This implies segregation of exchange into certain groups, with their own coordinates on what is considered fair and what is considered cheating. Strong reciprocity and multilateral punishment can actually sustain organic, non-designed institutions, but Greif (2005) claims that the efficiency of these institutions declines as populations and market grows.

The main difference with the designed institutions considered here is that individuals do not engage in multilateral punishment, but instead give up some share of their resources to a non-food-producing authority that has taken over this task for them. This mechanism relies heavily on a shared vision on what is considered cheating and how this should be punished; a vision that has developed through these organic institutions. Besides actual punishment, a mechanism that helps to keep players from defecting is the reputation mechanism. Although the increased size of the community after the Neolithic Revolution meant that kinship ties were less prevalent and the number of strangers increased, still most transactions would have a repetitive character. Being known as a cheater may decrease future payoffs to the extent that it is more beneficial to incur the costs of cooperation.

Wilson (1980) found that in the New England fresh-fish market, which consisted of a sizeable 1800 privately owned-vessels, reputation effects and mutual dependence still provided the most important enforcement mechanism for the smooth functioning of the market. A final mechanism that helps to establish continuous investment in institutions is that in cases where individuals do not have all the information, or where it is too costly to evaluate all the available information, the learning mechanism of copying the majority or averaging what the most important individuals are doing proves to be an effective way of aggregating information and extract adaptive behaviors to the new situation. This is called *conformist transmission* (Richerson and Henrich, 2009) and it implies that as soon as a certain proportion of society or some prestigious members of society support the institution, the rest of the community follows.

The final problem that communities may face are principal-agent problems. In section 2.3 we showed that those communities that succeed in

collecting the optimal tax rate are able to maximize the growth of society. However, in the real world an authority that is trusted to carry out this task may be hard to find, and other motivations may prevail. The high fixed costs of setting up designed institutions are not only due to technological determinants, there is also a mark-up over the 'production costs'. Generally a monopolistic organization takes upon itself the task of establishing the institution and enforcing the contracts, therefore those in power may also have strategic considerations or rent-seeking behavior, which incurs costs on the community. This means that the production function in section 2.3 would need to be altered as an optimal investment in institutions requires a mark-up on top of the optimal tax rate, τ^*. Because economic growth requires secure property rights and investments into appropriate public goods that can often most efficiently be undertaken by a central government, there is a trade-off between having an unchecked strong state that is very well able to collect taxes but may also abuse this power to enrich the elites and a weak government that is unwilling to invest in public goods because those controlling the state are not able to tax future revenues created by the public goods investments they do (Acemoglu, 2009). An analysis of economic institutions and policies should therefore take into account both individual incentives for investment and incentives for those in charge of public goods provision.

The strength of the state or the government depends on the information asymmetry between the elite or government and the rest of the community (Mayshar et al., 2011), and on the concentration of power (Guimaraes and Sheedy, 2012). The more transparent the taxable production is realized and the more power is shared, the harder it is for the elite to overtax as members of society will choose to overthrow them. Summarizing, sufficient conditions for the establishment of designed, agent-intensive[12] institutions include the presence of strong reciprocators or diffusion agents, a set of organic self-enforcing institutions that support the type of designed institutions that society wants to set up and finally a high transparency of production such that the authority assigned to executing the collection of resources and the enforcement of the institution has an optimal power balance with regards to the members of society.

[12]Meaning that they are not self-enforcing but require active investment and reinforcement.

2.6 Stable and unstable institutional environments

As explained by Diamond (2005) in his book "Collapse", the fact that a society has reached the critical mass to create a demand for institutions and succeeded in setting them up, does not mean that this institutional environment will remain in place. In this section we elaborate on some of the factors that determine whether investment in institutions and their enforcement will be stable or not.

Ostrom (1990) argues that all known self-organizing resource governance systems that have survived for multiple generations rely on the investment of participants in monitoring and sanctioning the actions of others, to reduce the probability of cheating. Once certain institutions have evolved, how are they actually reproduced and how is sustained investment ensured? Cosmides et al. (1992) show that in hunter-gatherer communities, survival depended partially on solving day-to-day collective action problems and quickly learning who was deceitful and who was a trustworthy reciprocator. Therefore those individuals who were best able to select the trustworthy partners from the deceitful ones had a selective advantage over the rest. Evidence from evolutionary theories and cognitive psychology reveals that humans have inherited the propensity to learn social norms and that this is how institutions stabilize within societies (Ostrom, 2014). Some cheaters will always be present in the population, but as long as members of society receive a noisy signal about whether another member is a cheating type or a cooperative type, a proportion of trustworthy types will survive and so will cooperation and investment in institutions.

Factors that have been mentioned to be conducive or detrimental to cooperation are, among other things: the type of production and allocation functions; the predictability of resource flows; the size of the group involved; the heterogeneity of the group; the dependence of the group on the good that can be acquired through complex transactions, and the relative scarcity of this good; the common understanding of the group; the size of the total collective benefit; the marginal contribution by one person to the collective good; the size of the temptation to free ride; the loss to cooperators when others do not cooperate; having a choice to participate or not; the presence of leadership; past experience as a group and the level of social capital; the autonomy to make binding rules (Ostrom, 2014).

We highlight some of these factors here. Ostrom (2014, p. 244) states:

> *"If a group can determine its own membership —including those who agree to the use of the resource according to their agreed upon*

> *rules and excluding those who do not agree to these rules —the group has made an important step toward the development of greater trust and reciprocity."*

The ability to do so depends largely on two things: geography and technology. Geography plays a role because natural boundaries can help to create and defend boundaries for the society. Technology also helps because for hunter-gatherers it is very hard to defend the rights to hunt in a certain area, as animals move around freely and unexpectedly. Agricultural technology enables society to cultivate and later also store their resources in a central place that can be defended more easily. Due to the importance of geography, different degrees of cooperation and joint investment in institutions across area are expected. Due to the importance of technology we expect to see an increase in cooperation in a society after it experiences the Neolithic Revolution, although these propositions are hard to test empirically. Furthermore, Gintis (2000) shows that cooperation is enforced by the synergistic effects between group level characteristics and individual behavior; humans have the distinctive capacity to construct environments that limit within-group competition and reduce the genetic and social variation within groups. This process of group selection allows individually costly but in-group beneficial behaviors to coevolve with the supporting institutional environment. This means that the group dynamics are getting stronger and the group is getting more distinctive from non-members as it keeps on developing more elaborate mechanisms to commit its members to the rules and to keep outsiders out. One of these mechanisms is monogamy, reducing phenotypic variation and raising the threshold for new members to be introduced into the society. The exact ways in which institutions are reproduced through culture-gene interactions (genetic mediation, cultural mediation, enhancement, opposition and neutrality) coevolution and dual inheritance are beyond the scope of this study but are insightfully described in i.e. Durham (1991); Norgaard (1994); Shennan (2000); Van Den Bergh and Stagl (2003).

Empirical research has found that major migration and (too) rapid technological change, as well as periods of war, pestilence and famine can threaten cooperation (Sengupta, 1991; Gintis, 2000; Ostrom, 2014). These forces affect the coherence of the group and make the benefits of future cooperation more uncertain, thereby making cheating more attractive as compared to cooperating. It then depends on the extent of codification and enforcement of these norms and institutions whether they survive. At the same time, in these uncertain situations the benefits of cooperation are also

largest, so as long as a small group of strong reciprocators is present, this small group will outcompete other self-interested groups and their share in the population will grow.

What about the optimality or Pareto Efficiency of institutions? Is it the case that only Pareto Efficient institutions will be stable or can suboptimal institutions persist in certain situations? Boyer and Orléan (1992) explains that a convention is a social arrangement that allows people to cooperate with each other, comparable to how institutions are approached in this study. Once the convention has been established, no individual agent has an incentive to deviate from it and there are increasing returns to adoption that ensure the self-enforcing quality of the convention, this is named *economics of conformism*. Therefore, replacing an existing convention with a more appropriate one is difficult due to its reinforcing nature. It is hard to achieve cooperation, but once present, agreements are typically harder to change than individual decisions, even if a change would match the desired values by many or all involved. Boyer and Orléan (1992) explain that generally, conventions that are adopted on a large scale can only change radically through a general collapse, external invasion or war or through collective agreements, which are very hard to reach. However, there are some counterexamples of an endogenous diffusion to a superior convention. Although evolutionary forces could in some cases favor the societies with more optimal institutional arrangements, the reinforcing forces and path dependencies are very strong, and in fact, Coase (1964, p. 195) claims that *"[...] until we realize that we are choosing between social arrangements which are all more or less failures, we are not likely to make much headway."* Institutions, as Fukuyama (2011) describes it, are 'sticky', they persist over time and are changed only with great difficulty. This implies that institutions that were created to meet a set of conditions can survive even when those conditions change or disappear. This inability to adapt is often the cause behind political decay.

2.7 Historical examples

The theoretical model in section 2.3 does not predict the exact timing of the emergence of a demand for institutions. We have merely argued that this timing is linked to the timing of the Neolithic Transition. The process of the emergence of transaction cost-lowering institutions and the trajectory to sustained investment in and stabilization of these institutions is displayed in Figure 2.7.1.

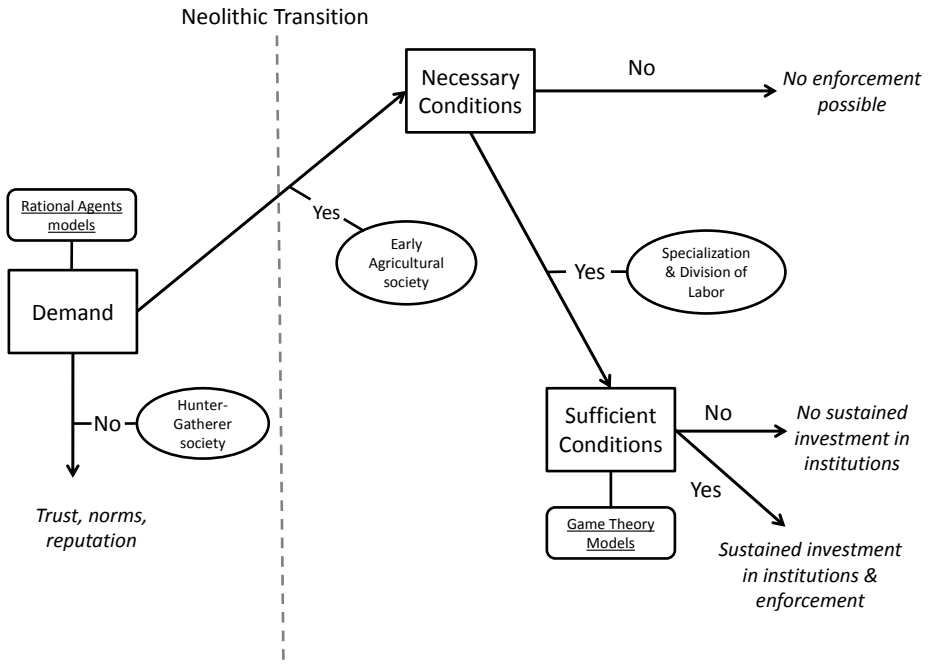


Figure 2.7.1: Theoretical flowchart of stages of institutional Development and Enforcement

We show in figure 2.7.1 that societies need to overcome many obstacles in order to successfully establish market-supporting institutions and guarantee a sustained investment and enforcement. The Neolithic Transition leads to a demand for these institutions, whereas hunter-gatherer societies can still suffice with organic, self-enforcing institutions. This stage is modeled in a representative-agent model. If the necessary conditions —food storage and market exchange leading to specialization —are not met, society does not require these institutions. Even if the demand and the means for the establishment and enforcement of market-supporting institutions exist, societies may still not be able to establish them because of principal-agent problems and freeriding behavior, if it does not meet the sufficient conditions —the presence of diffusion agents and strong reciprocators and the possibility to exclude nonmembers. If these conditions are all met, institutions that have been established tend to stay in place because of their reinforcing characteristics, even if they are no longer optimal given changed

circumstances.

We do not want to claim that the trajectory through the different stages of institutional development —the creation of demand, meeting the necessary and sufficient conditions and finally ensuring stable investment —is always a directed, chronological movement. In fact, societies can regress back from one stage to the previous one and it may take centuries or even millennia before they successfully set up a stable institutional structure. Furthermore, the nature of the transactions and the gains that can be made from either cooperating to establish institutions or from cheating or shirking behavior may differ between societies and over time. Therefore, we need casewise evidence to compare different societies, facing a variety of biogeographical conditions, and see which obstacles they faced and overcame in setting up certain institutions. This is not meant to be an in-depth analysis of these societies nor an exhaustive and chronological list of the possible stages and settings of institutional development. We merely try to illustrate with anecdotal evidence how the hurdles that we discussed above were manifested in reality and how societies were able to overcome them (or not).

!Kung Bushmen

Renfrew et al. (1974) show that when comparing early agricultural communities in the New World and Old World to the hunting and foraging communities that preceded them, the main distinguishing feature that allowed the farming communities to move beyond subsistence was the existence of specialist craftsmanship, which did not exist in the nomadic communities. These were flat societies with the absence of any grave goods or storages of value which could indicate disparities in individual wealth. A society that never developed the storage mechanisms to enable a non-food-producing elite is that of !Kung bushmen. Sahlins (1974) describes some features of this African hunter-gatherer society. A division of labour between men and women existed, but both contributed to food production; the men provided meat and maintained their weaponry while the women were foraging for roots, nuts and berries for between two and three days per week. Furthermore, there was a late assumption of adulthood in this society: adolescents were not expected to gather their own food until they were married. Child labour was not common and children were not seen as a method of extending food production. Access to natural resources was abundant and free for everyone to use, therefore the need to develop storage mechanisms was

low.

!Kung bushmen also do not consider it necessary to hoard objects, as status is not related to material possession. The !Kung historically lived in tribes of about 10-30 people and resided for some time around a water body. Once the water and resources around the village were depleted, the band would relocate to another, more resource-rich area. This shows that the !Kung bushmen lived in a way that neither caused the demand for designed, more complex institutions, nor did they meet the necessary conditions for setting up institutions. Transactions were not sequential and restricted to people within the small social or kinship circle. All fit and able members of society contributed to food production, with the exception of children, who were thus also not seen as a means of increasing wealth. Many hunter-gatherer societies consider food storage mechanisms as superfluous because of the trade-off between mobility and food security. Gathering more food to store it for later periods provides more food security, but also means that the area is depleted of natural resources sooner, forcing the community to move on. Some nomadic communities still rely mainly on the hunting-gathering lifestyle today, such as the Hadza people in Tanzania, although most of these societies have had much contact with modern civilization and do not necessarily live under the same conditions as the pristine Palaeolithic societies. In general, though, most hunter-gatherer communities in the past and in the present have not reached a level of population density and transactions high enough to cause a demand for complex institutions.

African Pastoralist

In our model in section 2.3, we argue that the increase in demand follows from a decreasing trust level when kinship ties become weaker. However, designed institutions always build on the organic, self-enforcing norms and institutions that are already present. Societies may differ in starting levels of demand for designed institutions due to different preferences or norms. Swallow and Bromley (1995) and Ensminger (1997) find that African pastoralists still rely very much on systems of inheritance and common property and prefer those to modern registration systems that could ensure private property rights. As this system works well within the still largely kin-based communities they live in, the transaction costs of changing to more designed institutions are higher for them as compared to other societies. In our theoretical model, this implies that they have a higher threshold from which investing in designed institutions becomes worthwhile.

Mesopotamia and Egypt, 3000-4000 B.C.

The distinction between hunter-gatherer, nomadic societies and agricultural, sedentary societies was not always that clear. Hunter-gatherer communities which were relatively specialized and sedentary existed, as well as farming communities that did not experience high specialization and institutional development. Mayshar et al. (2011) argue that the development of a mere agricultural surplus is not enough for state formation, since Malthusian forces would have annihilated those surpluses. Therefore, the binary existence or non-existence of food storage systems cannot explain fundamental regional variations among early states. The authors compare the rainy northern highlands of northern Mesopotamia with the alluvial planes of southern Mesopotamia and the Nile Valley in Egypt. Agriculture was taken up in northern Mesopotamia long before southern Mesopotamia and the Nile Valley joined, while it was in these latter two regions that the first city-states and central territorial states were formed and bloomed into an 'Urban Revolution'.

Mayshar et al. (2011) argue that these differences in institutional development stem from differences in the degree of transparency of farming activity. The higher transparency in Egypt and Southern Mesopotamia, as compared to Northern Mesopotamia, allowed the elite in those places to effectively collect taxes, putting downward pressure on population growth and increasing the size of the non-food-producing elite and the institutional system. Thus, although all three places went through an agricultural revolution, creating a demand for designed institutions, in Northern Mesopotamia the intransparency of agricultural production prevented this society from fulfilling the necessary conditions for further institutional development.

America Southwest and Illinois, 1-1300 A.D.

As explained in section 2.4, the degree of specialization and integration may depend on biogeographical conditions. Climatic shocks can cause a society to reach the necessary conditions for institutional development where it did not meet those before, or the other way around. Braun and Plog (1982) show how increasing environmental unpredictability in America Southwest in 700 to 1300 A.D. and Illinois between 1 to 700 A.D., respectively, intensified regional integration in these areas. The higher region-

wide uncertainty or risk arising from the environmental change led to more exchange and to increasing, and increasingly formalized regional connectedness and exchange. The more sustained the environmental change, the more permanent these effects. In addition to external climate shocks, the authors argue that the increased risk can also stem from the transition to food production itself: agriculture yields more nutrition from smaller areas of land than does foraging, but it can also entail greater local risks as society is bound to that specific environment and its climatic development.

China, 19th century A.D.

Even when a society meets the necessary criteria for institutional development, it may still find it impossible to actually execute these investments and reforms, due to freeriding problems or agency problems. Sng (2014) argues that these problems increase with community size and shows this in a theoretical model of China and its relative decline in the nineteenth century. China was a large dictatorship, but due to its size the central ruler had to rely on a large group of tax-collecting agents. However, monitoring this group is nearly impossible, giving the agents strong incentives to extort the taxpayers and leading to high corruption. This forces the central ruler to keep taxes low to prevent revolts. Economic expansion even aggravates the corruption and causes further fiscal weakening. His model is supported by evidence that the Chinese state taxed and administered sparingly, especially in regions far from Beijing, and that its fiscal capacity contracted steadily during the prosperous eighteenth century, leading to crises in the nineteenth century. This example shows that although all the necessary factors —specialization, high productivity, large population —were present in China, the country could not overcome the principal-agent problems in order to successfully further the investment in institutions. Bernstein and Lü (2003) show that high corruption and high taxation are still problems in contemporary China, providing a worrisome source of political and social instability for the Chinese government that is not able to solve these problems.

Greece, 300 B.C.

Even societies that were able to complete the whole trajectory necessary for

institution-building as described in Figure 2.7.1 can regress back in terms of societal organization and complexity. Ancient Greece was a highly developed, complex society with high levels of hierarchy and well-developed political systems, yet its democracy fell in 322 B.C. Tridimas (2015) shows in a theoretical model that war and internal conflict, rather than taxes and public expenditures, were the main redistribution means in this society. War inflicted proportional losses to the rich and the poor, but if won, greater benefits would fall on the poor than on the rich. Therefore it was a rational choice for Athenian citizens to pursue and renew war and conflicts, despite defeats. However, in the year 322 B.C. this gamble was catastrophic, when Athens capitulated and the democracy was deposed off. Democracy did re-emerge, but much later and in a different form. This is an example that shows that institutional settings that were beneficial to those involved can sometimes collapse under the influence of war or external shock. There are also numerous examples where the opposite is true, and suboptimal institutional arrangements remain in place for a very long time, either because they benefit a small elite, which through those institutions is able to increase its power (e.g. Acemoglu, 2009), or due to forces of path dependency and the economics of conformism.

We are aware that the above cases neither prove nor disprove the assumptions that feed into our model. Furthermore, all these examples and many more have been much more thoroughly studied as case studies by historians to provide insights on their institutional settings, but going to that level of depth is beyond the scope of this chapter. The purpose of this selection of anecdotal evidence here is to review some historical cases that illustrate how failing to meet the necessary conditions indeed implied not developing the institutions that we talk about and that meeting the necessary conditions enabled but did not guarantee such emergence.

2.8 Conclusion

In this chapter we focused on the mechanisms through which institutions emerge, develop, and possibly collapse. Why did certain institutions arise in one society at a given time but not, or much later, in another society? We reviewed the present literature on institutional development and showed that no single model could capture the full complexity of institutional emergence or development. Instead, in order to successfully build a stable institutional structure, a community needs to overcome many obstacles. We showed this in a flowchart and showed how each of these stages or ob-

stacles asks for different explanatory factors and different types of models. We also explained that there are no institution-free societies: when we talk about the emergence of institutions, this concerns designed institutions, always building on the organic institutions already present in the community. However, the latter are self-enforcing while the former require a costly enforcement through monitoring and punishing. These are therefore harder to establish and more interesting to analyze: a society needs to have demand for these institutions, but also meet the necessary and sufficient conditions to create them.

Demand for more complex institutions arises in those societies that pass through the Neolithic Transition and experience the related technological change. These societies still find themselves in a Malthusian trap at this point so all productivity gains are translated into larger community sizes. This technological change also needs to be complementary in the sense that it provides the necessary conditions to free resources necessary for institutional development, which is only possible with specialization enabling the existence of a non-food-producing class. The extent of specialization depends among other things on the degree of isolation and the variety and types of crops that can be grown in an area. But a bureaucratic hierarchy can also form a brake on specialization and trade, by monopolizing all transactions through a feudal system. Different factors explain why one society may progress faster to the next phase of the flowchart than others: the culture or informal institutions that are already present and natural together with perceived boundaries that help the society to be selective in their members determine how long the society can function with the self-enforcing organic institutions, before the transaction costs become too high and create demand for more complex institutions.

We assumed in our theoretical model that technological change, translating into larger, denser communities leads to higher transaction costs, but that this effect can be reduced by collectively taxing resources to set up institutions. However, the gains from institutional set-up are at the community, not the individual level, and although evolution may favor large societies that are able to control their population through 'ultrasocial institutions', these evolutionary forces may only be strong enough under competitive and uncertain circumstances, such as war. Even if resources are available and competition drives individuals together to collect these community level gains, this is no guarantee for the successful establishment and enforcement of complex institutions. The economic literature has been zealous in describing many different types of agency problems; cooperation may provide benefits for the society as a whole but individuals may

have incentives to freeride or defect, hence the community needs tools to enforce the contracts or agreements that have been made and to prevent shirking. These are the sufficient conditions for sustained investment in building and maintaining institutions, and these conditions are related to the transparency of production and the possibility of setting up 'boundaries' for membership of the community.

Therefore, in order to model the establishment and continuous investment in institutions, a multi-layered approach needs to be taken. An overlapping-generations model with representative agents can describe why at a certain point the demand for designed institutions emerges. Models that explain trade and specialization show which societies preceded in maintaining a non-food-producing elite. Game Theory models and Principal-agent models help to describe the problems that may arise in setting up these institutions and appointing an authority that acts in the best interest of the community and what mechanisms or tools exist to overcome these problems. Evolutionary theories are necessary to describe whether the institutions will be stable or not. Future theories on institutional development need to not only build on economic theory, but also shed light on the means through which institutions or conventions are transmitted from one generation to the next. Is this happening through cultural transmission, genetic transmission or the interaction effect of those? How does this affect the stability of institutions?

Our theoretical model can be extended in several ways. First of all, the threshold was taken as a fixed parameter here, but economists could use insights from behavioral and neurological sciences to introduce in the model the exact dynamics through which this limit evolves over time or differs between individuals. Secondly, we assumed here that transaction costs rise rapidly as soon as community size surpasses the limit of the social circle, but the model could also be extended using evidence from experiments on how much trust is extended to strangers in different categorizations and how much cheating actually occurs. Finally, to model a completely endogenous emergence of institutions in early agricultural societies, we need to come up with more evidence on how a production function for institution-building looks. We did not go into the question whether investing in institutions renders decreasing marginal or not, and whether a society can also overinvest in institutions. These are very important and interesting questions that remain open for future research.

The second stage of the flowchart described the necessary conditions for institutional development, which can be reduced to three factors: specialization and the division of labour, the transparency of production and the

degree to which market exchange can grow or is limited due to isolation or bureaucratic forces. Technological developments together with geographical conditions determine the availability of and need for storage mechanisms, enabling specialization and a non-producing class. However, this class can be supported only if production is transparent enough for taxes to be collected. Furthermore, if bureaucratic forces hinder the development of trade and exchange, for instance in a rigid feudalist system, institutional development is also slower than possible. These insights could be used in comparative analyses on current-day societies. Is corruption taking place because production is more intransparent or is it the other way around? Does stronger competition, for example in the case of formerly protective economies opening up, always lead to more market exchange and therefore institutional development?

When a society meets all the necessary conditions for institutional development, the actual establishment can still be endangered by different problems of coordination, freeriding and by principal-agent problems. Diffusion agents or lobbyists can ensure collective action towards a common goal, but sometimes the tendency of people to freeride or defect cannot be overcome and the public good or institution will not be successfully established. Additionally, the population and productivity growth that increase transaction costs and demand more elaborate institutions can also be counterproductive in their establishment. Monitoring problems and gains from rent-seeking also grow with population size, leading to corruption rather than a stable and supporting institutional setting. The ability to set limits on membership and internally punish members while excluding nonmembers from transactions is important for the survival of social institutions. This ability is increasingly under tension due to globalization and migration. Opening up the borders means a larger population but also an influx of people from different cultures, bringing with them different institutions. This heterogeneity and sometimes mismatch of informal institutions may make it more difficult to maintain the formal institutions that are present in the country under the new situation.

Finally we asked whether a society that successfully established complex, designed institutions will always maintain this societal organization, even if new technological changes have made the institutional arrangements obsolete or no longer optimal. Institutions or conventions, once established, have a strong path-dependent and reinforcing nature, regardless of whether they are still serving the general interest or not. They can be changed, but this might take an extreme shock or an external invasion.

3

Mechanisms of institutional diffusion between societies

An empirical study of the adoption of female voting rights across 101 countries, 1893-today

Objectives

- To shed light on the process of institutional diffusion between societies
- Particularly, to see whether the adoption of full voting rights for all women can be linked to a diffusion process besides hinging on within-country factors and external shocks
- To see if there is a spillover effect of the adoption of full voting rights for all women, i.e. if distance to other adopters predicts adoption of this institution
- To see how different types of inter-country contact influence the adoption of the extension of the franchise to women

This chapter is co-authored by Arjen van Witteloostuijn and will soon be available as a working paper.

- To analyze whether the importance of contact and diffusion versus within-country factors changes over time, i.e. to see if effects differ between early and late adopters

3.1 Introduction

In the previous chapter, we looked at the 'ex genesis emergence' of institutions as a possible consequence of demographic and technological changes after the Neolithic Revolution. As more societies became sedentary and communities grew larger in size, the incidence of contact between different societies became much higher. Therefore, this chapter is concerned with what happens to institutions once different societies come into contact with each other. Does the adoption of formal institutions hinge on the number and type of contacts between nations? Specifically, we look at whether the timing of the extension of franchise to all women is affected by a country's geographical neighbours, trade partners and war enemies or allies that already adopted this institution. While the adoption of voting rights for women has been analyzed, looking at country-specific factors, the novelty of this research is in studying diffusion as a process, focusing on the effect of contact. In a duration analysis, we test whether (a) diffusion is taking place and (b) the speed of diffusion depends on the intensity and type of contact between nations that do not (yet) and those that do have introduced this institution. We examine the relative importance of the following factors: geographical distance, international trade, diplomatic exchange, colonial ties, and inter-country warfare. In supplementary analyses, the sample is split into early and late adopters to analyze whether the same factors are important in both groups. To investigate whether inter-country contact has an effect beyond external shocks common to all countries and within-country predictors, we control for several key catalyst moments and for well-established predictors of women's rights. Additionally, we add continental fixed effects. We test our hypotheses in an event history analysis, with the year of adopting full female voting rights as the dependent variable.

The period under study ranges from 1893 when New Zealand was the first country to extend the franchise to women, until 2008. At this moment, all but a few countries have introduced voting rights for all women.

The sample consists of 101 countries all over the world.[1] We find strong evidence for diffusion, as the number of neighbours that already adopted female voting rights is a positive and significant predictor of franchise extension in all specifications. Other types of contact, the amount of trade, colonial relations and diplomatic exchanges with prior adopters, are also positive predictors but these are less robust to the inclusion of external shocks and within-country control variables. When late adopters and early adopters are analyzed separately, there are some remarkable differences: the effect of diffusion through neighbouring countries is only significant for late adopters, although trade with prior adopters and being a more open economy in general are positive predictors of the extension of the franchise for early adopters. The diffusion is mainly occurring between the late adopters, which is not that surprising as the number of prior adopters is rising over time so chances of contagion through contact are rising.

Institutional change is the subject of revived interest. Williamson (2000) argued that the more deeply embedded institutions are society, the harder they are to change. Others have shown that sometimes outside influences can function as a 'shock' large enough to fundamentally change the set of political and economic institutions in a given country (Acemoglu, 2009; Nunn, 2008). Within sociology and political science, the observation that institutions can diffuse over countries is nothing new. As Gleditsch (2002) states, democracies rise in "system-wide waves", implying that this cannot be solely due to independent internal processes. The Arab Spring of 2012 provides anecdotal evidence that reform and revolution may be contagious. Simmons et al. (2007, p. 450) argue that: "Diffusion theorists of different stripes share the view that the policy choices of one country are shaped by the choices of others, whereas conventional accounts of policy choices point only to domestic conditions." There is thus consensus that the influence of other states has to be taken into account when looking at institutional change within a country, and that this diffusion is anything but a homogeneous process. This fact that choices of actors from one particular spatial unit are affected by those of actors in other spatial units is also called 'spatial dependence' (Perkins and Neumayer, 2009). The power prestige hypothesis argues that the characteristics of the major powers that

[1] As these are not all countries, the issue of selection bias warrants some consideration. The sample includes all countries for which there is full data. We have tested for selection on the observables, which does not appear to influence our results, but we cannot test for selection on the unobservables. The countries that drop out of the sample are mainly the smaller countries where data collection proves an overall problem. The sample does include countries from all continents and across all 'waves of adoption'.

originally adopt a policy affect the way in which this policy is diffused to other states (Fordham and Asal, 2007). Acemoglu et al. (2001) look not at the characteristics of the transmitter, but instead at the characteristics of the recipient, arguing that the disease environment in colonies determined the type of institutions they ended up with.

Although the characteristics of the recipient country are important, one piece of the puzzle has been less prominent in empirical research to date: the effect that the adoption of the institution by other nations has, through different channels of contact, on the chances that a particular country introduces this institution as well. We seek to examine the impact of direct contact between two countries, and analyze whether the intensity and type of contact has an effect on the diffusion of an institution, controlling for within-country factors such as changes in the political regime, as well as external shocks such as inter-country wars.[2] We would thus like to go beyond describing the diffusion, and also find out the causal mechanisms behind this diffusion, paralleling Salehyan and Gleditsch (2006) in their research on the spread of Civil War. We analyze how the intensity and type of contact between two states (war, diplomatic bonds, colonial dependencies or trade) influences the sudden emergence of voting rights for all women in those states.

When studying the effects of inter-country contact on the adoption of an institution, many different institutions could be considered. We decided to focus on the adoption of voting rights for all women, as this institution has now been established in a large part of the world. This makes cross-country comparisons most meaningful. The first country to introduce full woman franchise was New Zealand in 1893; in 2008 all but a few countries have adopted this institution. This institutional adoption thus took place during a period for which historical data on different types of contact are reasonably well recorded. Furthermore, the extension of franchise to all women is a clear change that is formally documented. This means that our study is restricted to de jure and not necessarily de facto institutional change of women's rights. This is unavoidable as this is the only objective way to compare institutional change across countries. In future research, the analysis might be extended to include more institutions, possibly including a more recently adopted or more historical institution, or comparing the adoption of an economic and a political institution. We focus on whether and to what extent the type of contact between two states

[2]These large wars are also a form of contact, but we interpret them as shocks here because they also influenced countries not directly involved in the war.

(inter-country warfare, diplomatic bonds, colonial dependencies, and international trade) influence the sudden emergence of voting rights for all women in those states.

This chapter is structured as follows. In the next section, we briefly develop our hypotheses, and explain how they follow from theory. As a second step, we elaborate on the history and spread of female voting rights and summarize the insights from extant work on this institution. In the following section, we describe the data and our sample, as well as our empirical strategy. We describe our results and discuss these results and their interpretation in the final section.

3.2 The adoption of female voting rights

The nineteenth and twentieth century were characterized by different "waves of democratization", including extensions of franchise to men of all ranks and status and sequentially or simultaneously to women. In 1893, New Zealand was the first country to grant all women the right to vote in national elections, followed shortly by Australia in 1902. The extension of franchise was a more gradual process in some countries and involved an abrupt change in others. In the United States, women were already allowed to vote in several states by 1910, but the constitutional amendment affecting all states was only passed in 1920. In Europe the extension of franchise started in Finland in 1907. Several other European countries followed before or shortly after World War I, like Norway, Denmark, Sweden, Great Britain and Germany.[3] In the years 1918 and 1919, Austria, many Eastern European countries and Belgium and the Netherlands also joined. Another wave of extensions happened in Southern Europe; Spain and Portugal enfranchised women in the 1930s, but France, Italy and Greece only did so after World War II. The exception in Europe is Switzerland, where women were only enfranchised in 1971. Between 1940 and 1950 several Asian countries extended franchise to women. In Latin America we observe two waves: the first in the 1930s (Ecuador, Chile, Brazil), and the second after World War II when several other countries joined. On the African continent, South Africa was the exception with an early extension in 1930 (to white women only). For most African countries, female vot-

[3]This is the German Weimar Republic. We do not consider the constitutions of East and West Germany that followed as they did not remain independent countries until the end of our data period. Both countries did have full voting rights for women in place during the whole period of their existence.

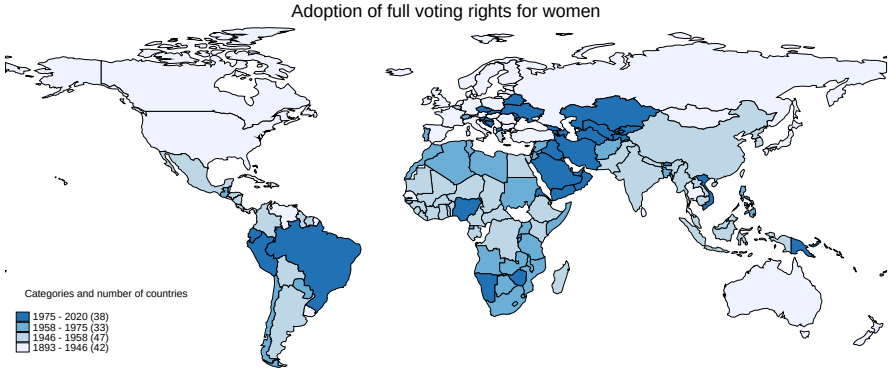
ing rights came with the post-decolonization constitutions, between 1950 and 1970. After the ending of the Cold War, several new countries came into existence that implemented universal suffrage directly in their constitution. Currently, only a handful of countries in the Persian Gulf do not grant women voting rights. There have been a few recent extensions, notably in Kuwait in 2005, the United Arab Emirates in 2006 and Saudi Arabia in 2015.

The literature on the evolution of women's political rights is substantial. The extension of franchise to include women has been studied extensively in economics as well as political science, as it has been linked to structural changes in policy, such as the composition and size of government spending, and other variables such as child health (Aidt et al., 2006; Aidt and Dallal, 2008; Docquier et al., 2015; Hofer, 2015). These studies focus on the consequences of the extension of franchise, rather than on the antecedents. Añón Higón and Minguela (2014) look at women's suffrage and political representation in the period 1893-2010, relating this to economic development and women empowerment. They reveal a nonlinear relation between economic growth and women's suffrage, but show that external causes such as treaties and international conflicts also have strong effects. Importantly, they report a difference between early adopters (where economic development constitutes the main factor) and latecomers (where global demand for women's rights, expressed in treaties, is more important). They were not able to show the process of diffusion yet, as their way of measuring a 'neighbour effect' depends only on the total share of countries already having adopted full suffrage, not taking into account the distance to the country under observation.

Damjanovic et al. (2015) that economic development matters, for the following reason: as a society develops, it accumulates physical and human capital, which feeds into higher productivity of females, equalizing the gender balance of power. McCammon and Campbell (2001) analyze the differences between the Western States in the USA, being frontrunners with regards to full suffrage, and the rest of the country. They find that neighbouring states having adopted woman suffrage already definitely matters, as well as the presence of females in mainly male dominated domains such as higher education and specific professions (i.e., lawyers and physicians). The strategies of political movements are also found to be influential. King et al. (2005) find that, while controlling for other political variables, women suffrage movements are especially important in the early stages of the legislative process, and not so much in the later stages.

Other studies have linked specific types of contact to changes in democ-

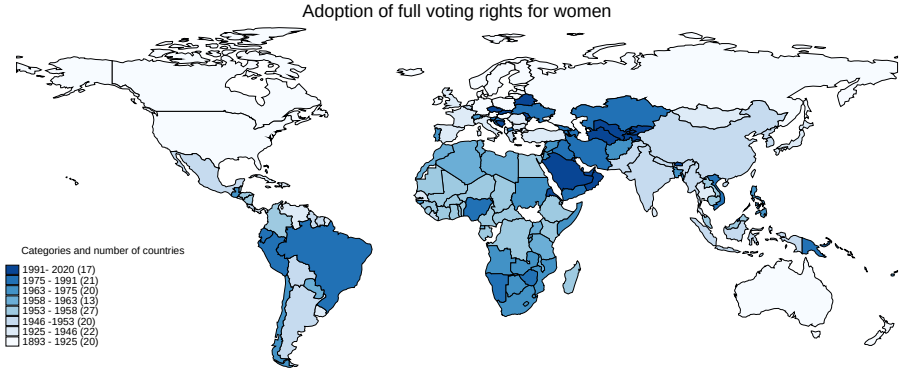
Figure 3.2.1: Geographic representation of adoption year



racy. For example, the migration between two countries may also affect the intensity and type of contact between those two countries, opening gateways for diffusion. Beine et al. (2013) find that there are behavioral transfers between migrants' receiving and sending countries, and that a larger share of migration to countries more democratic than the home country therefore also increases democracy in the home country. Docquier et al. (2015) confirm this finding for developing countries specifically, and find that openness to emigration increases the institutional quality in those countries.

To get an idea of the geographical diffusion process of female voting rights, we visualize the timing of adopting full female suffrage on a world map. The grouping of categories influences the way the geographical diffusion process looks like. As a start, we show a graph with four categories based on quartiles. For the purpose of creating these maps, the four countries that did not adopt full female suffrage yet were assigned the year 2020 for adoption, such that a darker color represents later adoption. The numbers behind the categories in figure 3.2.1 indicate how many countries fall into that category. A general pattern emerges. However, the first category includes more than 50 years as adoption was more spaced in the beginning. It may be more meaningful to increase the number of categories. In Figure 3.2.2 we divide the countries into eight categories based on an equal distribution. Now we can see that the first 30 years were grouped together, but we do see more variation, especially in the Middle East region. The pattern that emerges is that the early adopters were mainly located in Europe and the Western Offshoots, followed by Asia and some countries in Central and South America, after which Africa followed, then the rest of South America and finally the Middle East region. However, instead of

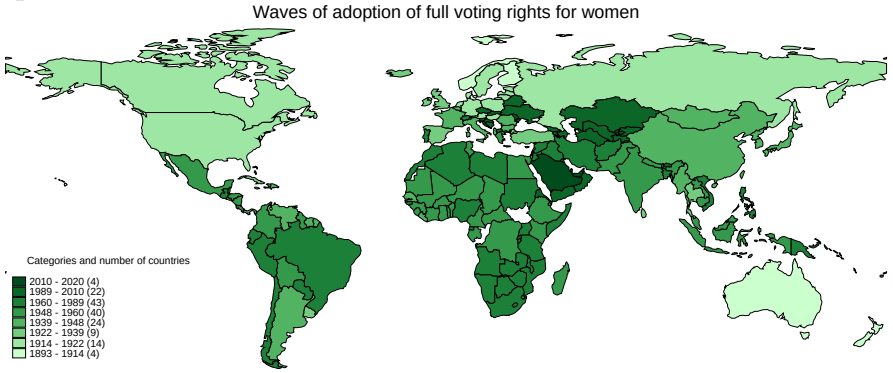
Figure 3.2.2: Geographic representation of adoption year, quartile intervals



categorizing into equal groups based on the distribution, it may be more insightful to combine historical events together with a visual picture of the distribution over time to identify 'waves of adoption' (see Figure 3.2.3). Aidt et al. (2006) divide the extension of the female franchise in two waves (after World War I and after World War II), but a more detailed division may provide further insights. We define the following eight categories: 1893-1914 (Pre-World War I), 1913-1922 (World War I and aftermath), 1922-1939 (interwar period), 1940-1948 (World War II and aftermath), 1949-1960 (Decolonialization), 1961-1989 (Cold War), 1990-2010 (end of Cold war) and 2010-2020[4], as is shown in figure 3.2.3. The effect of geographical proximity on adoption and diffusion seems to be clearer with this categorization. We can now distinguish between the four very early adopters (New Zealand, Australia, Finland and Norway), and the other parts of Europe and the Western Offshoots that adopted full franchise during or after World War I and the interbellum. Many Asian countries adopted full franchise during World War II and its aftermath. The majority of Africa and South America are clustered in the Decolonization wave, the newly formed Eastern European countries fall into the end of the Cold War wave, and the Middle East region either adopted full female suffrage after the end of the Cold War or has not adopted it yet. However, we do see from these three graphs together that geography seems to play a role, but that it cannot be the only determinant. For this, the geographical pattern, both within continents and across the world, is not clear enough.

[4]This is a fictitious category as it starts when the data collection ends. So, this time window includes countries that do not have suffrage yet at the end of the data period.

Figure 3.2.3: Geographic representation of adoption year, custom waves of adoption



3.3 Hypotheses

Many studies have examined the adoption of voting rights for women, but they have tended to focus only on within-country factors, such as democracy and the (labor-market) position of women. Our analysis controls for the most important factors mentioned in this literature, replicating this prior work. The added value of our study is that we also add the lens implied by the process of (spatial) diffusion among countries by including the type and intensity of contact between nation-states. In doing so, we connect to relational models that suggest that rates of diffusion vary with the level of interaction between prior and potential adopters (Strang and Meyer, 1993). In the context of the diffusion of female suffrage, we formulated seven hypotheses.

The first hypothesis relates to the effect of geographical distance. We hypothesize that geographical distance has a negative effect on adoption. This prediction follows from gravity logic, widely used in the social sciences, arguing that the volume of a certain variable (traditionally trade) is inversely related to the distance between the two entities (Tinbergen, 1962).

Hypothesis 1. *The shorter the distance between a country A that has not yet adopted female voting rights and all prior adopters, the larger the probability that this country A will also adopt female suffrage.*

Moreover, contagion is more likely to happen if there are more neighbouring countries that have already adopted the institution. The effect of local neighbours was found to be important in the adoption of universal suffrage by Ramirez et al. (1997) for the period 1930 to 1990, reporting that

regional neighbours with women's suffrage influenced holdout in this period. Brinks and Coppedge (2006) confirm the influence of the average degree of democracy among contiguous neighbours on the degree of democracy for countries between 1972 to 1996. Therefore our second hypothesis involves this 'neighbour effect' (Dietz, 2002):

Hypothesis 2. *The more prior adopters are neighbouring a country A that has not yet adopted female voting rights, the larger the probability that this country A will also adopt female suffrage.*

Cao (2010) explains that policy learning and emulation is more likely to occur across two countries that are close to each other, and that closeness can have different meanings. Next to geographical closeness, countries can be close in a cultural sense, when they share a language, religion and customs. As a robustness check, we include cultural distance to prior adopters. Besides geographical and cultural closeness, we examine how the type and intensity of contact influences the diffusion of the institution. We distinguish four types of contact: trade, inter-country war, diplomatic ties and colonial heritage. We argue that all four types of contact positively impact diffusion, but for different reasons. In sociology, four mechanisms of diffusion are identified: coercion, competition, constructivism and learning (Simmons et al., 2007). If diffusion happens through coercion, the powerful nation-states or international institutions threaten with sanctions or promise aid in return for the adoption of a certain institution. Competition implies that diffusion emerges because countries are competing for investment and export, and therefore adjust their national norms and rules to lower the cost of doing business in the hope of reciprocity. Constructivism involves policy and institutional changes promoted by expert epistemic communities and international organizations, who define economic progress and human rights. Finally, learning has to do with countries that learn from their own experiences and from those of other countries, especially their well-respected peers. In the context of the extension of franchise to women, it is not apparent that one of these mechanisms would be more relevant than any or all of the others. This may rather depend on the ways in which a country is exposed to others that already adopted the institution, and thus of the nature of contact with those others, i.e. friendly or unfriendly, and intense or superficial. The four types of contact that our study analyzes are trade, diplomatic ties, war and colonial heritage.

Regarding international trade, we argue that the mechanisms of com-

petition and learning play an important role. First of all, as soon as a country interacts with an international trading partner, the domestic producers have to compete for demand. Under these forces of competition, the country will likely be tempted to copy institutional successful features from its trading partners to not to fall behind (Simmons et al., 2006). Additionally, learning may emerge in the form of seeing how specific institutions lead to certain outcomes for the trading partner, which may trigger the home country to also adopt the institution.

Hypothesis 3. *The more a country A that has not yet adopted female voting rights trades with prior adopters, the larger the probability that this country A will also adopt the female suffrage.*

With regards to diplomatic relationships, the argument is that the mechanisms of (soft) coercion and learning will be important for diffusion. The diplomatic partners may coerce the home country to adopt similar institutions, and again a positive learning mechanism may also operate, known as the cohesion theory among organizations (Davis, 1991).

Hypothesis 4. *The more diplomatic ties a country A that has not yet adopted female voting rights has with prior adopters, the larger the probability that this country A will also adopt the female suffrage.*

However, as Simmons et al. (2007) state, *"Coercion can be exercised by governments, international organizations, and nongovernmental actors through physical force (Owen, 2002), the manipulation of economic costs and benefits, and even the monopolization of information or expertise"*. Besides the soft coercion taking place in diplomatic negotiations, more forceful ways of coercion arise when countries are at war. In that case, the institution may be imposed by the more powerful player, the victor, upon the less powerful player (Peic and Reiter, 2011).

Hypothesis 5. *If a country A that has not yet adopted female voting rights lost a war or yielded to a prior adopter, the probability that this country A will also adopt female suffrage is higher.*

As a robustness check, we will also explore the effect of being engaged in war in general, as this may affect the position of women in the home country regardless of the outcome of the war and the institutional quality of

the enemy. If a country is a colony of another, more powerful country, this could also lead to forceful coercion by the colonizer, imposing institutions upon the colony. Furthermore, a constructivist mechanism might be forceful when authorities of the colonizing country design certain institutions in the colonized country.

Hypothesis 6. *If a country A that has not yet adopted female voting rights has colonial ties with a prior adopter, the probability that this country A will also adopt female suffrage is higher.*

The four types of contact were of varying importance in different periods of time. Globalization studies argue that higher interdependency between countries due to increased international trade decreases the frequency of conflicts between countries, although empirical evidence remains mixed in this area (Barbieri and Schneider, 1999; Gissinger and Gleditsch, 1999).

Hypothesis 7. *The relative importance of unfriendly contact (inter-country war and colonial ties) will decrease over time, whereas the relative importance of friendly contact (international trade and diplomatic relations) will increase over time.*

There are theories, however, that state that the effect of contact on diffusion may not always be as expected. Strang and Meyer (1993) argue that exchange can also generate cultural divisions of labour and therefore the effect of theorization should be included when looking at the diffusion of practices. The institution that we look at, however, has been elaborately theorized and the diffusion is taking place in nation-states, which is also conducive to diffusion according to the authors. The next section describes the data and methods used to test these hypotheses.

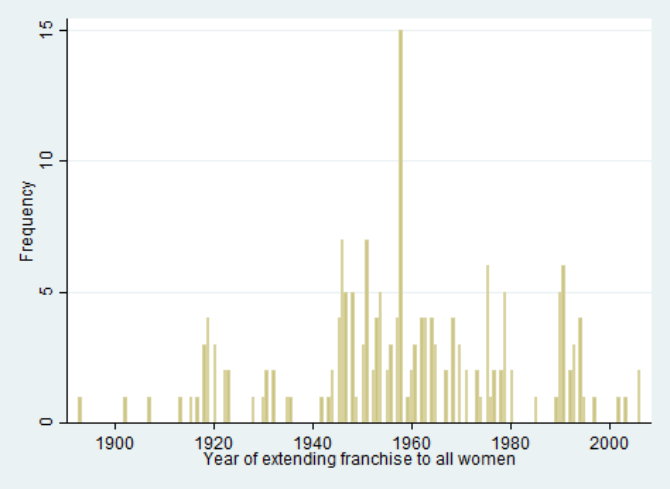
3.4 Methods and data

3.4.1 Dependent variable

The dependent variable is the adoption of universal suffrage for women. This institution is widespread nowadays, as only seven out of the 202 countries in the dataset have not yet adopted female suffrage. In the literature, an ongoing discussion involves the question as whether or not democracy is welfare-enhancing. Equal rights for men and women are generally seen as good for welfare in the broader sense and equal treatment of men and women is a fundamental right. Figure 3.4.1 displays the distribution over

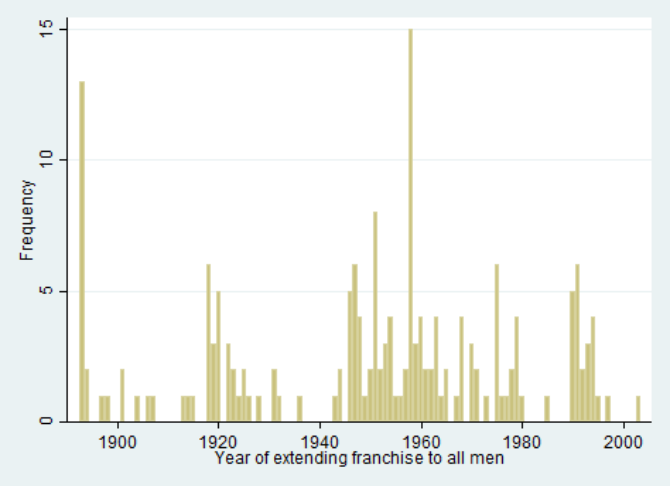
time of the extension of suffrage to women. A visual inspection shows that one can distinguish several waves or patterns in the data set. First, early adopters emerged, before the First World War. Second, shortly after World War I, we notice a small spike, after which larger spike just occurred just after the Second World War. These two periods are often described as two important waves of democratization. Third, after that, the real diffusion seems to kick in, with a large but decreasing density of countries adopting full female franchise until the last year in the dataset (2008). The year 1958 strikes out with 15 countries adopting the institution in that year; these are all African countries.

Figure 3.4.1: Distribution over time of the first year of full female franchise



Data on universal franchise are taken from the Political Institutions and Political Events dataset (Beck et al., 2001). In the variable 'F' (franchise), franchise is coded from 0 to 7, where 7 indicates that all men in society can vote. A second digit codes suffrage for women, where 0 indicates that no women can vote, 1 indicates that women are qualified on a narrower basis than males, and 2 indicates that women are qualified equally as males. We decided to look at the year in which a country achieves the status of 72, such that all men and women are allowed to vote. The alternative could be to include all cases where the second digit was a 2, but because this can refer to strongly limited voting rights (for both man and woman), we decided to focus on full universal suffrage only. Limited voting rights exist when there is some threshold on education, property, income or taxes, such as in Portugal. Here, women were granted the right to vote in 1931, but

Figure 3.4.2: Distribution over time of the first year of full male franchise



with even stronger restrictions than man; women were only allowed to vote if they had completed secondary or higher education, whereas the only requirement for men was that they knew how to read and write.

The first country to receive the status of full and equal voting rights was New Zealand in 1893, and the last country that obtained that status without censoring was Kuwait in 2006. On average, countries achieved full universal suffrage in 1952; the median year of adoption is 1958. Figure 3.4.2 shows the distribution of first occurrence of full male voting rights. Here, the mean year of adoption is 1949, and the median year is 1954. The average lag between male and female suffrage is nine years, but the median is zero years, as many countries introduced democracy simultaneously for both sexes. The largest lag that can be found is 68 years (for Paraguay). For 10 per cent of the countries, the lag is 45 years or more. We see a large spike in 1893, but this is by default as this is the first year in the dataset. So countries that already had full voting rights before 1893 are given that value in this first year. Furthermore, we observe, again, the two waves of extension of franchise, after the two World Wars, and the year 1958 is again marked by a large set of African countries. In the Appendix, in figures 6.2.1 and 6.2.2, the same two graphs can be found for only the 98 countries in our sample. Here we see that not all the African countries that extended their voting rights in 1958 are included, showing a smoother pattern over time.

3.4.2 Independent variables

Geographical distance

The data on geographical distance measures, related to Hypothesis 1, are taken from the Gleditsch and Ward (2001) dataset. They provide two variables: the first is the kilometer distance between the capital cities of a country-pair, and the second is the minimum distance between any part of the two countries.[5] We generate two time-varying variables of geographical distance to prior adopters. The first gives the average capital-city distance to all countries that have the institution already in that given year (*Mean distance*). The second variable gives the number of countries within a maximum range of 640 kilometer over water or lad (400 miles, see Gleditsch and Ward (2001) on their treatment of distance variables) that have adopted the institution already, in any given year (*Close neighbours*). The dyadic distance dataset consists of 41,006 country-pairs, out of which 1,944 country-pairs have a minimum distance of less than 640 kilometer between them. 21 out of 202 countries have no neighbours within the 640 kilometer distance circle. With regards to the mean distance to prior adopters, we expect that the distance depends mainly on New Zealand in early periods, hence being large. Over time, as more countries start adopting full suffrage, the distance starts declining rapidly. However, as time progresses and almost every country worldwide adopts full suffrage, the distance rises again, albeit slowly. The nonlinear nature of this variable makes the estimations for this variable hard to interpret. Therefore, as a robustness check we include cultural distance to prior adopters. Cultural distance is proxied by the linguistic proximity index, taken from Melitz and Toubal (2014). Taking linguistic proximity as a measure of cultural distance is common in studies explaining trade between countries (Felbermayr and Toubal, 2010; Isphording and Otten, 2013).

International trade

The independent variable relating to Hypothesis 2 is the share of total trade volume (in millions of dollars) that is carried out with prior adopters. Trade flows are taken from the Correlates of War dataset (Barbieri et al., 2008). The maximum trade flow is between China and the USA in 2008. In addi-

[5] In the data we find several pairs with distance 0. These are all previous names of the same country. These cases are dropped out of the sample.

tion to that, openness, proxied by total trade volume in millions of dollars as percentage of GDP is included as a control variable. A problem with these two variables is that for many countries, trade is not measured yet in the years prior to adoption of the institution. Including these variables thus reduces the number of events in the analysis by 50. Therefore, we run analyses after both including and excluding these two trade-related variables.

Inter-country war

To test Hypothesis 3, we use data from the Correlates of War project (Sarkees and Wayman, 2010). These data relate to dyadic and monadic wars. However, the dyadic data do not go back far enough in history so we used the software from EUgene to convert the monadic war into a dyadic war dataset (Bennett and Stam, 2000). Building on the coercion hypothesis, we argue that the diffusion probability increases if a country loses a war or yields to another country that already extended the franchise to women. The first variable, *Waryieldlost* takes on a value of 1 if the country either lost from or yielded to another country that already adopted the institution in a given year. We also control for the general effect of being engaged in war, as will be explained in the section on control variables.

Diplomatic exchange

Data on diplomatic bonds are taken from the Correlates of War dataset (Bayer, 2006), to explore Hypothesis 4. This includes data on diplomatic exchanges (*Diplomatic exchange*), which captures all diplomatic representations between states from 1817-2005. We transform this data into a variable counting the yearly contacts with countries that already adopted the institution. We do not distinguish between different types of diplomatic exchanges, but merely provide a count of the total number of exchanges per year with countries that already adopted the institutions.[6]

[6]The data is not available for every year, especially in early years, therefore missing observations were filled by using the observation for the last recorded year for that country.

Colonial dependence

We include colonial ties to examine Hypothesis 5, again using data from the Correlates of War project (CoW, 2002). This is an undirected variable: whether country A was a colony of country B or vice versa does not matter. The aim is to merely analyze whether this type of relationship is conducive to the diffusion of the institution, compared to other types of contact. *Colonial dependence* is coded as 1 if, in that year, there was a colonial dependency relation with another country that already adopted female voting rights. We test the effect of the total number of colonial ties with prior adopters for every year.

In Table 3.4.2, we report the bivariate correlations between all our main

Table 3.1: Correlation table contact variables

	Mean. Dist.	Close Neigh.	Cult. Dist.	Trade	War	Col. Dep.	Dip. Exch.
Mean Dist.	1.000						
Close Neigh.	0.231*	1.000					
Cult. Dist.	0.832*	0.241*	1.000				
Trade	0.722*	0.264*	0.660*	1.000			
War	-0.076*	0.033*	-0.006	-0.081*	1.000		
Col. Dep.	0.053*	0.089*	0.154*	0.204*	0.052*	1.000	
Dip. Exch.	0.459	0.209*	0.610*	0.520*	-0.007	0.379*	1.000

independent variables. The bivariate correlations are generally quite high and significant, but only the correlation between mean geographical distance and mean cultural distance to prior adopters is worryingly high. However, cultural distance will be introduced as a substitute for geographical distance in a separate analysis so these two variables will not be introduced in the same regression. With regards to the other covariates, we add them stepwise to make sure that there is no bias introduced due to multicollinearity.

3.4.3 Control variables

Besides contact with prior adopters, within-country factors and general events[7] could influence the adoption hazard rate. Our first control variable is annual GDP per capita *(GDPcap)*, which is a time-varying covariate. The level of development is linked to the extension of franchise in a theoretical model by Lizzeri and Persico (2004). These data are taken from the

[7]Meaning not specific to a pair of countries.

PIPE dataset (Beck et al., 2001), and are measured in dollars. The mean GDP per capita is 2,891 dollar. The highest GDP per capita in the dataset is Qatar in 1973, with a GDP of 42,916 dollars. This may seem peculiar, but remember that countries drop out of the dataset as soon as they adopt full female voting rights. So, for many rich Western countries, the latest year of measurement is somewhere in the early twentieth century, when GDP per capita was much lower all over the world. Our second control variable is openness *(openness)*, which is the total trade volume over GDP. This variable is calculated using national trade data from the Correlates of War project (Barbieri et al., 2008). Countries that are more open could have a higher tendency to adopt new institutions in general. The most open country is Saudi Arabia in 1992,[8] with a ratio of 29 of trade volume to GDP. Our third control variable is legal origin. These data are taken from the CIA World Factbook CIA (2014). The common law system is different from the civil law system and other mixed or religious legal systems by generally allowing for more political flexibility. This is captured by a dummy variable, with 1 for common law systems *(Common law)* and 0 for civil or other (mixed or religious) legal systems.

Our fourth control variable involves differences in the position of women between countries. A possibility is to look at female labour participation, or composite gender equality indices that change over time. However, most measures of gender equality in different domains are (i) based on very recent data, which is pointless in a longitudinal context, and (ii) highly endogenous to the dependent variable, which would lead to biased estimates. Therefore, we instead use an instrument that has been linked in prior work to time-invariant differences in gender roles between countries. These persistent differences have been ascribed to differences in cultural beliefs about the appropriate role of women in society (Fortin, 2005; Fernandez, 2007; Borck, 2014). We follow Alesina et al. (2013) and take traditional plough use as our proxy for the position of women in society. The link between plough use and gender role differences was originally put forward by Boserup (1970): in societies where plough agriculture is practiced, the argument goes, there is more need for physical upper body strength, implying that men have an advantage in farming vis-à-vis women. This triggered a division of labour and specialization along gender lines, and the persistent belief that the place for women to work is at home, even

[8] Again, this may seem peculiar, but we observe this because (1) oil revenues are important here and (2) countries have become more open over time, and developed countries dropped out of the analysis early in our time window.

when the main production in that society is no longer in agriculture. Data on plough use is available at the ethnic group level from the ethnographic atlas but were transformed into country-level data by Bolt (2012).

Ethnic groups are classified into one of three mutually exclusive categories: i) absence of plow animals/plow cultivation, ii) plow cultivation was not aboriginal but well established at the period of observation or iii) plow cultivation was aboriginal prior to contact. Differences may exist in the extent of gender equality between societies where the plow was aboriginal and those where the plow was introduced after contact, but prior work did not establish this as of date. For our research, the total absence of plow cultivation is used as a measure of higher persistent gender equality. In the country-conversed data, each country has a weighted average between 0 and 1 of the values for this variable of the different ethnic societies that pertained to this country. A higher value for this variable, *(plow absent)* means higher gender equality. We also test whether a variable of contemporary gender equality, the 2008 Gender Equality Index from the Human Development Report, performs better than this proxy for historical gender roles, by way of robustness check. Of course, we do not argue that the gender equality in 2008 can directly explain the adoption of female voting rights in the preceding years, but rather assume that differences in gender equality today between countries may also reflect persistent differences in gender equality in the years before.

Additionally, we control for the shares of each of the four major religions (Hindus, Muslims, Christians and Buddhists) as well as the share of atheists and agnostics in the country, as religion has also been linked to the position of women in countries (Bertocchi, 2011). These variables are taken from ARDA and are collected for the year 2010 (Association of Religon Data ARDA, 2011). Hence, this pair of control variables only work if they reflect time-invariant inter-country differences. This makes sense, given the long and persistent history of dominant (non-)religions across different societies.

Finally, a few important events during the time period under study have likely been catalysts for institutional change: the ending of the First, Second and Cold Wars, respectively. Therefore, we include dummy variables for the years 1919, 1946 and 1990 *(WWI, WWII* and *Fall Wall)*. Given that these effects may need some time to materialize , we also create a five-year lag of these variables, as a robustness check. We also control for the general effect of being involved in a war. For example, Milkman (1987) has shown that wartime usually stimulates feminism and promotes gender equality as women are required to take over traditionally male tasks when

a large share of the male population is away in the militant fighting force. This variable, *(ongoing war)* is a dummy that takes the value of 1 for each year in which the country is at war.

We ran a few robustness checks with additional control variables. First, we add a proxy for the level of democracy, the *Polity* index from Marshall et al. (2014), which is available from 1800 until 2013 and ranges from -10 (strongly autocratic) to +10 (strongly democratic). Second, we control for the net migration rate in 2010 *(migrate)*, as e.g. Spilimbergo (2009) and-Docquier et al. (2015) link migration to changes in democratic institutions. Third and finally, we check whether our results are robust to the inclusion of continent fixed effects. Descriptive statistics for all control variables can be found in Table 6.2 in the Appendix.[9]

3.4.4 Empirical strategy

To analyze the event of interest, the adoption of full female franchise, and to look at the process of diffusion behind this event, normal OLS procedures are not appropriate for two reasons. First, the dependent variable is associated with a large number of zeroes and therefore the distribution of event times is far from normal. Second, ordinary linear regression methods cannot effectively handle the censoring of observations that is prominent in the type of data that we analyze, and this will lead to biased estimates. Left censoring occurs when the event emerged before the start of the dataset's time window, but it is unknown when exactly (here, in which year). Right censoring occurs when the event has not emerged yet at the end of the measuring period. Left censoring is not present in this dataset, as we know when each country adopted full female franchise. Right censoring is present, but only for a small share of the countries. A more appropriate estimation method to analyze this type of data is event history analysis, which is a form of duration analysis. This estimation strategy is based on a hazard rate function, indicating the risk of the event happening at time t, given that it has not happened yet. This hazard rate function can be made conditional on different time-constant and time-varying covariates, as well as on a function of time itself. One can either specify the time distribution completely or opt for a semi-parametric model, which is the preferred model here. By choosing a proportional hazards model, which is a semi-parametric model, the time distribution consists of two components:

[9]The correlation table including the bivariate correlations between all the independent variables and all control variables is very extensive and therefore not displayed here, but it is available upon request.

one that varies over time identically for all members of the population, and another one that depends on the measured characteristics of the individual members of the risk set (Cox, 1972).

An assumption of this model is that all events are single, unrepeatable events. Therefore, the first time that a country moves to full female voting rights is chosen as the event, even if there would be a regression in female franchise later in history. Indeed, this is very rare in the dataset. Another feature of the proportional hazards model is its log-linear nature, implying that the effects combine multiplicatively and that the magnitude of effects of covariates may vary over the range of other covariates. This means that a country's risk factors can be updated over time, such that as more of the country's contacts adopt female franchise, the country's estimated risk can change accordingly. This feature makes event history analysis the suitable method for analyzing our data, as one country's risk of adoption is hypothesized to depend on the adoption patterns of partners with which the focal country has contact. As adoption of voting rights for women is measured annually only, this implies the possible occurrence of tied survival times. So, survival time should be treated as discrete in the analysis (Jenkins, 2005). We use the discrete analog of the continuous proportional hazards model: a complementary log-logistic (clog-log) model. The spatial element of this model is captured in two independent variables, rather than estimating a full spatial model. We believe that these two variables together capture the elements of geographical proximity that are relevant. This keeps the model parsimonious. As there is no theoretical basis to assume that the time-dependence of the adoption of full female voting rights follows a certain distribution, we do not assume any distribution a priori, as in the Cox proportional hazards model. The proportional hazard rate model can be written as:

$$\lambda_i(t \mid x_i) = \lambda_0(t) \exp(x_i'\beta) \tag{3.4.1}$$

Here, $\lambda = \Pr(\mathrm{T}{=}\mathrm{t} \mid \mathrm{T} \geq \mathrm{t})$ and $\lambda_0(t)$ is a baseline hazard function that describes the risk for individual countries with $x_i = 0$, that serve as a reference cell or pivot, and $\exp\{x_i'\beta\}$ is the relative risk, associated with the set of characteristics x_i. In this proportional hazard model, the increase or reduction in risk, dependent on the values of the covariates, is the same at all durations t. Adding the set of covariates, the empirical model becomes the following:

$$\lambda_i(t \mid x_i) = \lambda_0(t)\exp(\beta_0 + \beta_1 * CatalystYears + \beta_2 * GDPcapita_t + \beta_3 * TradeOpenness_t + \beta_4 * MeanDistance + \beta_5 * CloseNeighbours_t + \beta_6 * Tradeshare_t + \beta_7 * ColonialTies_t + \beta_8 * DiplomaticTies_t \beta_9 * WarYieldLost_t + \beta_{10} * Ongoingwar_t + \beta_{11} * Legalsystem + \beta_{12} * GenderEquality) \tag{3.4.2}$$

In a survival analysis it is necessary to assume the risk set. We assume that all countries are at risk starting in 1893, which is when the institution of full female voting rights was first adopted, regardless of whether data collection for this country starts before or after 1893. When a country has adopted the institution, it is no longer at risk and exits the dataset. Duration thus ranges from 0 to 115 years. One could argue that countries can only be at risk if they are an independent nation-state during the whole period. We check for this in two ways. First, we run a separate analysis only for countries that were independent before 1893. This greatly reduces the sample size and may lead to selection bias. Second, therefore, we include the year of independence in our standard regressions to see if that changes the results. Data on independence is taken from the PIPE dataset (Beck et al., 2001). For countries that were independent before the year 1000 the date is set arbitrarily at 1000. For countries that temporarily lost independence having been annexed to other countries, the most recent date of independence is taken.

Another issue of concern is the fact that we have an unbalanced panel. Not all countries have complete data for the full period during which they are at risk. Countries are deleted casewise if they have missing data on any of the variables included in the regression. For our most basic model, we have a sample of 101 countries, out of which 98 have adopted female voting rights, but including all the relevant covariates reduces this sample to 58 observations. There is a small difference in the mean year of extension to full franchise between the large and small sample, respectively 1961 and 1960, but this is not significant. So, selection on the dependent variable is not an issue. We do check whether our most important results are robust to the different specifications. We also observe a slight difference in the mean year of adoption between countries that are in the sample as compared to the full population, respectively 1961.1 and 1961.8, but this is again not significant. Our sample consists of a large range of countries, developed and developing, spread out over all the continents and with a

Table 3.2: Descriptive statistics for main contact variables

	Sample 1 (N=101)					Sample 1 (N=55)				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Mean Distance	2650	1496	1669	0	8724	1827	1437	1619	0	8409
Close Neighbours	2650	0.495	1.251	0	12	1827	0.588	1.387	0	12
Cultural Distance	1414	0.022	0.028	0	0.158	939	0.023	0.027	0	0.158
Tradeshare	2650	43.619	38.782	0	100	1827	44.349	37.802	0	100
Colonial Dependency	2650	0.323	0.970	0	16	1827	0.419	1.115	0	16
Diplomatic Exchange	2650	5.830	9.025	0	84	1827	6.843	10.017	0	84
Waryieldlost	2650	0.065	0.386	0	6	1827	0.076	0.430	0	6

considerable variation in the year of the extension of franchise to females. Hence, these results are representative to a large extent. The list of countries in the largest sample and the smallest sample can be found in table 6.2 in the Appendix.

We analyze separately the countries that extended the female franchise before or after the median year of extension in the population. This way, we can determine whether effects differ between early and late adopters. As a robustness check, rather than taking the median year of adoption, we split up the sample between adopters before the end of WWII and its aftermath (1949), as this is often claimed to be the largest wave of democratization. With regards to causality, we admit that we cannot identify the direction of the relationship in this model and that the female franchise extension is likely endogenous to several covariates in our model. For example, countries may trade more with each other exactly because they share the same norms regarding female rights, and there may be other latent variables capturing similarity between countries, making them having more contact as well as adopting the institution around the same year in history. A standard approach in the literature to circumvent this problem is estimating the independent variable with a gravity model based on exogenous predictors. However, this would mean that, by and large, the same predictors would need to be used for the different contact variables. We aim to show that contact variables indeed matter for the adoption of female voting rights, beyond the standard within-country effects and external shocks, and to analyze whether the importance of different types of contact changes over time. We stress that these data only allow us to establish correlations. Identifying causal links between each of these types of contact individually and

the adoption of voting rights for women is something that could be taken up by future research.

3.5 Results

Table 3.3 shows the results for the complementary log-logistic model. The coefficients reported are hazard ratios, meaning that a number greater than 1 increases the hazard and a number smaller than 1 decreases the hazard for the event happening. We build the model by first estimating a baseline specification, including a constant term as well as the time variable that counts the years the country is at risk. In the second model,we focus on the effect of the contact variables, which have been added one by one to make sure that the coefficients were not affected by multicollinearity.

Focusing on the geographical variables first, Model 1 reveals that both the mean distance to other adopters as well as the number of close neighbours significantly affect the chance of adopting female voting rights. The coefficient for mean distance is rounded to 1, but is actually slightly larger than 1, implying a small positive effect of the mean distance to prior adopters on the chances of extending the franchise. This is counterintuitive, but this could be due to the nonlinear nature of this variable, as explained above. Therefore, as a robustness check, we substitute this variable for another measure of distance, cultural distance. Due to data limitations, this reduces the sample by half. The number of neighbours within 640km still positively affects the chances of extending franchise to females, but the other contact variables are no longer significant. Leaving out the distance variable altogether does not change the coefficients by much, but does make the effect of colonial dependence insignificant. Furthermore, the model with the best fit is the one with all six contact variables included.Therefore we choose model 1 as our main specification.

The effects of neighbours, trading with prior adopters, and having colonial ties with prior adopters are all in the expected direction, but having more diplomatic relations with prior adopters actually decreases the chance of adopting the institution. The effect of the mean distance to other prior adopters is also in the opposite direction as expected, but this could be due to the fact that initially, countries all over the world adopted the institution. The effect of this variable is also very close to unity: so, the economic significance is low. However, in this first model we clearly have an omitted variable bias, given that we do not yet control for other within-

Table 3.3: Complementary log-logistic regression of contact on women's suffrage adoption

	Women's Suffrage			
	Baseline	Model 1	Model 2	Model 3
Mean distance		1.000(0.000)***		
Cultural Distance			1314.354(0.147)	
Close neighbours		1.403(0.073)***	1.707(0.205)***	1.362(0.069)***
Tradeshare		1.009(0.004)*	1.006(0.253)	1.017(0.003)***
Colonial dependence		1.170(0.080)**	1.242(0.289)	1.112(0.081)
Diplomatic exchange		0.970(0.012)**	0.985(0.022)	0.977(0.013)*
War yielded/lost		1.205(0.287)	1.373(0.529)	1.159(0.274)
Time	1.005(0.003)	0.998(0.004)	1.004(0.006)	0.999(0.004)
Constant	0.022(0.008)***	0.142(0.006)***	0.010(0.006)***	0.012(0.005)***
No. of obs	2650	2650	1414	2660
Nonzero outcomes	98	98	46	98
AIC	839.958	**766.682**	384.532[1]	776.481

***=<.010, **=<.050, *=<.1, Standard errors between brackets.
[1]This AIC value cannot be compared to the other values as the number of observations varies

country factors or external shocks. So, these results should be interpreted with caution.

In Table 3.4 we extend our analysis by looking more carefully at the risk set. We assumed that all countries were at risk from the moment the first country extended franchise to women, but one could also argue that countries only become at risk when they are independent. Therefore, we also run the analysis with only those countries that were independent before 1893. This leaves 40 countries in the analysis. However, this assumption may be somewhat too strict, as non-independent countries could in fact adopt the institution as well. As a robustness check we include the year of independence. Furthermore, one could argue that the extension of franchise to all women only follows after the extension of franchise to all men. Hence, we also include the years that passed since full suffrage for men was introduced. As there is still no full suffrage in Qatar and Saudi Arabia, according to the Political Institutions and Political Events dataset, this reduces the sample by two observations.

Table 3.4 shows that the results for the contact variables are largely robust. The coefficients for the effects of close neighbours, tradeshare and colonial ties remain stable and highly significant in all specifications. The effects of geographical distance to prior adopters and diplomatic relations with prior adopters lose their significance when only looking at countries that were independent before 1893. Strikingly, losing a war or yielding to a prior adopter now is a positive and significant predictor of adopting female voting rights. This could be due to the fact that many of these relatively old and established countries were involved in the two world wars. Hence, they may have adopted female voting rights in either of the two waves that followed the two world wars. When including the year of independence as a control variable for all countries, this is a significant predictor of adopting female voting rights and it improves the fit of the model so we keep this variable as a control variable. The years since full male suffrage are not significant. Also, adding this variable does not improve the model.

So far we have only looked at contact and other variables that may determine whether countries are at risk to adopt the full female franchise institution. Now, we will examine whether the effect of contact is robust to including a large set of covariates. We do this in three steps. First, we introduce time dummies referring to three catalyst events: the ending of World War I, World War II and the Cold War.[10] The dummy variable capturing

[10]We also tried different lags for these dummies but this did not significantly change the results.

Table 3.4: Complementary log-logistic regression of contact on women's suffrage adoption including riskset controls

	Women's Suffrage			
	Model 1	Model 2	Model 3	Model 4
Mean distance	1.000(0.000)***	1.000(0.000)	1.000(0.000)*	1.000(0.000)*
Close neighbours	1.344(0.071)***	1.283(0.118)***	1.394(0.078)***	1.397(0.079)***
Tradeshare	1.010(0.004)***	1.023(0.008)***	1.010(0.004)***	1.010(0.021)***
Colonial dependence	1.166(0.081)***	1.127(0.077)*	1.213(0.090)***	1.207(0.093)***
Diplomatic exchange	0.970(0.012)***	0.998(0.018)	0.972(0.013)***	0.972(0.130)***
War yielded/lost	1.212(0.292)	1.579(0.359)**	1.194(0.276)	1.190(0.274)
Year of Independence			1.001(0.000)***	1.001(0.000)***
Years since male suffrage				0.998(0.792)
Time	0.996(0.004)	1.013(0.008)	1.004(0.005)	1.004(0.005)
Constant	0.016(0.006)***	0.001(0.000)***	0.001(0.001)***	0.001(0.001)***
No. of obs	2482	1917	2482	2482
Nonzero outcomes	96	40	96	98
AIC	735.442	344.487[1]	**729.586**	731.516

***=<.010, **=<.050, *=<.1, Standard errors between brackets.

[1]This AIC value cannot be compared to the other values as the number of observations changes, so these are not nested models

the end of the Cold War was omitted from all specifications, however, due to lack of variation. Second, we add time-dependent and time-invariant within-country factors, that may affect the chances of adopting female suffrage. One of those variables is the effect of gender equality, which we will test with two different proxies: the absence of historical plough agriculture, which is argued to lead to persistent differences in gender equality between countries, or a contemporary indicator of gender equality. Finally, we include continent fixed effects. The model with all covariates included reduces the sample to 58 countries, so this smaller sample is used.

As can be seen from Table 3.5, the effect of neighbours extending the franchise is very robust across all specifications. If one extra neighbour adopts the institution in a given year, the focal country is 1.4 times as likely to do that as well. The same holds for the effect of having colonial ties with other prior adopters, which is also stable, positive and significant across all specifications. The effects of trading with prior adopters, having diplomatic exchanges with prior adopters and the mean distance to prior adopters are not robust at the 5% alpha level. The effects of the catalyst years are clearly important, positively affecting the chance of adopting female suffrage. When adding the country-specific variables, only the effect of being involved in a war is strongly significant and positive. This confirms previous findings that the position of women is strengthened when the country is at war. The extent to which the plough was used in historic societies is predicted to be related to gender equality as well, as plough agriculture requires physical strength and therefore creates traditional gender roles. This effect is confirmed, but disappears when continent fixed effects are added. This could be due to the high correlation between absence of the plow and being an African country (the correlation coefficient is 0.69). As a robustness check, we included the contemporary gender inequality index. This variable is also not significant in the full model and does not substantially change the results of our variables of interest, so we only report the specification with traditional plough use here. The religion share variables are jointly significant, and the share of Buddhists, Hindus and Atheists in the country are also individually significant, positively affecting the likelihood of extending franchise to women. Being an autocratic or democratic country does not affect the likelihood of adopting the institution, and neither does being a more open country with regards to trade or migration, or being a more developed country in terms of GDP. The con-

Table 3.5: Complementary log-logistic regression of contact on women's suffrage adoption including controls

	Women's Suffrage			
	Model 1	Model 2	Model 3	Model 4
Mean distance	1.000(0.000)	1.000(0.000)	0.999(0.000)	0.999(0.000)
Close neighbours	1.435(0.110)***	1.423(0.109)***	1.329(0.116)***	1.265(0.123)***
Tradeshare	1.004(0.006)	1.000(0.006)	1.005(0.007)	1.003(0.007)
Colonial dependence	1.276(0.107)***	1.280(0.108)***	1.219(0.113)**	1.199(0.110)***
Diplomatic exchange	0.984(0.017)	0.982(0.017)	1.021(0.018)	1.033(0.018)*
War yielded/lost	0.810(0.405)	0.726(0.399)	0.553(0.346)	0.487(0.340)
Year of Independence	1.002(0.001)***	1.002(0.001)***	1.002(0.001)**	1.003(0.001)***
WWI		5.398(4.080)**	5.955(4.786)**	6.500(5.503)**
WWII		3.368(2.137)*	4.094(2.660)**	3.318(2.133)*
Openness			0.975(0.076)	0.936(0.088)
GDP per capita			0.999(0.000)	0.999(0.000)
Common law			2.282(2.075)	2.236(2.241)
Gender Equality/Plow Absent			9.947(5.637)***	1.100(0.866)
Ongoing war			2.631(0.964)***	2.629(0.989)***
Democracy			0.992(0.027)	0.994(0.029)
Migrate			0.983(0.035)	0.976(0.034)
Religion Dummies			***	***
Continent Dummies				***
Time	1.005(0.007)	1.004(0.007)	1.013(0.008)	1.033(0.010)***
Constant	0.000(0.000)***	0.001(0.000)***	0.000(0.000)***	0.000(0.000)***
No. of obs	1955	1950	1950	1950
Nonzero outcomes	58	58	58	58
AIC	499.111	497.017	481.769	**469.394**

***=<.010, **=<.050, *=<.1, Standard errors between brackets.

tinent dummies are also jointly significant.[11]

When comparing the nested models, the specification with all the control variables is the preferred model, having the lowest Akaike Information Criterion. Therefore we can conclude that even when controlling for external shocks and within-country factors that are said to influence the chances of extending the franchise to women, contact is still an important predictor. Having more neighbours that are prior adopters, colonial ties with prior adopters and more diplomatic exchanges with prior adopters all positively affect the chance for a country to also adopt the institution in a given year. Furthermore, the set of contact variables is jointly significant in the full model at the 1% alpha level.

3.5.1 Changing effects over time

The effect of certain types of contact may change over time. To examine this, the sample is split into two subgroups: early and late adopters. The cutoff point is the median year of adopting full female voting rights in the full population, 1960. As a robustness check, we split up the sample in two groups using the theory about waves of franchise extension, in this case using the end the second wave of democratization (1949), as the cut off point. Table 3.6 displays the result for both subgroup analyses. The most striking finding is that the effect of having more neighbours that are prior adopters either becomes insignificant or negative for the early adopters. This is the variable that was most robust across all the different specifications discussed above. It has to be noted we only have 15 early adopters when using the 1949 cutoff. Hence, this result should be treated with caution. However, it is safe to say that in the early stages of franchise extension, the relationship between neighbours' and own adoption is much weaker than for the late adopters. This is quite an intuitive result, as diffusion forces tend to cumulate over time when there are more other adopters present. The results for trade share with prior adopters are interesting as well, as trading more with prior adopters seems to positively predict adoption for early adopters, but not for late adopters. The effect of diplomatic exchanges with prior adopters is positive for both early and late adopters when 1949 is the cutoff point, but insignificant when 1960 is the cutoff point. This may mean that something particular is going on in the period 1949-1960. This would warrant further study, looking more in depth into the types of exchanges and geopolitical relationships between countries. This is beyond

[11] We also tried to run separate regressions for each continent, but these models cannot be identified as the number of observations is too low.

Table 3.6: Analysis split into early and late adopters

	Women's Suffrage			
	Early (<1960)	*Late (>1960)*	*Early (<1949)*	*Late (>1960)*
Mean distance	1.000(0.000)***	1.000(0.000)	1.000(0.002)	0.999(0.000)
Close neighbours	0.818(0.188)	1.494(0.272)**	0.056(0.067)**	1.370(0.174)**
Tradeshare	1.026(0.012)**	0.980(0.010)*	1.147(0.067)**	0.993(0.008)
Colonial dependence	1.174(0.330)	1.033(0.150)	1.189(0.856)	1.180(0.139)
Diplomatic exchange	1.057(0.081)	1.042(0.027)	1.520(0.325)*	1.046(0.022)**
War yielded/lost	0.555(0.417)	Omitted	0.619(1.459)	0.504(0.559)
Year of Independence	1.002(0.002)	1.006(0.002)	1.011(0.005)**	1.004(0.001)
WWI	8.505(7.733)**	Omitted	13.927(20.789)*	Omitted
WWII	2.853(2.102)	Omitted	5.654(6.695)	Omitted
Openness	24.013(34.645)**	0.789(0.112)*	37.645(94.806)	0.829(0.112)
GDP per capita	1.000(0.000)	1.000(0.000)	0.999(0.001)	4.914(2.053)
Common law	0.487(1.170)	Omitted	779.671(4122.386)	Omitted
Gender Equality/Plow Absent	104.651(153.516)***	0.169(0.215)	2.127(8.627)	0.680(0.589)
Ongoing war	1.385(1.034)	5.815(3.293)***	0.016(0.054)	4.914(2.053)***
Democracy	1.079(0.061)	1.011(0.059)	2.016(0.646)**	0.998(0.041)
Migrate	0.882(0.132)	0.928(0.044)	1.663(2.353)	0.947(0.038)
Religion Dummies	**	-	-	***
Continent Dummies	-	***	-	***
Time	0.983(0.016)	1.061(0.023)***	1.003(0.104)	1.048(0.014)***
Constant	0.000(0.000)	0.000(0.000)***	0.000(0.000)**	0.000(0.000)***
No. of obs	920	998	593	1331
Nonzero outcomes	29	29	15	43

***=<.010, **=<.050, *=<.1, Standard errors between brackets.

the scope of the current study.

We observe a few interesting results in the control variables as well. The effect of being involved in a war is only significant for the late adopters, suggesting that the wars before 1960 did not do as much for the position of women. Furthermore, being a more open country with regards to trade has a large positive effect for early adopters, but not for late adopters. We also notice a sign switch for the gender equality variable, positively affecting the chance of adoption for early but not for late adopters.

Summarizing, the pure diffusion effect where a country that has more neighbours that adopt female voting rights is associated with higher chances of adopting the institution as well seems to be driven mainly by the late adopters in the sample. The evidence for the effect of contact in general is strongest for these late adopters, although being more open to trade is associated with higher chances of adopting female voting rights for early adopters.

3.6 Conclusion

We analyzed the adoption of full voting rights for women across countries. While female suffrage has been extensively researched, the contribution of this study is to see this as a process of diffusion, and to focus on the effects of cross-country contact on the timing of adoption. The effects of distance and contiguity, as well as four types of contact were analyzed. Furthermore we examined whether the relative importance of the different types of contact and their effect changed over time. We hypothesized that as the mean geographic distance to all prior adopters decreases, the probability of establishing full female voting rights increases. Furthermore, as the number of neighbours within a radius of 640 kilometer that are prior adopters increases, the probability of extending franchise to women increases. The first hypothesis was not confirmed: the mean distance to prior adopters was not significant in most specifications, even negative in some models. Hence, as mean distance to other adopters increases, the chances of adopting the institution for a country in a given year increase. This is a counterintuitive result. However, the coefficient of this variable is in all specifications very close to unity, implying that the economic significance of this result is negligible. Moreover, the non-linear nature of this variable over time makes it hard to interpret the coefficient. Take the following example. For European countries, initially the distance to other prior adopters was very large, driven by the distance to New Zealand. Still, many of those

European countries were early adopters. We tried an alternative specification where we looked at cultural distance to prior adopters rather than geographical distance, but this variable was not significant, substantially reducing the number of observations. The second part of the hypothesis was confirmed, with stable positive coefficients that were robust across all specifications. When the number of neighbours of a given country within 640 km that are prior adopters increases by 1, the likelihood of the focal country extending the franchise in that year is approximately 1.3 times as large.

The second hypothesis argues that having a larger share of trade volume with prior adopters increases the probability of establishing full voting rights for women in the country at risk. This was confirmed when studying only contact, but turned out not to be robust to including within-country control factors and general shocks.

The third hypothesis relates to the effect of war. We argued that losing a war or yielding to a country might lead to the forceful imposition of this latter's country's institutions upon the losing party —in this case, voting rights for women. This hypothesis was not confirmed in the analysis. However, being involved in a war in general does positively influence the probability of adoption, and this is a large and robust effect. A speculative explanation is that the channel is indirect by affecting the position of women in the country, rather than the direct channel where victors impose their institutions on losers. Furthermore, war creates a period of instability, during which the diffusion of new norms and values disrupts the traditional equilibrium, which is almost impossible during periods of peace due to institutional inertia (Zhao and Cao, 2010).

With regards to diplomatic exchange with prior adopters, it is hypothesized that through coercion this would lead to earlier adoption of the institution. This positive effect was confirmed when looking only at contact variables, but is not robust to the inclusion of within-country control variables and external shocks. We also tested whether the year of independence and the years that have passed since the extension of franchise to all males alter our results. The first variable is robust across all specifications, but the latter is not. If we would be strict in assuming that only independent countries before 1893 are at risk starting in 1893, our results still carry through, but this assumption seems unrealistic.

Within-country factors seem to matter less than previously expected in prior work. These variables are jointly significant, but only the effect of being involved in a war is robust across all specifications. Historical absence of plow agriculture, a measure for persistent gender equality, positively in-

fluences the chance of adopting female suffrage, but this effect disappears when continent fixed effects are added. These continent dummies, as well as the share of each religion in a country, are jointly significant but do not affect the magnitude or significance of our important findings. External shocks are important, as was already shown in previous studies. Particularly, extension of the franchise to women went in waves: the dummies for the 'catalyst years' 1919 and 1946 had large positive effects on the probability of adopting the institution in most specifications. However, even when these catalyst years are controlled for, contact variables still have a significant effect.

We also split the sample into early and late adopters of full female voting rights. We concluded that the pure diffusion effect, as measured by neighbours that are prior adopters, is only important for late adopters. For early adopters, contact matters less, although being a more open country does positively affect the chance of adopting the institution. We also expected that 'unfriendly' contact would be relatively more important for early adopters, while 'friendly' contact would matter more for late adopters, as globalization causes countries to be more dependent on each other for trade and investment. However, we could not find evidence to confirm this hypothesis. In fact, trade matters more for early adopters, whereas being involved in a war increases the likelihood of adopting female voting rights only for late adopters.

In conclusion, while large external shocks such as world wars are indeed important catalysts for institutional change, and although within-country factors specific to the institution under analysis should not be overlooked, this study shows that cross-country contact is positively related to the diffusion of an institution and should therefore be included when analyzing institutional adoption processes. The process of diffusion seems to be strongest within the group of late adopters. Having prior adopters as neighbours is the most robust factor that promotes diffusion, but other types of contact do matter as well.

We acknowledge that this study has several limitations. Strikingly, in our study, the position of women in a focal country does not significantly affect the timing of adoption of voting rights for women. Although there is some evidence for the persistent nature of gender equality, future studies could attempt to find time-varying instruments for gender equality to see whether this has more prominent effects when the assumption that it is constant is removed. Another extension is the analysis of other institutions to explore whether the same types of contact matter for the process of diffusion. This way, research could draw comparisons on the effects of contact

between historical versus modern institutions, economic versus political institutions and more versus less embedded institutions. In so doing, our theory's external validity could be examined, probably leading to further theory development, and to specifying boundary conditions.

Another possible limitation of this study is that we do not consider the role of contact between two countries on third parties, or the existence of possible 'diffusion professionals' that do not adopt the institution themselves but do promote its advancement. This channel is hard to disentangle in cross-country empirical studies, but possibly case studies can be used for this. Finally, in the current specification it is difficult to identify causal links, as causality might run in the opposite direction and more similar countries may be affected by the same processes. To identify causal links, future studies should focus on one type of contact at the time, as the traditional gravity-equation approach cannot be applied to multiple contact variables that are predicted by the same factors. What we can conclude though, is that countries do not choose their neighbours, and that having more prior adopters among neighbours positively affects the chance of extending franchise to females. This is an important conclusion indicating that diffusion processes should not be overlooked when analyzing the adoption of institutions.

4

Searching for the existence of entrepreneurial ecosystems

A regional cross-section growth regression approach

Objectives

- To find empirical evidence for the way in which institutions may lead through differences in economic growth through the channel of entrepreneurial activity
- To find evidence for the existence of different institutional regimes or entrepreneurial ecosystems in European regions
- To propose a methodology through which institutional regimes may be empirically captured and later decomposed, at different levels of aggregations and in different samples
- To see if the institutional regimes moderate the relationship between entrepreneurial activity and economic growth
- To allow for heterogeneity in the effect of entrepreneurial activity on economic growth between countries but also between supernational

This chapter is co-authored by Mark Sanders, Marc Schramm and Niels Bosma and is available as a TKI Working Paper.

classes of regions that may differ on other characteristics

4.1 Introduction

The term entrepreneurial ecosystem was coined to refer to those elements in the entrepreneurs' environment that help them to succeed (or not) in their efforts to grow a new venture. Even though the analogy with the natural and biological ecosystems cannot be upheld in all details (Holling, 2001), the concept has proven very useful to signify those elements directly relevant to the entrepreneurial value chain (Adner and Kapoor, 2010; Spigel, 2015). The term 'ecosystem', however, also suggests a broader scope and extending the concept to the entire set of interdependent actors and factors coordinated in such a way that they enable productive entrepreneurship is in fact required (Stam, 2015). This encompassing definition has proven useful in theorizing about the importance of the context in which entrepreneurs operate and led both academics and policy makers to zoom in on the local, regional and national enablers of entrepreneurship (Audretsch et al., 2015). Measuring (the size and/or quality of) entrepreneurial ecosystems, however, continues to proves a challenge. A first problem with measuring the quality of an ecosystem is that we first have to establish the appropriate ecosystem boundaries.

In nature, ecosystems overlap. Climate is a global system that overlaps with more localized weather systems and the relevant ecosystem boundaries for the habitat of the eagle clearly span the ecosystems of many mice. Economic or entrepreneurial ecosystems share that property. The local entrepreneurial ecosystem is part of the larger regional, national, international and perhaps truly global ecosystem and although all economic activity is in the end local, there is no reason to assume that the most relevant ecosystem boundaries coincide with any administrative geographic units, however (dis)aggregated they may be. Instead, the relevant ecosystem boundaries depend on the question at hand. Do we study eagles or the mice?

The evidence that entrepreneurship contributes to economic growth through innovation and creative destruction seems mixed at the aggregate, regional and local levels (see below) and it has proven hard to convincingly identify causal links at all aggregation levels. The existence of heterogeneous overlapping and nested entrepreneurial ecosystems at all these levels can of course explain this.

We follow Stam (2015), who put forward that the entrepreneurial ecosys-

tem is multilayered. And propose an approach building on Andersson and Henrekson (2015), who first define (local) competitiveness as the ability to continuously renew the economic base of the local economy. Then they argue that entrepreneurship (at the local level) is instrumental in maintaining that ability. Following Baumol (1990) institutions, broadly defined as "the rules of the game" (North, 1990), will influence the (local) supply and direction of entrepreneurship. The relevant institutions Andersson and Henrekson (2015) then identify are clearly "located" at the national level (e.g. rule of law, ease of starting a firm, taxes, property rights, labour market institutions, social security) or the local level (e.g. local and regional taxes and regulations, presence of e.g. universities and research institutes and local entrepreneurship cultures and skills). Andersson and Henrekson (2015) also argue that one should be able to identify the effects of local entrepreneurial ecosystems in the different effect entrepreneurial activity has on growth across localities, regions and countries. They conclude that their arguments: "*[...] suggest that the magnitude of local multipliers and growth effects associated with the local presence of entrepreneurial and knowledge-intensive activities are a function of the local institutional environment. This is a hypothesis that remains to be rigorously tested in empirical work.*" (Andersson and Henrekson, 2015, p. 189)

In this study we aim to take a first step in empirically testing that hypothesis. More specifically we test whether the contribution of entrepreneurship differs significantly across regions. If we may refer to the "local institutional environment" as the ecosystem, we focus on the predicted effect of local entrepreneurial ecosystems on the local multipliers/growth effects. If such effects exist, the marginal effect of entrepreneurial activity (measured in several standard ways) on local economic growth would vary systematically with the quality of the (unobserved) entrepreneurial ecosystem. By investigating at what level of geographic aggregation the marginal impact of entrepreneurial activity on growth shows this predicted variation, we can show whether entrepreneurial ecosystem quality matters most at the regional or national level when economic growth is our variable of interest.[1]

We estimate a standard growth regression, both in an OLS specification and in a multilevel specification for 107 European regions. We find that a multilevel model with different intercepts (average growth rates) per country fits our data best. That specification is preferred over the OLS and mul-

[1] Unfortunately our data do not yet allow us to investigate this at the local level, but our approach can easily be applied to such more disaggregated data as well.

tilevel specification with varying slope coefficients across countries. In the OLS regression, the average rate of those entrepreneurs that expect high job growth is a positive predictor of economic growth. In the multilevel model the effect of high job-growth entrepreneurship is weaker, only significant at the 10% alpha level at the regional level, and the country-mean of entrepreneurship does not affect regional growth. To see whether any heterogeneous effect of entrepreneurship exists, we use residuals of the different specifications, either controlling for unobserved country differences and without controlling for those. We find no evidence, circumstantial or otherwise, for entrepreneurial ecosystems operating either at the regional or national level.

As other research suggests that the local context does matter and entrepreneurs may have a key role in shaping and developing local entrepreneurial ecosystems (see e.g. Feldman (2014)), we suspect that our data is geographically too aggregated to distinguish the differential impact of local entrepreneurial ecosystems. Our NUTS-1 and 2 level regions may simply be covering diverse local ecosystems, obscuring and confounding the differences that exist within these regions. Given the limited granularity available in the GEM-data we use for our entrepreneurship variables, however, this is not something we can resolve in this cross country study. Our main contribution is the empirical method we propose. Therefore we very much encourage those with access to more disaggregated data to repeat our analysis and gladly make our Stata-code available to them on request. In the current data set we must conclude that we fail to find support for any theory that predicts different marginal effects of entrepreneurial activity on growth across (groups of) NUTS-1/2 EU regions. But the search should continue and we hope our method contributes to that.

The remainder of this chapter is structured as follows. In the next section, we position our study in the extant literature. Section 3 presents our empirical strategy, briefly derives the main equations we estimate and motivates the controls we add to the standard cross country growth regression model. Section 4 presents our data and shows the empirical results. We conclude with a discussion of our results and agenda for further research.

4.2 Entrepreneurial ecosystems, institutions and regional growth

Theoretical interest in the role of entrepreneurship in creating economic growth goes far back. Schumpeter (1912) already put entrepreneurship

center stage in his Theory of Economic Development by coining the term "creative destruction". In more modern endogenous growth models, such as those developed by Romer (1986) and Lucas (1988), there is no role for entrepreneurship, as they passively commercialize all knowledge that is generated. But Aghion and Howitt (1992) created a model where "entrepreneurs" both generate the knowledge and appropriate entrepreneurial rents through entry, while Michelacci (2003) and Acs and Sanders (2012) distinguished between invention and innovation and show that in such a model entrepreneurship is complementary to knowledge creation and contributes to growth.

The empirical evidence in support of such predictions, however, remains mixed (e.g. Van Praag and Versloot, 2007). As Glaeser et al. (2010, p. 3) stated: "While it would be hard to imagine a world in which an abundance of entrepreneurs did not strengthen the local economy, the literature documenting this effect is still in its infancy". Empirical results to date are inconclusive for three reasons. First, there is a measurement problem of entrepreneurship, which led to the use of many different proxies for entrepreneurship (Hechavarria and Reynolds, 2009; Stenholm et al., 2011). Studies using varying proxies for entrepreneurship (Audretsch, 1995; Carree and Thurik, 1998, e.g.) have found a positive relation overall, but often these results are multi-interpretable.[2]

Second, and closely related, there is heterogeneity within entrepreneurial activity and the effect on growth is not equal for all types of entrepreneurship (Santarelli and Vivarelli, 2007; Stam, 2009). Research makes a distinction between ambitious entrepreneurship or gazelles (Henrekson and Johansson, 2010; Stam et al., 2012) and businesses started at the local level that create little employment growth or economic expansion. E.g. Wennekers et al. (2005) find a U-shaped relationship between entrepreneurial activity and economic development, proxied by GDP per capita levels. This finding was explained as labour moving from unproductive entrepreneurship into employment in early stages of development (higher GDP/capita reduces entrepreneurship) and from employment into growth enhancing innovative entrepreneurship (entrepreneurship causing GDP/capita growth) as development increases.

Third, the two-way causality already referred to above already indicates

[2]For example Glaeser et al. (2010) and Carree and Thurik (1998) find a negative correlation between establishment size and employment or productivity growth, respectively. And one of the possible explanations is that small firms are a sign of more entrepreneurship and new startups. But there could be other explanations such as an increase in competition (Nickell, 1996; Nickell et al., 1997).

that the effect of entrepreneurship on growth is hard to isolate empirically (E.g. Baumol, 1990; Carree and Thurik, 2010; Stephens and Partridge, 2011; Fritsch and Wyrwich, 2014). Entrepreneurship both causes and is caused by GDP (growth), creating an identification problem. In sum, the literature is only beginning to build a consensus on how to measure entrepreneurial activity, what outcomes to expect from what types of entrepreneurial activity and how to properly identify such effects. But even if that consensus were achieved, the concept of "entrepreneurial ecosystems" implies that the same entrepreneurship, accurately measured and properly instrumented, should be expected to affect economic growth differently in different contexts. The entrepreneurial ecosystem concept stresses the complex interactions between various environmental factors and entrepreneurial activity in creating economic dynamics. As such the concept is closely related to literatures on the national (Freeman, 1987; Lundvall, 1992; Freeman, 1995; Edquist, 1997) and regional systems of innovation (Autio, 1998; Cooke, 2002; Doloreux, 2002) as well as the more mainstream empirical growth literature that is linking economic growth and development to institutional quality (e.g. Acemoglu et al., 2001; Easterly, 2001; Sokoloff and Engerman, 2002; Nunn, 2008). These strands of literature share a focus on the enabling conditions for Schumpeter's "gales of creative destruction" (1942). It is beyond the scope of this chapter to survey all that literature here. Instead we restrict ourselves to establishing the connection with more mainstream empirical growth literature.

The main problem in this literature is that institutions develop endogenously with economic growth, and therefore, to identify and isolate the effect empirically, the researcher has to find an exogenous instrument for institutional quality (e.g. Acemoglu et al., 2001). But even if such a satisfactory instrument can be found at the regional and local levels, there is still the problem that no satisfactory measure of institutional quality exist. This problem, we argue, will also haunt empirical research into the impacts and relevance of the entrepreneurial ecosystem. In fact, both concepts are multidimensional complex systems that prove too elusive to be measured directly in a one dimensional, cardinal number. The proxies that have been used for institutional quality, such as indices of property rights protection, economic freedom and corruption, inevitably only "measure" a part of the relevant institutional complex and not even accurately so (Rodríguez-Pose, 2013), based as they often are on expert opinion and surveys. The same holds for measuring the quality of the entrepreneurial ecosystem. Elaborate indices as for example proposed by Acs et al. (2014); Szerb et al. (2015) do give us a useful first cross regional comparison of ecosystem quality.

But such composite indices, by the way they have been constructed, have the disadvantage that they cannot be used as independent variable to properly identify the moderating effect of ecosystem quality. It will prove even more complicated to find suitable instruments for such complex multidimensional indices if we aim to use them as an independent variable in cross sectional growth regressions.

Here we therefore propose an alternative strategy for "measuring" the effects of the entrepreneurial ecosystem. To do so we exploit the assumption that the entrepreneurial ecosystem affects output and growth *through* its effect on entrepreneurial activity, which we can observe. That is, assuming the entrepreneurial ecosystem matters for growth, its quality will be systematically correlated with growth and works as a *moderating* variable on entrepreneurial activity. To give some examples of such moderating mechanisms: better protection of property rights increases the future expected private gains from entrepreneurship. This may attract more talented entrepreneurs to the region (or country) to start their own business (see e.g. Johnson et al. (1999), Sobel (2008)). Another example is an improvement of financial institutions, enabling a flow of funds from investors to potential entrepreneurs such that they can undertake larger and more ambitious projects, thereby also increasing the effect of entrepreneurial activity on growth (e.g. King and Levine, 1993). Informal institutions can also provide a moderating mechanism; an entrepreneurial culture, trust, social capital and a dense local knowledge network are all likely to contribute to more productive entrepreneurship in a region or locality (Bosma, 2013; Andersson and Henrekson, 2015). Our approach thus distinguishes itself by researching a channel through which the moderating effect of the entrepreneurial ecosystem on growth will become visible in variation in the contribution of observed entrepreneurial activity to growth.

We conclude from the above review that entrepreneurship takes on many shapes and manifests itself in different contextual settings. Consequently, the relation between entrepreneurship and growth is heterogeneous and not all types of entrepreneurship create economic growth to the same extent in different institutional contexts. As we are not primarily interested in the exact value or even sign of the marginal effect of entrepreneurship itself, however, we may side-step these issues for now. We make the case that, even if the estimated marginal effect itself is inefficient or biased due to measurement error, impact heterogeneity and reverse causation, possible differences in the marginal effect of entrepreneurship on growth are caused, *by definition* by differences in the regional entrepreneurial ecosystem. In other words, we assume that the presence of an

entrepreneurial ecosystem can be deduced from the way it transforms locally available entrepreneurial talent into activity and ultimately economic growth. Thus defined, the first step that is needed to prove the existence of an entrepreneurial ecosystem at the aggregation levels investigated is to show that the marginal effects differ significantly across geographies. Our proposition is that the entrepreneurial ecosystem is what (ever) moderates the effect of entrepreneurial activity on growth. Of course, we do need to consider relevant controls and specify our regional growth regression in such a way that we do not allow omitted variables that explain cross regional variation in growth and might be correlated with entrepreneurial ecosystem quality to create the illusion of an ecosystem at work. Therefore all the usual suspects and controls that the literature suggests should be included. If multiple regimes can be identified with regards to the relationship between entrepreneurship and economic growth, we can introduce variables to predict regime membership, in order to find out what constitutes the entrepreneurial ecosystem and to rule out other possible explanations.

Capital formation as well as population growth obviously play a role, in the neoclassical growth model and in regional growth accounting (Ciccone, 2002). E.g. Sala-i Martin (1996) also finds evidence for beta-convergence between regions in the USA, Canada, Japan and Europe, indicating that initial income should be controlled for in regional growth regressions. However, the largest share of variance in growth across regions has been attributed to productivity differences and innovation. In empirical work on regional growth we can then distinguish i) the linear model ii) the innovation-systems or learning-region model and iii) the knowledge spillovers model. Rodríguez-Pose and Crescenzi (2008) are the first to have nested these models in one empirical specification. They find that knowledge spillovers indeed happen. But they also show that the transmission of tacit knowledge strongly decays with distance, supporting the innovation systems approach. Furthermore they find that R&D expenditures only lead to actual innovation and growth if a) a region can also benefit from extra-regional spillovers, measured by R&D in neighbouring regions and b) the territorially embedded institutional environment is conducive to the diffusion of innovation. These results can be related to the geographical clustering literature (Jaffe et al., 1993; Breschi and Lissoni, 2003; Bosma et al., 2008) that links the success of certain regions to the formation of a cluster within which companies of the same sector thrive and benefit from agglomeration advantages and knowledge spillovers. Frenken et al. (2007) show that such spillovers are stronger for "related varieties" in "smartly specialized" regional clusters.

Reynolds et al. (1994) find that urbanization promotes spillovers. A result that Bosma (2009) confirmed in a multilevel analysis for European regions as also presented in our study.

From the regional growth literature we take away that it makes sense to control for regional population density, age structure and industry diversification (Dogaru et al., 2011; Thissen et al., 2011). Other variables, such as the degree of urbanization, labour productivity, the level of human capital, the investment in R&D or the level and type of competition, have been mentioned in the literature as possible moderating variables on the effect of entrepreneurship on (job) growth (Fritsch and Mueller, 2008; Fritsch and Schroeter, 2011; Unger et al., 2011; Fritsch and Noseleit, 2013b,a).

Our strategy, however, involves first estimating a clean growth residual. That is, we estimate growth that is not explained by entrepreneurship. Including variables that are complement to and therefore possibly endogenously correlated with entrepreneurship, would then be undesirable here. It therefore seems inappropriate to include these variables in the first stage of our analysis. Rather, these variables would become relevant if we find different classes of regions as they would then potentially explain the class membership in the second stage.

The above contributions led us to expect that at the regional level entrepreneurship is positively correlated with growth at the regional level. There are also reasons to expect this effect to differ across regions. Helsley and Strange (2011) for example show that urbanized cultures are often more conducive to growth-oriented entrepreneurship. When looking at regional performance in general, Eriksson et al. (2013) have distinguished between regions with a 'business climate' and regions with a 'people climate', finding that both matter for regional performance, but through different channels. Audretsch et al. (2012) classify regions in an explicit way, using the functional specialization theory. In a sample of German regions they use regional characteristics to determine the industry structure, and thereby the specialization into either production or innovation. They find five groups of regions, Industrial Districts, Periphery, Industrial Agglomeration, Urban Agglomeration and Urban Periphery. The propensity to start a business is highest in groups 1, 4 and 5, and these types of regions are thus said to have an entrepreneurial regime. These papers, however, focus entirely on the effects of specific aspects of the ecosystem on observed entrepreneurial activity. Our study investigates the existence of latent regional ecosystem characteristics that affect the contribution to growth of such early stage entrepreneurship (that is, conditional on observing such activity). We can summarize this in the following hypothesis:

Hypothesis. *Entrepreneurship is positively correlated to regional growth, but this relationship is moderated by the regional entrepreneurial ecosystem and therefore differs across regions. The quality of the entrepreneurial ecosystem cannot be measured directly, this moderation will manifest itself through classes of regions that differ with respect to the strength and possibly the sign of this relationship.*

4.3 Empirical strategy

In our aim to localize the moderating effect of entrepreneurial ecosystems on the relationship between entrepreneurial activity and economic growth, keeping the ecosystem as a latent concept, we employ a three-step empirical process. First we need to account for economic growth caused by the 'usual suspects' at the national level (Mankiw et al., 1992). We extend their standard growth model with entrepreneurship and the relevant regional controls in an OLS estimation to establish the effect of entrepreneurship on growth at the regional level. Next we estimate this model in a multilevel specification. Paterson and Goldstein (1992) and Leckie et al. (2010) show that multilevel models are very suitable when a complex web of influences at different levels of layers are suspected. This is definitely the case when looking at entrepreneurial ecosystems and formal and informal institutions at the regional and national level in Europe. Estimating our regression in a multilevel model allows us to distinguish between variance in growth rates at the regional and national level, and the extent to which entrepreneurial activity can explain this variance. We then exclude the entrepreneurship variables in both growth regressions, and regress the residuals on our entrepreneurship variable in an unconditional latent class specification. That allows us to test the hypothesis that the marginal contribution of entrepreneurship to (unexplained) growth that differs systematically between (groups of) regions that do not necessarily coincide with countries. This somewhat roundabout way implies we do not allow the coefficients on the usual suspects in the standard growth regression to differ across (classes of) regions. If we find significant latent classes in this approach this proves that it is the contribution of entrepreneurship to growth differs significantly between endogenously determined groups of regions.

The growth literature has emphasized the importance of considering all relevant variables in a growth regression (Sala-i Martin, 1994). We base our model on the Mankiw et al. (1992), hereafter MRW model and we estimate a cross-regional growth regression including the "usual suspects", adding also relevant regional controls. The regression equation can be de-

rived from the production function:

$$Y(t) = K(t)^{\alpha} H(t)^{\alpha} (A(t)L(t))^{1-\alpha-\beta} \tag{4.3.1}$$

where Y is output, K is capital, L is labour, H is the stock of human capital and A is the level of labour augmenting technology.[3] From this specification we can derive (see Appendix) the MRW specification for cross sectional growth regression:

$$\frac{lnY_i(T) - lnY_i(T0)}{T - T0} = \alpha_0 + \alpha_1 s_i^k + \alpha_2 s_i^h + \alpha_3 n_i + \alpha_4 lnY_i(T0) + \epsilon_i \tag{4.3.2}$$

where i indexes regions, s_i^k and s_i^h are the average shares of income invested in human and physical capital (in the period T0-T) respectively, n is the average growth rate of the population over that period. In any cross-country growth regression, endogeneity may be an issue. We follow MRW in assuming that the rates of saving and population growth are independent of country-specific factors shifting the production factors, meaning that they are independent of ϵ and that we can estimate equation 4.3.2 with ordinary least squares. This is common in growth regressions and the assumption is met when saving and population growth are endogenous but preferences are inelastic. These models build on solid theoretical foundations, but we have to acknowledge that endogeneity still makes interpreting causal effects hard.

By including initial income regions are allowed to be out of the steady state and we can test for convergence. The disadvantage is that permanent productivity differences in the production function, $A_i(0)$, across countries and regions end up in the error term. We control for these differences by estimating the country fixed effect in a multilevel model, as in equation 4.3.3.

$$\frac{lnY_{ij}(T) - lnY_{ij}(T0)}{T - T0} = \alpha_0 + \alpha_{0j} + \alpha_1 s_{ij}^k + \alpha_2 s_{ij}^h + \alpha_3 n_{ij} + \alpha_4 lnY_{ij}(T0) + \epsilon_{ij} \tag{4.3.3}$$

where j indexes the country and the specification allows for country specific intercepts α_{0j}. The error term, ϵ_{ij}, can be decomposed in a fixed and

[3] In a Cobb-Douglas it does not matter if we specify technology as purely labour augmenting, purely capital augmenting or as total factor productivity as the elasticity of substitution is 1. This is easily verified by defining $B_i \equiv A_i^{1-\alpha-\beta}$, where B_i is then total factor productivity.

a random part, which can be used to calculate between and within country variation using the Intra Class Correlation (ICC) (Snijders and Bosker, 2012). This gives us insight into how similar regions within a country are with regards to the level of entrepreneurship. The ICC is the ratio of the variance between countries to the total variance, so it tells us how much of the variance in economic growth is accounted for by clustering on the country-level.

We first extend the OLS growth regression by adding the entrepreneurial variables of interest. The regression equation that we estimate becomes:

$$\frac{lnY_i(T) - lnY_i(T0)}{T - T0} = \alpha_0 + \alpha_1 s_i^k + \alpha_2 s_i^h + \alpha_3 n_i + \alpha_4 lnY_i(T0) + \alpha_5 E_i + \epsilon_i \tag{4.3.4}$$

where E_i is our proxy for entrepreneurship.[4] By comparing the results from this regression to a multilevel specification (that is equivalent to the one in equation 4.3.3 but we do not write out to save on notation), we can test the hypothesis that the marginal effect of entrepreneurship on growth, α_5, by controlling for the usual variables included in this type of research, is different across countries and/or regions within countries. The MRW model was developed to model cross-country differences in growth rates by differences in physical and human capital and initial income. However, from the regional growth literature, we know that several other factors are important when explaining regional growth. These main factors are population density industry structure and the demographic structure of the region and they are found to aid the conversion of innovation to growth at the regional level. Population density is a commonly used proxy to control for the effect of urbanization economies. The industry structure is important because related variety has been found to contribute to regional growth. Finally, the demographic structure tells us something about the share of working-age individuals and possible emigration of young individuals.

We then take the residuals from regressions 4.3.2 and 4.3.3 and regress these on regional entrepreneurship in a latent class specification that allows for endogenous sorting into regimes $r(= 1, ...R)$.[5]

[4] We test three.

[5] Our data do not allow for running the full model in a latent class specification as we would have to estimate too many parameters on too little data and exhaust our degrees of freedom. With this two step procedure we assumed that the slope coefficients on traditional production factors are identical across regions and countries and investigate the effect of entrepreneurship on the estimated Solow-residuals or total factor productivity growth.

$$\begin{aligned} \epsilon_i &= \beta_{0r} + \beta_{1r} E_i + \varepsilon_i \\ \epsilon_{ij} &= \beta_{0r} + \beta_{1r} E_{ij} + \varepsilon_{ij} \end{aligned} \tag{4.3.5}$$

where $r = 1, ..., R$ indicates the regime and R refers to the (exogenous) total number of regimes. Each regime, r has its own β_0 and β_1. We can then test for the existence of latent classes in the data following Greene (2007) who suggests a "test-down" strategy to identify the correct number of regimes. By regressing the residual on entrepreneurship only, we restrict this analysis to finding groups of regions that differ in the marginal contribution of entrepreneurship to unexplained, residual growth variation. We want to stress that the estimated coefficients are likely to be biased by endogeneity (entrepreneurship is endogenous to GDP growth) and omitted variable bias (our specification is parsimonious and variables, possibly correlated with the entrepreneurial ecosystem quality, such as religion and geography, are not included). For the purpose of this research, however, we are only interested in the possible clustering of data points the latent class model would suggest. In the next section we describe our data and present our results.

4.4 Data and results

4.4.1 Data

Our data comes from Eurostat (2016) and GEM (2016), as is described in table 4.1. Our final sample consists of 107 regions in 16 European countries. The bottleneck for our data is the GEM data on entrepreneurship. To get reliable measures of regional entrepreneurship, this data had to be gathered during a 5 year period and mostly at the NUTS-2 level. For some countries this still did not provide enough observations so the aggregation was scaled up to NUTS-1. In the GEM data set there are 128 regions in 18 countries and some city regions have been identified at the NUTS-3 level. We reaggregated the GEM data such that these cities are also included in the NUTS-2 region in which they are found, this decreases the number of regions slightly but is necessary for consistency. Furthermore, Switzerland and Croatia are in the GEM data set but the first is not part of the European Union and the latter only became a member in 2013, implying that there is no comparable data on the control variables for these countries that can be used together with the available data from Eurostat. A list of all the countries and regions in our sample can be found in the Appendix. Our sam-

ple is a good representation of regions in Western, Northern and Southern Europe. Eastern Europe is somewhat under-represented as many Eastern European countries only joined the European Union in the last few years.

We collected data on GDP and the variables in the MRW growth regression for the period 2001 and 2014, or as long as our data sources allowed. As it is a cross-section analysis, the variables that change over time are averages over these years.

Our dependent variable is computed from regional GDP data for each year between 2006 and 2014, taking the difference between the logarithm of GDP in 2006 and 2014 and dividing that by the number of years, $\Delta lnY/(T - T_0)$.
We also take the logarithm of initial GDP, which is the year 2005, $lnY(2005)$. For the share of physical capital investment in income, s^k, we take the average investment rate, that is the average of the values of Gross fixed capital formation divided by the GDP in that year. At the national level, this is found to be around 0.3, in our data set the mean is 0.21. Our physical capital data comes from Eurostat (2016).[6]

For the investment in human capital, s^h, MRW propose a variable "school", which is formed by multiplying the fraction of the 12-17 aged population that is enrolled in secondary school by the fraction of the working age population that is in that age bracket, to proxy for the time spent in secondary education. For most European regions, however, secondary education is obligatory until the age of 16 or 17, so we cannot expect much variation there. Therefore, we take the share of the working-age population aged 20-24 and multiply it by the enrolment rates in tertiary education (ISCED 5-6) as % of the population aged 20-24. This variable could be seen as a proxy for foregone earnings on the part of the working-age population that is in school. For n we take the logarithmic average population growth.[7]

As the literature suggests not all types of entrepreneurship make the same contribution to growth, we also use three measures of entrepreneurship. Our entrepreneurship variables are taken from the Global Entrepreneurship Monitor (GEM, 2016), that conducts representative surveys among the

[6]For all the averaged variables, we take the average of all available years in the sample period for that region.

[7]In contrast to Mankiw et al. (1992) we do not include g, the average growth rate of technical progress or δ, the average depreciation rate in our regression. Mankiw et al. (1992) assume these variables to be constant across all countries in their data set. That assumption would then hold *a fortiori* for EU-regions and including these constants is only important if the actual size of the estimated coefficients is of interest. We are mostly interested in the between regional variation in the marginal effect of entrepreneurship, which is unaffected by the constant.

Table 4.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N	Unit	Year
ΔlnY	1.278	1.406	-2.6095 (GR: Kentriki Ellada)	6.463 (NOR: Vestlandet)	107	% per annum	2006-2014
Initial GDP ($lnY(2005)$)	10.936	1.103	8.708 (HU: Dél-Dunántúl)	13.137 (GER: Nordrein-Westfalen)	107	MLN 2000 EUR (log)	2005
Physical capital (s^k)	0.214	0.073	0.040 (GER: Bremen)	0.793 (GER: Hamburg)	107	share of GDP	average of 2006-2014
Human capital (s^h)	3.829	1.603	1.553 (DE: Rheinland-Pfalz)	12.771 (GR: Nisia Aigaiou/Kriti)	107	% of population	average of 2006-2012
Population growth (n)	2.918	5.200	-9.116 (DE: Sachsen-Anhalt)	16.397 (NOR: Oslo og Akershus)	107	% change	2001-2014
Share of 18-34 (YNG)	6.199	1.470	2.960 (FR: Mediterrannee)	19.104 (GR: Nisia Aigaiou/Kriti)	107	% of population	average of 2004-2014
Specialization	42.728	5.811	33 (HU: Dél-Dunántúl)	62.967 (GR: Nisia Aigaiou/Kriti)	107	Share	2012
Population density	401.182	1012.129	3.317 (SE: Övre Norrland)	6808.738 (UK: Greater London)	107	Share	2012
$NEIGH_{sk}$	0.209	0.028	0.099 (UK: Northern Ireland)	0.325 (GER: Schleswig-Holstein)	107	%	average of 2006-2014
$NEIGH_{sh}$	3.637	1.072	2.149 (GER: Berlin)	8.805 (GR: Kentriki Ellada)	107	%	average of 2006-2014
$NEIGH_n$	2.443	3.630	-5.333 (GER: Sachsen-Anhalt)	9.466 (UK: Northern Ireland)	107	%	average of 2006-2014
TEA	5.458	1.584	1.871 (FR: Center-East)	9.699 (DE: Hamburg)	107	% of population	2001-2006
TEA_{in}	1.243	0.542	0 (PT: Norte, Algarve, Lisboa, Alentejo)	2.954 (DE: Hamburg)	107	% of population	2001-2006
TEA_{hj}	0.689	0.454	0 (PT: Norte, Algarve, Alentejo)	2.334 (GER: Hamburg)	107	% of population	2001-2006

Eurostat data are taken from Eurostat (2016). GEM data from GEM (2016)

adult population and measures Total Entrepreneurial Activity (TEA) as the share of people who select themselves either as nascent entrepreneurs (actively involved in setting up a business), or as an owner-manager of a new firm existing up to 42 months. For an elaborate account of the methodology and operational details we refer to Reynolds et al. (2005); Bosma (2013). GEM also considers various subcategories of entrepreneurship, of which two are of particular interest for our purpose. Under innovative entrepreneurship (TEA_{in}), only the entrepreneurs who indicated that their product or service is new to customers and reported that not many competitors are offering the same product or service are included. For entrepreneurs with a high job-growth-orientation (TEA_{hj}), early stage entrepreneurs who expect to have 20 or more jobs in their company in the next 5 years are included. To obtain measures at the regional level with acceptable statistical precision, we follow Bosma (2009) who computed regional indicators over six annual GEM waves for our European regions. The correlations between these variables are displayed in table 6.5 in the appendix. We see that there are quite some significant correlations between the variables we are interested in. However, none of these correlations are extremely high so multicollinearity is not a problem when running our regressions.

Following the literature on regional growth (Dogaru et al., 2011; Thissen et al., 2011), we added controls on population density, regional diversification over industries ($SPEC$)[8] and the share of young people between 18 and 34 in the population (YNG). Finally, we added a vector of variables, $NEIGH$, to control for spillover effects. We identified the neighbouring regions for each region and took the average saving rate of those regions ($NEIGH_{s_k}$), the average human capital share ($NEIGH_{s_h}$) and the average population growth ($NEIGH_n$). The neighbouring region's are the same for any region for each of the three variables, so one might expect a high correlation between these three variables but this problem seems to be limited due to variation in the variables of interest. The highest correlation is between the physical capital of a region's neighbours and the population growth of neighbours, and this is -0.44. We are not interested in the individual signs of these neighbour effects, so we just report whether they are jointly significant.

We observe that the variation in GDP growth rates is large and that all

[8]The degree of regional specialization is measured by share in employment of the 5 largest industries out of total employment. It actually measures specialization, so a high score means high specialization and thus low diversity.

variables fall in the expected range (the mean growth rate in the period 2006-2014 is 1.28 per cent, with a minimum of -2.61 per cent for Central Greece and a maximum of 6.46 for West Norway. Growth rates in this period are severely affected by the financial and Eurocrisis, naturally, so these rates are much lower than in preceding years.

No severe outliers were detected. In table 6.5 in the appendix, we see that of the usual suspects, only population growth has a significant and positive correlation with GDP growth, suggesting that at the regional level, the Mankiw-Romer-Weil model is less potent. We also have to note, however, that 9 years is a very short time for the type of long-run growth model that they describe. All three entrepreneurial variables are positively correlated with GDP growth, but only entrepreneurship with a focus on high growth is significant. Naturally, the three types of entrepreneurship are also highly correlated amongst each other. However, we cannot interpret these correlations as causal relations, as interdependencies between regions in the same country are not accounted for yet and causality in these variables is very likely to run both ways. This will bias the estimated coefficients later on, but should not affect our results on between country or region differences in the correlation coefficients of entrepreneurship with growth.

4.4.2 Linear regression and multilevel results

We first run a standard OLS regression with robust standard errors of GDP growth with the variables of interest as regressors, following the specification from the MRW model. The results are displayed in table 4.2. The standard OLS regression (model 1) does not find evidence for a convergence effect of regional GDP between 2006 and 2014, the coefficient is insignificant. This indicates that 9 year growth rates are simply too short to detect significant convergence among regions and the fact that these years are mostly crisis years also obscures the standard convergence dynamics. As our correlation table (table 6.5 in the Appendix on page 149) already suggested, the rate of population growth has a positive effect on the average GDP growth rate. But again, we cannot claim any causality from this finding. It could for example be that fast growing regions also attract more migrants, explaining this positive coefficient. Low migration overall in Europe, however, does suggest the causality runs mostly from population growth to GDP growth. We observe that s^k and s^h do not have significant effects on growth rates. This might be explained by the fact that human capital is high in all European regions or that the mobility of human and physical capital is high

such that the explanatory power on growth is not large. Also, the rates of capital and human capital accumulation and specifically physical capital utilization rates have been seriously affected by the crisis.

We then ran the OLS regression including the controls suggested in the regional growth literature (model 2). The regional controls include an index of how specialized a region is across sectors ($SPEC$), the share of the working age population that is between 18 and 34, to take into account the demographic structure of the region (YNG) and population density ($POPDENS$), which is a measure of urbanization and market size. Furthermore, to capture spillover effects we control for the mean levels of the three production inputs of the Mankiw-Romer-Weil model (physical and human capital and population growth) of a region's neighbouring regions ($NEIGH$).[9] The rationale behind this is that factor mobility between regions may be so high that the growth of these factors in neighbouring regions may actually produce growth in other regions.

[9] Both within the same country and across national borders, where applicable.

Table 4.2: OLS regression of MRW model with regional controls and entrepreneurship

$\Delta lnY(2006-2014)$	1	2	3	4	5	6
Initial GDP $(lnY(2005))$	-0.056	-0.135	-0.219*	-0.143	-0.158	-0.138
	(0.134)	(0.134)	(0.125)	(0.121)	(0.135)	(0.135)
Physical capital (s^k)	0.337	-1.261	-1.573	-1.291	-1.357	-1.316
	(1.968)	(1.408)	(1.245)	(1.424)	(1.273)	(1.361)
Human capital (s^h)	-0.096	-0.077	-0.042	-0.076	-0.059	-0.080
	(0.087)	(0.117)	(0.115)	(0.118)	(0.125)	(0.117)
Population growth (n)	0.077***	0.137***	0.133***	0.139***	0.129***	0.131***
	(0.027)	(0.047)	(0.044)	(0.046)	(0.044)	(0.046)
Share of 18-34 (YNG)		-4.164	-5.106	-4.432	-3.513	-3.545
		(9.329)	(8.904)	(9.274)	(8.686)	(9.338)
Specialization		-0.118***	-0.094***	-0.117***	-0.103***	-0.117***
		(0.019)	(0.021)	(0.020)	(0.019)	(0.019)
Population density		-0.000	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Neighbour$^+$		***	***	***	***	***
TEA			-13.318	-2.117		
			(8.0078)	(8.420)		
TEA_{hj}			95.453**		64.604*	
			(37.601)		(34.253)	
TEA_{in}			-5.809			16.165
			(31.589)			(24.092)
$Constant$	2.107	9.606***	8.808***	9.871***	7.111***	9.166***
	(1.732)	(3.073)	(2.556)	(2.682)	(1.751)	(2.890)
Model Fit Statistics						
R-squared	0.075	0.417	0.462	0.417	0.447	0.420
Observations	107	107	107	107	107	107
AIC	377.296	339.947	337.225	341.858	**336.374**	341.373

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01
$^+$ Vector of spillover variables: neighbours' average s^k, s^h and n, *** jointly significant at the 1% level

Out of the regional control variables, $SPEC$ is always negative and highly significant, indicating that more diversified regions have grown faster in these years, which is an intuitive result during crisis years. Furthermore, the effects of neighbour physical, human capital and population growth are always jointly significant. This effect is mainly driven by population growth in neighbouring regions.

The fact that the share of young people has no effect on economic growth is in line with the findings of Rodríguez-Pose and Crescenzi (2008), who also showed that e.g. life-long learning, the percentage of the labour-force

working in agriculture, and the demographic structure of the population are not significant in explaining growth across European regions during roughly the same period. Spillover effects are always jointly significant, and this is mainly driven by the effect of $NEIGH_n$, the average population growth of neighbouring regions.

We then included our proxies for entrepreneurship in columns 3-6. We deviate here from the standard MRW specification, as the other variables are averages over the sample period, whereas our entrepreneurship measures have been collected between 2001 and 2006, so in the preceding period. We do this because Fritsch and Mueller (2008) have found that the effect of entrepreneurship on growth rates often takes up to ten years to materialize. Furthermore we want to show that our results are not caused by a simultaneity bias.

Adding the three entrepreneurship measures to the MRW model alters the estimated coefficients of other variables only marginally. Only entrepreneurship with a focus on high growth (TEA_{hj}) is significant when introduced to the model both individually and jointly with the other two measures. To put things in perspective, note that adding the different entrepreneurship variables to the standard MRW specification explains a mere additional 3-5% of the unexplained variance. Investment in physical and human capital remain statistically insignificant when entrepreneurship variables are added.

The results obtained from the OLS estimations above, are, however, not to be taken at face value. Contrary to the cross-country analysis of MRW, we are dealing with regions in countries and this creates two challenges; first of all, cross country variation in the production function will now be absorbed in the constant and residuals. Furthermore, when regions in country A all suffer from lower productivity relative to regions in country B, the error terms are no longer independently and identically distributed. To incorporate country effects, we therefore estimate our model in a multilevel regression, using restricted maximum likelihood estimation. In this multilevel specification we allow the intercept, and thus the mean growth rate, to vary between countries. We tested the standard linear model against the varying-intercept model for all specifications and the Likelihood Ratio test statistics show that the country multilevel model should be preferred, both including and excluding entrepreneurship proxies. For several specifications we also tested whether a varying-slope model should be used for entrepreneurship. These models were rejected in favor of the varying-intercept model.

Table 4.3: Multilevel models

$\Delta lnY(2006-2014)$	0	1	2	3	4
Initial GDP $(lnY(2005))$		0.185***	0.106	0.070	0.083
		(0.071)	(0.075)	(0.076)	(0.075)
Physical Capital (s^k)		-0.249	-0.367	-0.952	-0.735
		(0.690)	(0.672)	(0.736)	(0.702)
Human Capital (s^h)		0.024	-0.082	-0.083	-0.093*
		(0.037)	(0.054)	(0.055)	(0.054)
Population Growth (n)		0.077***	0.097***	0.091***	0.093***
		(0.014)	(0.016)	(0.016)	(0.016)
Share of 18-34 (YNG)			6.480	7.493	7.794*
			(4.609)	(4.574)	(4.530)
Specialization			-0.042***	-0.043***	-0.044***
			(0.013)	(0.013)	(0.013)
Population density			-1.88x10^{-5}	-4.1x10^{-5}	-1.95x10^{-5}
			(0.000)	(0.000)	(0.000)
Neighbour^{+}				*	*
TEA				-5.02	
				(5.415)	
TEA_{hj}				29.577	30.307*
				(20.003)	(17.917)
TEA_{in}				21.830	
				(19.137)	
Constant	1.279***	-1.035	1.486	1.582	1.599
	(0.384)	(0.919)	(1.372)	(1.379)	(1.361)
SD (Country level)	1.514	1.303	1.301	1.316	1.296
SD (Regional level)	0.567	0.421	0.417	0.417	0.417
Model Fit Statistics					
Observations	107	107	107	107	107
ICC	0.728	0.774	0.756	0.759	0.757
AIC	246.563	**217.534**	219.420	229.619	258.682

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

$^{+}$ Vector of spillover variables: neighbours' average s^k, s^h and n, * jointly significant at the 10% level

The results from the multilevel growth regression are presented in table 4.3. Again we first present the MRW model in a multilevel specification (column 2), the convergence effect actually becomes positively significant, meaning that once country-level variation is accounted for in the multilevel model, regions within that were richer within the country in 2005 are also the regions that achieved the highest growth rates.[10] This effect does not remain significant, however, when regional controls are introduced. Population growth has again a positive significant effect, and investment in physical and human capital still do not contribute to GDP growth in the period studied. Recall these are crisis years in which under utilization of capital is not uncommon.

The regional controls behave the same as in our OLS model; the degree of specialization has a significant negative effect, population density and the demographic structure do not influence regional growth rates. By controlling for unobserved heterogeneity between countries, the spillover effects become much weaker, although they are still significant at the 10% level. Adding all three entrepreneurship measures jointly does not lead to a significant effect of any of the measures, but when high job-growth entrepreneurship is entered individually it is still significant at the 10% level.

We use the Akaike Information Criterion (AIC), a measure based on the log-likelihood of the model, to see which model is preferred. The standard MRW-model without regional controls is the model with the lowest Akaike Information Criterion (AIC). This means that, within this range of nested models, it is the model with the best fit. It also means that neither regional controls nor entrepreneurship measures can significantly improve upon the standard growth model explaining patterns of regional growth in our sample. Following this, we use the residuals from the MRW model to test for latent classes, although as a robustness check we also used the residuals from model 2. Our multilevel specification also allows us to see whether the added variables explain relatively more level 1 (country level) or level 2 (regional level) variation. The Intra Class Correlation (ICC) can be interpreted as the degree of similarity between two regions that are in the same country (Snijders and Bosker, 2012). Or alternatively as the share of the variance that can be attributed to differences at the country level. SD(country level) reports the estimated variance that is attributable to the

[10]This result does not necessarily contradict previous findings from the literature. The empirical growth literature has not yet resolved the issue whether an initially higher income is reinforced over time or whether convergence takes place, as predicted by theoretical models. Furthermore, most of the theoretical convergence literature focuses on lagging countries, not regions within countries.

country level, while SD(Regional Level) reports the residual variance at the regional level. When the former is significantly different from zero, as it is for all models including regional controls, this implies country level differences are significant and a multilevel estimation is appropriate.

From the ICC we see that adding entrepreneurship variable TEA_{hj} to Model 2, we see that this slightly decreases the ICC, although the difference is very small. This means that entrepreneurship thus picks up some variance at the between-country level, while the regional variance within countries remains equal. Adding regional control variables to Model 2 decreases the ICC. This implies that these controls do better in explaining between-country regional differences than within-country differences, which is not necessarily what we would expect. Again it may be explained by high mobility within countries.

We can now investigate whether the effect of entrepreneurship varies between regions inside countries. We do this only for high job-growth-oriented entrepreneurship, as this was the variable with a positive significant effect in the OLS estimation. The intuition is that even though the homogeneous effect of high job-growth-oriented entrepreneurship is not significant, it may be significant if we allow for more heterogeneity, for example distinguishing between urban and rural regions. A first way to measure the possible interaction effect is to add the country-level mean of high job-growth entrepreneurship to the model, as well as the interaction effect between a country's average entrepreneurship rates and the region's deviation from that country mean.

Table 4.4 below shows how the multilevel model is built up. Column 1 presents the Null model, including just the intercept. This model shows (again) that the average GDP growth rate over this period is 1.28%. In column 2, we add the mean rate of high job-growth entrepreneurship per country to the model, in order to explain some of this country-level variation. The constant, which now corresponds to the GDP growth rate in a country with average entrepreneurship levels, is 0.788%. The country-level of entrepreneurship with a focus on high job growth is not significant. Adding the mean rate of entrepreneurship to the multilevel model does increase the ICC, the share of the variance that is explained by the country level, and this makes sense because we added a country-level explanatory variable. In the third column we add the regional high-job growth focused entrepreneurship rate as deviation from the country mean and an interaction effect with that country-mean. Both are not significant. In column 4 we added all other variables. The control variables that were significant in previous estimations, population growth, the share of young people

and the specialization rate, are still significant and have the same sign, but the country-level entrepreneurship does not help explain regional growth in this model. The model with the best model fit is the model with only the country-mean, deviation from the mean and interaction effect of high-job growth focused entrepreneurship. We conclude from this that when country-level heterogeneity is controlled for, there is no significant homogeneous effect of entrepreneurship on regional growth, and we cannot find evidence for heterogeneous effects of entrepreneurship on growth between regions in different countries.

We obtained the same results when we add the group mean of total entrepreneurial activity or high job growth entrepreneurship in this multilevel regression; the effect of the entrepreneurship is non-significant also when we take the mean, the deviation from the mean or the interaction effect of those. Therefore we may conclude that we do not find a robust positive effect of adding entrepreneurship rates to a regional growth regression. There is a positive correlation between regional growth and the number of entrepreneurs that expect high job growth in the future. This variable also has a significant effect in a simple OLS regression, but the effect becomes weaker when we control for unobserved heterogeneity between countries through a multilevel analysis. We cannot yet reject the hypothesis of a heterogeneous relationship between entrepreneurship and economic growth, however. Our multilevel specification does not pick up systematic differences between regions that cluster on other groupings than countries. One might, for example, think of urban versus rural regions, industrial versus agricultural or larger regional clusters such as Scandinavian, North-Western, Southern and Eastern EU-regions. To test for the existence of such clusters we best employ a latent class model.

4.4.3 Latent class analysis

From the standard MRW model, excluding entrepreneurship, we can take the residuals and save these as a measure of 'unexplained growth'. We know from our analysis above that entrepreneurship has some explanatory power, but so far we were unable to establish a strong link at the regional level. This should make us pessimistic about finding strong variation in estimated coefficients on entrepreneurship across regions. Still, the objective in study is to illustrate a method that would uncover differences in the contribution of entrepreneurship to residual growth across latent classes of regions. A straightforward way of testing this would be to interact the entrepreneurship variable with our moderator of interest, the entrepreneurial

Table 4.4: Multilevel model with fixed and random effects

$\Delta lnY(2006-2014)$	1	2	3	4
Constant	1.279***	0.788	0.786	1.945
	(0.384)	(0.824)	(0.824)	(1.510)
TEA_{jg} (mean)		69.359	69.541	18.330
		(102.387)	(102.232)	(89.954)
TEA_{jg} (deviation)			25.151	-47.638
			(86.380)	(63.719)
TEA_{jg} (mean*deviation)			6575.155	11300.55
			(10651.03)	(7721.956)
Initial GDP $(lnY(2005))$				0.066
				(0.071)
Physical capital (s^k)				-0.585
				(0.735)
Human capital (s^h)				-0.091
				(0.053)
Population growth (n)				0.093***
				(0.016)
Share of 18-34 (YNG)				6.794*
				(4.407)
Specialization				-0.040***
				(0.012)
Population density				-5.08x10^{-5}
				(0.000)
Neighbour^{+}				*
S.D. (Country level)	1.514	1.540	1.548	1.353
S.D. (Regional level)	0.527	0.567	0.626	0.378
Model Fit Statistics				
ICC	0.728	0.731	0.712	0.781
AIC	246.564	237.018	**191.448**	200.882

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

$^{+}$ Vector of spillover variables: neighbours' average s^k, s^h and n,

* jointly significant at the 10% level

ecosystem. However, as already mentioned this concept cannot (yet) be empirically captured in a meaningful way. Our method allows for this heterogeneity to be found in the data without imposing any regimes or interactions beforehand. We do this in a two-step procedure, using the residuals from standard regressions instead of estimating our whole model in a Latent Class Analysis because the number of parameters that the LCA needs to estimate increases quickly with the number of classes that we want to test. Introducing the whole set of controls in the LCA would leave very few degrees of freedom as we only have 107 observations. We use residuals from the OLS model, where we do not control for country effects, and from the multilevel model, where we do. The regressions above showed that in the OLS regressions, the model with the best fit was the model with regional controls. In the multilevel model it was the model with only the MRW variables, so we use the residuals from these models.[11]

For the OLS residuals, the correlation between unexplained growth and high job-growth oriented entrepreneurship is significant and 18.8%. If we take the residuals from the MRW model the correlation is even 32.2%. For the multilevel residuals the correlation with high job-growth oriented entrepreneurship are respectively 5.9% and 8.3%, both not significant at the 10% alpha level. Graphs 4.4.1 and 4.4.2 show the bivariate relation between residual GDP growth and job-growth oriented entrepreneurship, respectively and this shows that there is a positive relation between entrepreneurship and cleaned growth, but it mostly disappears once unobserved country-level heterogeneity is controlled for. A visual inspection does not directly hint at the existence of multiple 'clouds' of data points that could indicate varying relations between entrepreneurship and regional growth, but the Latent Class Analysis in the next section provides a formal test of this. In order to test this hypothesis we regressed the two measures for unexplained growth on our three measures of entrepreneurship in six latent class specifications and tested for the existence of multiple regimes in our sample. We compared models with 1, 2, and 3 classes on the base of three information criteria. The model that scores the lowest values on these criteria is to be preferred (see table 4.5 and 4.6).

For both the OLS model residuals and the multilevel residuals the model selection criteria show that for all three explanatory variables, the model with one class is preferred. We do see that in several of the specifications,

[11] As a robustness test, we also used the residuals from the MRW-only model for OLS and for the model with all the controls in the multilevel specification, but this does not alter the findings.

Table 4.5: Selecting regimes: Multilevel residuals

TEA	Log-likelihood	Obs	Parameters	AIC	Significance of entrepreneurship
1 class	-52.719	107	2	**-1.815**	No
2 classes	-49.475	107	7	1.056	No, No
3 classes	-44.181	107	11	1.031	No, No, No
TEA_{in}					
1 class	-52.510	107	2	**-1.819**	No
2 classes	-49.891	107	7	1.063	No, No
3 classes	-41.965	107	11	1.016	No, No, No
TEA_{hj}					
1 class	-52.506	107	2	**-1.819**	No
2 classes	-49.510	107	7	1.052	No, -
3 classes	-45.286	107	11	1.052	No, +, No

+ = positive at the 5% level, - = negative at the 5% level, No = not significant.

Table 4.6: Selecting regimes: OLS residuals

TEA	Log-likelihood	Obs	Parameters	AIC	Significance of entrepreneurship
1 class	-158.936	107	2	**0.170**	No
2 classes	-146.109	107	7	2.862	-, +
3 classes	-142.776	107	11	2.874	-, +, No
TEA_{in}					
1 class	-158.743	107	2	**0.166**	No
2 classes	-148.831	107	7	2.913	+, No
3 classes	-142.095	107	11	2.862	+, -, No
TEA_{hj}					
1 class	-157.096	107	2	**0.136**	+
2 classes	-151.868	107	7	2.969	No, No
3 classes	-147.137	107	11	2.956	+, No, No

+ = positive at the 5% level, - = negative at the 5% level, No = not significant.

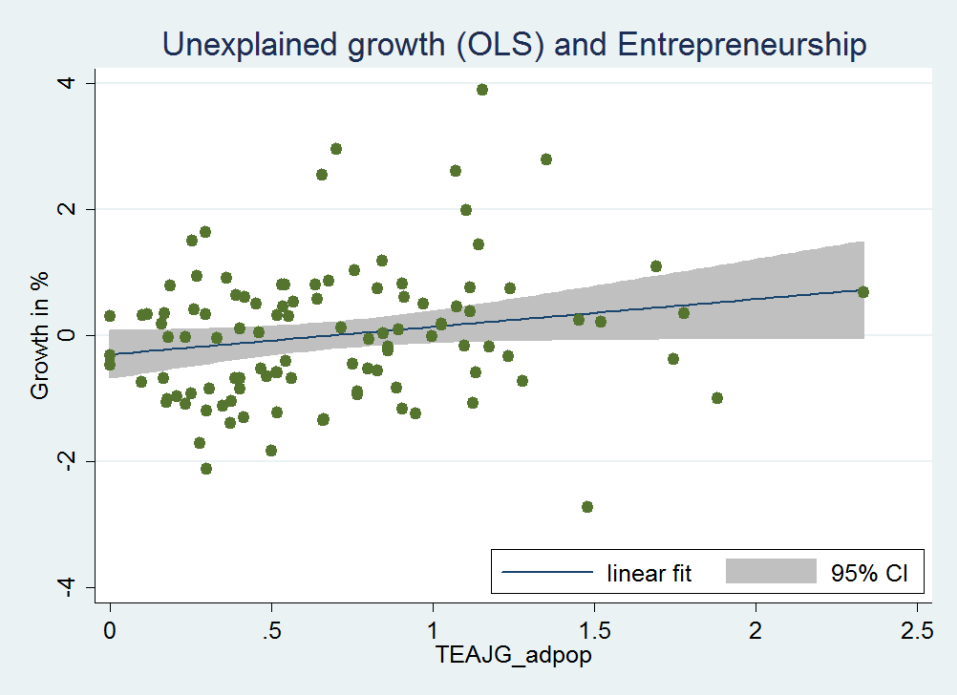


Figure 4.4.1: Relation between total early stage job-growth-oriented entrepreneurial activity and economic growth

there can be found groups of observations with a positive and a negative effect, and considerably large prior probability for observations to fall into those classes. But the parsimonious model with only one class still fits the data better. Therefore, in this data set we cannot reject the hypothesis that the marginal effect of entrepreneurship on growth is homogenous across all regions in our sample, controlling for country fixed effects.[12]

[12]In the Appendix we show in table 6.4 the same analysis when using the results of our OLS analysis with only the MRW controls, but the conclusion remains that one class is preferred.

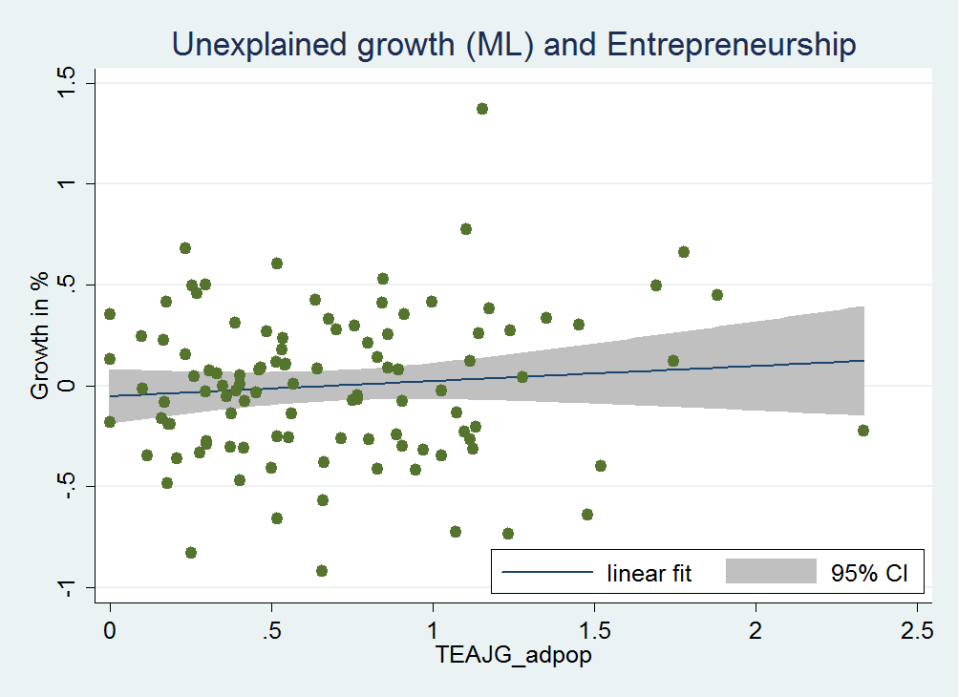


Figure 4.4.2: Relation between total early stage job-growth-oriented entrepreneurial activity and economic growth

4.5 Discussion

In this chapter we argue that an entrepreneurial ecosystem must reveal itself in the data through affecting the impact of observed entrepreneurial activity on GDP-growth. Then, if ecosystems differ in quality across space, we should be able to reveal the existence and relevance of entrepreneurial ecosystems in the heterogeneity of marginal impacts of entrepreneurial activity on GDP growth at the appropriate geographical level. Finding evidence of such heterogeneity would prove that something is affecting the contribution of entrepreneurship to economic growth differently in different places. We then collected data and designed an empirical strategy to reveal such heterogeneity, if present. We find no support for our hypothesis at the NUTS1-2 aggregation level in our data. But not finding such heterogeneity in this data set should certainly not be interpreted as proof that the entrepreneurial ecosystem is non-existent. Instead, we can only conclude it does not reveal itself at this level of aggregation.

We simply cannot reject the hypothesis of a homogeneous effect of entrepreneurship in our data. This, however, is not the same as rejecting the hypothesis of a heterogeneous effect. There may be a variety of reasons why our result was found. First, and we deem this most likely, our level of geographical disaggregation may simply be too high for the effects of local ecosystem differences to show up statistically significant. Second, the sample under study might not include enough countries to find a heterogeneous effect. The data is very noisy and from the literature we already knew that the impact of entrepreneurship itself is hard to properly identify. This then holds *a fortiori* for any moderation effects on entrepreneurship.

The regions we studied may simply not be different enough to find significantly different marginal effects. Third, the period under study is a short period when compared to other empirical growth studies, and the growth rates of European countries in this period have been severely affected by to the impacts of the Global financial crisis and the Eurocrisis. In an earlier version of this study we also ran the analysis on the average growth rate over a much longer period. There we face the problem that we regress long run growth on a relatively short sample of entrepreneurial variables observed towards the end of the period. But changing this does not qualitatively change our results. We only find a significant effect of entrepreneurship on growth in our OLS model, where unobserved heterogeneity between countries is not taken into account. The probable endogeneity between entrepreneurship and economic growth may also be why, despite the time lag, it is hard to disentangle the empirical effect of entrepreneurial activity on GDP. And fourth, the countries in the sample are all developed and very much integrated countries, which might imply entrepreneurship and other factors are highly mobile within and even across countries. We tried to control for that by taking into account spillover effects of factor accumulation in neighbouring regions, but such mobility would compound the endogeneity problems and reduce the variation in the marginal productivity of entrepreneurial activity across regions. An interesting finding is that the only type of entrepreneurship that has an effect on growth in this study is from the group of entrepreneurs that also expect high job growth in the next years, indicating that the 'gazelle hypothesis' (Henrekson and Johansson, 2010) also holds for European regions and that entrepreneurs can actually identify themselves as gazelles.

We carried out several robustness checks on our estimations. We tried alternative specifications to control for the country effect, respectively a varying-coefficient multilevel model and a standard OLS regression with clustered standard errors by country or country dummies. Both models

were rejected in favor of the varying-intercept model presented above. Following Paterson and Goldstein (1992), Gorard (2003) and Leckie et al. (2010) we also recommend in general the use of multilevel-models in studies on entrepreneurship and growth, as a multilayered web of influences at different levels of aggregation is at play.

Our results therefore only mean that the regional ecosystems considered here are either not that different or matter little for the contribution of entrepreneurship to regional growth in the period under study. Finally, during the sample period, most European regions have suffered from the global financial crisis, affecting growth rates more than entrepreneurship. This too may confound the long run effects of ecosystem quality differences. But even as our data do not justify strong conclusions at this stage, we do feel strongly about the appropriateness of our method and consider this the main contribution of our study. We feel that ecosystem quality proxies cannot (yet) capture the full richness of the entrepreneurial ecosystem, whereas more complex composite indices impose an assumed structure and are not suited as independent variables. Therefore more studies should make use of latent variables and indirect relations to show the possible moderating effects of ecosystems on innovation and growth before pushing proxies and measures that are supposed to explain that variation directly. Moreover, such analyses may inform data collection efforts in the future. If anything, our analysis has shown that collecting entrepreneurial ecosystem data at the NUTS1-2 level may not be the right level of aggregation, and research should focus on national or more local levels of aggregation instead. Establishing where the relevant ecosystem boundaries (do not) lie is an important contribution of our study also, as it is the first step in collecting better data efficiently in the future.

The challenge remains to make the entrepreneurial ecosystem visible in the data. With better and more complete data our method may allow us and others to do so in the future. In the light of this thesis, we have to admit that the differential structure of institutions as would theoretically follow from Chapters 2 and 3 cannot be revealed here. However, as mentioned above, this may be due to the level of aggregation that we are looking at due to data restrictions. The method as used here still provides a valid tool for disentangling latent entrepreneurial ecosystems in different samples and aggregation levels where data is available.

5

Conclusion

5.1 Motives and discussion of main results

Throughout this thesis I examined the relationship between institutions, demography and economic development. It is well established that institutions are relevant for development, but the questions regarding how institutions emerged and diffused or what causes differences in institutional strength remain largely unanswered. Institutions and economic development are two concepts that are interdependent and coevolve together. Therefore, to find an empirical answer to the questions above, the truly exogenous determinants of institutional emergence and development need to be studied. These exogenous components are proposed to be found in natural geography and genes (see e.g. Hibbs and Olsson 2004 and Ashraf and Galor 2013). Biogeographical initial conditions have led to different demographic processes across the world, which in turn put societies on different trajectories of institutional development. Thereby these initial conditions are fundamental in explaining inequality across the globe today. This should not be equated with geographic determinism; I argue that institutional development does not automatically follow even if the necessary conditions are met, as this may require the presence of diffusion agents or strong reciprocators. Furthermore, although societies develop their initial set of institutions largely in isolation, different forms of global contact can lead to the adoption of new institutions.

Specifically, in this dissertation I addressed three sets of questions in three chapters. The first chapter was concerned with the evolution of self-organizing, trust-based communities to societies with more complex forms of organization. In a theoretical model I assumed that the increase in pop-

ulation density following the Neolithic Revolution led to the emergence of a latent demand for institutions. If these institutions are implemented in a successful way this creates positive feedback loops between investment in institutions, population growth and economic development. However, the broad collection of empirical and theoretical evidence reviewed in chapter 2 shows that there is no linear path of institutional development. The successful setup and reinforcement of institutions hinges on whether a society is able to overcome problems of collective action and whether some natural boundaries keep others from freeriding on this society's institutions.

Chapter 3 addressed what happens once societies —having followed their own specific trajectories of institutional development —come into contact with each other. In particular, how do different types of contact between societies (war, trade, colonial ties, diplomatic exchange) lead to the diffusion of a specific institution. To illustrate how such diffusion processes might be studied empirically I present a study on the diffusion of the extension of franchise to include all women. This institution was chosen not because of its particular relevance to the process of economic growth as compared to other institutions, but because it provides a suitable case to study of the diffusion of a formal institution. This analysis could and should be reproduced using other institutions to confirm whether the same mechanisms prevail. I found that contact between countries is an important predictor of the timing that a country adopts this institution, even when within-country predictors and external shocks are controlled for. This suggests that, when studying institutional change and adoption, diffusion is an important channel and mechanisms of diffusion through contact should be more intensely studied. I also found that being engaged in a war, regardless of whether this is with prior adopters or not and regardless of the outcome, strongly predicts the adoption of full voting rights for women. This suggests that institutional change is more likely in turbulent and unstable situations. Other institutional changes have also been found to coincide with major conflicts, such as the waves of democratization and the constraints on power imposed by victors after war (Huntington, 1991; Gleditsch and Ward, 2006; Ikenberry, 2009).

In the fourth chapter, I asked how institutional frameworks or ecosystems may affect the relation between a region's production factors and economic growth. More specifically, is it possible to disentangle a moderating relationship of the quality of the institutional ecosystem on the contribution of entrepreneurial activity to economic growth? Again, entrepreneurship is chosen here for pragmatic reasons; to illustrate a method for investigating latent aspects of institutional complexes. I studied this in a sample of Euro-

pean regions at the NUTS-2 level for the years 2001-2014. For this specific level of aggregation I found no evidence for a moderating effect of latent institutions, in the form of an *'entrepreneurial ecosystem"* or a *'system of innovation'* (Lundvall, 1992; Nelson, 1993; Hwang and Horowitt, 2012). Not being able to confirm the hypothesis in this dataset does not imply that such ecosystems do not exist, however. It is clear that empirically capturing the effect of institutional frameworks on growth remains a challenge. As Farole et al. (2011) state, there is a macro-geography of nested core-periphery patterns, such that agglomerations and comparative advantages linked to institutional differences are at work over various, overlapping scales. If some regions do have more potential to create economic growth from their productive factors than others, due to these differences in institutional strength, we need to establish the level of geographical aggregation at which this heterogeneity can be found before we can try to measure what constitutes these differences. So far, I found a homogeneous effect of entrepreneurship on economic growth, although it has to be noted that the effect of entrepreneurship in this time period is very weak to begin with, and difficulties of measuring entrepreneurship[1] add to the challenge.

5.2 Limitations and opportunities

I aimed to disentangle the relations between demography, geography, institutions and growth, respectively. This is an ambitious goal, as not only are these concepts hard to measure empirically, they also coevolve endogenously and at different, overlapping scales. This makes it extremely complex to disentangle (causal) relations between them. Using a mix of theoretical, empirical and historical anecdotal evidence I hope to have added some understanding to these issues. However, it is clear that every research choice comes with limitations and that many opportunities for future research exist, to build and improve on my work.

In Chapter 2, I presented a theoretical model illustrating how demographic pressure leads to a growing latent demand for institutions. Under evolutionary pressure, societies that are able to successfully set up and invest in these institutions will overtake societies that do not. I did not incorporate into the model what determines which societies will be able to achieve this and which will not. Luck, biogeography, genetic variation can

[1] Especially regarding the type of entrepreneurship that is believed to lead to innovation and economic growth.

all cause variation that in the end drives these trajectories. Investment in institutions or public goods may also depend on the extent to which a society is able to exclude non-members. Mayshar et al. (2011) show that certain modes of food-production are more transparent and therefore taxes necessary to sustain investment in institutions are more easily collected. These factors are also related to the physical geography in which a society resides. Future models can try to incorporate some heterogeneities between societies in order to see when and where cooperation will be sustained and how strong evolutionary forces are required to be to for evolution to select on this trait. Another extension of this theoretical model is to endogenize technical change, which for convenience was assumed to be growing with periodic shocks in the background now. It can be argued that there exists an additional feedback mechanism between technological progress and institutional development.

In Chapter 3, I study a particular institution, the extension of voting rights to include all women, to gain insights into how contact between adopters and non-adopters can predict the spread of an institution. I show that contact should not be overlooked and that differences exist between predictors of adoption for early and late adopters. To generalize over multiple institutions and say something on the diffusion of growth-enhancing institutions in general, and on how contact between societies plays a role in this, more of these types of studies should be done. Institutions that differ in terms of the timing of adoption, embeddedness, level of controversy and legitimacy and the extent to which they are formal versus informal should be compared to see if the same mechanisms of diffusion operate. This chapter has put forward a way in which such studies can be conducted in a comparable way.

In Chapter 4, I looked for empirical evidence of core-periphery (or any other systematic) patterns of institutions. Institutions do not lead directly to higher production, but they can lower transaction costs and stimulate investment or innovation. Therefore, the effect of higher quality institutional frameworks can manifest itself through different indirect channels. In this thesis we chose to focus on the moderating effect of latent institutional quality on the relationship between entrepreneurial activity and economic growth, because these data are available at a relatively low level of aggregation and because it presents a suitable case, as argued by Baumol (1990). Repetition of our work with other data for which institutions are expected to moderate the effect of the independent variable on economic growth or development would be useful and desirable. Although no evidence for these types of institutional regimes was found, the hypothesis of

a core-periphery pattern of institutional strength that influences the extent to which a region can utilize its growth potential should not be dismissed. In fact, future research should intensify the search for these institutional regimes at different levels of aggregation to determine their geographical scales and boundaries. Latent Class estimation is a suitable tool to search for the relevant ecosystem boundaries in cross-country regional data sets. That way, the hypothesis that rural regions across countries have more in common than different types of regions within one country. Following that exercise, the factors that comprise those institutional regimes that produce high rates of economic growth could be determined.

This dissertation has tried to shed light on the complex and dynamic relation between people, geography and institutions and the way in which these three factors interact to contribute to sustained economic growth. It has to be acknowledged, however, that we are still far from understanding the full complexity of this relationship and even further from empirically disentangling it. Demographic change, institutional development and economic growth are very much intertwined and follow endogenous processes, by which it is hard to identify what causes what. Geography can provide the exogenous variation that gives us a starting point for finding causal relationships, but the importance of geography fades over time as these intertwining processes take place. In this dissertation I did not use instrumental variables to capture exogenous variation, because good exogenous proxies as an instrument for institutions are very hard to come by. Many instruments suffer from the problem that they explain little variation in the endogenous variable they are trying to capture (in this case the institutional framework), while a weak relationship between the instrument and the error term already leads to large inconsistencies in estimation. Another alternative would be to analyze a natural experiment, i.e. a case-study with an exogenous shock in institutional quality. Again, these experiments are extraordinary in history as most institutional changes are endogenous. The rare opportunity of a natural experiment showing the effect of institutional differences on economic growth, everything else being equal, is studied e.g. in the case of North and South Korea by Acemoglu et al. (2005) and provides detailed insights. In this dissertation I want to focus on general insights and mechanisms. An institutional framework is a rich, multidimensional complex concept that in our opinion can be best captured latently. An obvious critique is then that this dissertation cannot go further than describing correlations, rather than causal relations. While one study alone may not provide strong evidence for identifying these links, the cumulative nature of science can show the likeliness of certain

hypotheses. In my opinion, this dissertation should be considered together with other studies, theoretical studies, case studies and empirical studies using instrumental variables, several of which were reviewed in Chapter 2, to help build the argument that institutions indeed cause differences in economic growth and to get a sense of the mechanisms through which this works.

A problem that will be very familiar to all economists trying to conduct comparative research on economic growth is the fact that most data on relevant variables is collected at the national level. National borders in many cases do not correspond to natural borders and are not very stable in the long run. Spillovers may take place more between regions in different countries sharing a border than between regions in the same country. Furthermore, if we want to link economic growth to the quality of institutions, we also know that the national level may not be the relevant level; every country has core and peripheral regions with respect to the institutional framework or ecosystem, and again these ecosystems may cross national borders or find themselves at lower levels of aggregation within countries (e.g. Tabellini, 2010, and Andersson and Koster, 2011). There is also evidence that institutions are linked to and move with societies or groups of people, not countries (Levitt, 1998; Ashraf and Galor, 2013b), therefore cultural drift and admixture should be taken into account (Ashraf and Galor, 2013a). If we want to explain economic growth today using differences in historic exogenous conditions, these factors are usually linked to a geographic location, not a country, as the composition and boundaries of countries change over time. There have been attempts to generate a proxy for economic activity, which would then be correlated to GDP, using light intensity at the Grid cell level (Chen and Nordhaus, 2011; Ghosh et al., 2013), but these are not uncontested. Furthermore, these analyses then also require data on independent variables that can be disaggregated to this level. For some variables this may be easily done (such as education or demographic structure), but for other variables, especially investment in physical capital, this is a great challenge and future puzzle.

Finally, one can question whether growth in welfare is best measured using regional or national GDP per capita data. These data have several limitations. First, the measures do not take into account the uncertain and unequal distribution, so the average does not indicate the likelihood that any individual will share in the prosperity or the degree of income security that individuals have (Osberg and Sharpe, 2002). Eicher and Turnovsky (2003) review the various ways in which policies aimed at fostering growth may have an impact on inequality, as the benefits of growth are rarely

shared equally. Second, the Easterlin Paradox —the phenomenon that while GDP per capita has steadily increased in developed economies, measures of happiness do not show any improvement after a certain point (Easterlin, 1974) —suggests that other, more comprehensive measures of (subjective) wellbeing are warranted (Eckersley, 2000; Kubiszewski et al., 2013). Finally, due to scarcity of non-renewable resources and other negative externalities of the increase in production volume by Western economies, some question whether the development of the Western world has come at the expense of current developing economies. They suggest that it should not be sustained GDP growth but sustainable, inclusive and 'green' growth that we strive for (Fay, 2012). We focus on GDP measures here as we are mainly interested in inequality at the macro-level and these are the most comparable data, however, we note that they may not be linked one-to-one to the broader concept of human wellbeing.

5.2.1 Institutional change and institutional inertia

This dissertation focuses to a large extent on the history of institutions, how they have emerged and diffused across societies. But what can we learn from this regarding future institutional change or stability? If institutional development is indeed endogenous to economic growth, does that mean that rich countries will find themselves in an upward spiral with an institutional framework that functions better and better to meet the needs of that society and attracts even more growth in productivity? Or can there also be an overinvestment in institutions, to the extent that they limit the freedom of societies to innovate and to change radically if necessary? As Fukuyama (2011, p. 12) claims:

> *"It is quite legitimate to argue that modern governments have grown excessively large, and that they thereby limit economic growth and individual freedom. [...] But in the developed world, we take the existence of government so much for granted that we sometimes forget how important it is, and how difficult it was to create, and what the world would look like without certain basic political institutions."*

A well-functioning institutional framework and a strong state provide stability, but this stability can also become a brake on economic progress. Circumstances change while there is no automatic mechanism through which political systems or institutions adjust. Structural changes to the system are very hard to implement as political elites tend to focus on short-term fixes,

instead of implementing changes that are painful in the short run or to certain groups in society. This mismatch between the institutional framework and the new circumstances can lead to tensions in society, and possibly even the collapse of this society or the overtake by other societies.

A question for the future is whether institutions will be able to adapt to new circumstances in an endogenous manner, rather than changing only as a consequence of war or conflict. Are disruptive times necessary to break free from suboptimal institutions or can they only mean harm now that many states went through long path-dependent ways of institution-building and processes of incremental change? As Tilly (1985) argues, war made the states, and states make war, but institutionalized states have not found a solution to end wars. Some people see a stateless society as an utopia, however, Fukuyama (2011, p. 13) notes:

> *"A market economy and high levels of wealth don't magically appear when you 'get government out of the way'; they rest on a hidden institutional foundation of property rights, rule of law, and basic political order."*

In Chapter 2 I showed that a demand for institutions does not automatically lead to the successful establishment of and sustained investment in institutions. In order to direct the critical mass to collectively divert resources, diffusion agents are critical. Furthermore, a society needs tools to distinguish cooperating members from non-cooperating members and to punish and exclude defectors. Religion fulfilled this role and was crucial to state formation in the past, but now large parts of the world are secularizing. States are more and more separated from religion, although they are still embedded in cultural traditions. New mechanisms of legitimization, like cultural habits, ideology, and nationalism, have taken over the role of religion (Fukuyama, 2011), but it remains to be seem whether this will affect the willingness of members to adhere to the state-level institutions.

Historical evidence shows that institutional change usually does not happen purely endogenously. In the past, drastic changes such as climatic catastrophes or war were necessary triggers. Organization of society has moved from the family as the central unit, to religious leaders, to national political leaders. What is next? Geographical boundaries are arguably becoming less important (Friedman, 2005), societies are more heterogeneous (Putnam, 2007), therefore culture may be more based on similarities within groups in society than on a national culture. As of now, it is unclear which mechanisms of legitimization will gain the upper hand. The last decade has seen a resurgence of ethnic conflict and nationalism (Smith, 2013). At

the same time, a countermovement is present in the form of religious extremism in some parts of the world, which is non state-related.

Levitsky and Murillo (2009) argue that institutional strength depends on two aspects; stability and enforcement, and they show that neither of those aspects should be taken for granted when analyzing developing countries. The institutionalist literature mostly views institutional change as a gradual process consisting of long continuities, periodically interrupted by radical shifts as a consequence of external shocks. In the case of institutional weakness, institutional change may look very different. Vicious cycles in which parchment rules frequently change and are not enforced may arise, inducing actors to behave in expectation of this and thereby leading to a self-fulfilling prophecy of repeated institutional change and instability.

Too much stability can also provide a source of problems. In the aftermath of World War II, Western countries were looking for institutions that provided stability. People were willing to give up individual and political freedom for that cause. This stability proved to be a seeding ground for growth and prosperity. However, after the financial crisis, when the welfare growth came to an end, again a demand for more freedom and flexibility has developed and societies are trying to break out of these rules and institutions.[2] Another issue that creates tension is the recent migration crisis, set in motion by conflict and wars in the Middle East. As Itzigsohn (2000, p. 1126) states, "Modern politics and citizenship have been organized around the coincidence of citizenship rights and the boundaries of states". Episodes of large scale migration always create tensions within this form of organization of citizenship. Furthermore, many institutions in Europe are based on the Scandinavian model of the welfare states, requiring high levels of solidarity among participants. It is unclear whether this solidarity will be extended to the new population of migrants who originate from different institutional settings. In short, the institutional challenge in many Western countries has changed from providing stability and preventing wars between states to coping with changing circumstances and demands due to the forces of globalization. Formal institutions at the national level have to reinvent themselves in order to fit the new pattern of heterogeneity with respect to culture, religion and informal institutions within the state.

[2]One could see the 2016 referendum in which Great Britain voted to leave the European Union as an example of this, although other motivations also may have intervened.

5.2.2 Policy implications and research agenda

Economic activity is unevenly distributed across space; this is measured by gaps in wealth, the density of population and economic activity, and compositions of regional and national economies. The motivation for studying institutions and institutional change is to help eliminate worldwide differences in prosperity and to understand how developing regions and countries can achieve greater economic growth through well-functioning institutions. When it comes to policy recommendations, however, caution is warranted. There is no "one size fits all" recipe for implementing successful institutions and past experiments of top-down imposed policy reform in developing countries have not been universally successful. It seems that there are winning strategies for development, but their evolution and implementation will vary depending on the historical setting of the country (Boeninger, 1991). There is a consensus now that trust and formal institutions are not substitutes, but complements. Policy and formal laws can only contribute to welfare if they are embedded in the informal institutional setting. Some countries find themselves in a second-best equilibrium of stability under unenforced formal institutions. If this situation has persisted over a longer period of time, inhabitants adjust their expectations and behavior on that, and an initial period of institutional failure or instability can lock a polity into a path of institutional weakness or an institutional instability trap (Grzymala-Busse, 2006; Helmke, 2007). Escaping from such a trap requires a deeper understanding of the relation between formal and informal institutions. It is especially important to understand how informal institutions change or move together with movement of people and contact, in the context of migration.

Another issue that warrants further research is the tension between creating growth and reducing inequality. Alleged differences in institutional quality are an important cause of persistent inequality, and these do not seem to disappear with greater integration (Rodríguez-Pose, 1999; Puga, 2002). Policies that stimulate growth may therefore very well increase the level of inequality and underdevelopment (Farole et al., 2011). Institutions could improve economic performance and reduce underdevelopment through varying mechanisms, but this too may lead to conflicting policy goals. For example, institutions focused on long staying power and stability may promote investment and thereby growth, but they can also hamper innovation, which requires an institutional setting that is focused on novelty and openness. In order for research to come up with a manual for policy reform for regions or countries, both the starting position and the

objectives thus have to be clearly identified. Institutions may be either too inert to react to changing circumstances or too instable for actors to invest in sources of future growth. Furthermore, is the goal of policy reform reducing overall inequality, increasing economic growth (according to Grow the Pie theories) or fighting underdevelopment? Growth and equity may be hard to achieve at the same time, although suboptimal institutions hurt both these goals at the same time.

In this dissertation I have focused on the interactions between people and geography that created today's pattern of institutions. Geography determines the initial conditions upon which institutions emerge, and it also affects institutional change through the influence of neighbours. While the influence of geography is persistent, its importance may decrease over time as nation-states are becoming more fluid and barriers to free movement of people and production factors are disappearing at certain scales. Future research thus has to acknowledge, and deepen our understanding of, the ways in which contact between people and between nations influences the diffusion of institutions and how that may break or reinforce core-periphery patterns of institutional quality and prosperity. This will require a multidisciplinary approach that combines at the very least insights from economics, sociology and migration studies.

6

Appendix

6.1 Appendix Chapter 2

On page 34, the following household maximization problem is put forward:

$$\begin{aligned} &\underset{c_t, n_t}{\text{maximize}} \quad U_t = {n_t}^{\gamma} {c_t}^{1-\gamma} \\ &\text{subject to} \qquad\qquad (1-\tau_t) z_t = \rho n_t + c_t \end{aligned}$$

Where U is utility which is derived from n, the number of offspring and c, consumption where $0 < \gamma < 1$ is the elasticity for offspring in utility. The budget constraint states that after-tax disposable income per person, $(1-\tau_t)z_t$, is spent on consumption and child rearing, where we normalised the price of consumption to 1, ρ is the fixed cost of raising a child and $0 < \tau < 1$ is the tax rate.[1]

We formulate the following Lagrangean for the maximization problem:

$$\mathcal{L} = {n_t}^{\gamma} {c_t}^{1-\gamma} + \lambda \left[\frac{1}{1-\tau_t} (\rho n_t + c_t) - z_t \right]$$

This gives us three First Order Conditions:

$$\frac{\partial \mathcal{L}}{\partial c_t} = (1-\gamma) n_t^{\gamma} c^{-\gamma} + \frac{1}{1-\tau_t} \lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_t} = \gamma n_t^{\gamma-1} c_t^{1-\gamma} + \frac{\rho}{1-\tau_t} \lambda = 0$$

[1] To be very precise, it is the share of disposable income collected and allocated to building and maintaining institutions in society.

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \frac{1}{1-\tau_t}(\rho n_t + c_t) - z_t = 0$$

Rewriting the first two equations gives you two formulas for λ:

$$(1-\gamma) n_t^{\gamma} c_t^{-\gamma} (1-\tau_t) = -\lambda \tag{6.1.1}$$

$$\frac{1}{\rho} \gamma n_t^{\gamma-1} c_t^{1-\gamma} (1-\tau_t) = -\lambda \tag{6.1.2}$$

Which can be solved as

$$\rho n_t = \frac{\gamma}{1-\gamma} c_t \tag{6.1.3}$$

Using the third F.O.C. we can get another condition for c:

$$c_t = (1-\tau_t) z_t - \rho n_t \tag{6.1.4}$$

Using equation 6.1.3 and 6.1.4 together we can solve for consumption as a function of output, which gives equation 2.3.1:

$$c_t* = (1-\tau_t)(1-\gamma) z_t$$

We can then use equation 6.1.4 and the optimal consumption function to solve for n_t, which gives equation 2.3.2:

$$n_t* = \frac{\gamma}{\rho}(1-\tau_t) zt$$

6.1.1 Social Evolution

The derivation of equation 2.3.9 using equation 2.3.8 is shown below. We chose the following specification for θ_t:

$$\theta_t = (\frac{\tau_t z_t}{L_t})^{\xi} = (\frac{\tau_t \theta_t (AX)^{\alpha} L_t^{1-\alpha}}{L_t})^{\xi} \tag{6.1.5}$$

We solve for θ:

$$\theta_t = (\tau_t \theta_t (AX)^{\alpha} L_t^{-\alpha})^{\xi}$$

$$\theta_t^{1-\xi} = \tau_t^{\xi} (\frac{AX}{L_t})^{\alpha\xi}$$

Simplifying and solving gives equation 2.3.9:

$$\theta_t = \left(\tau_t (AX)^{\alpha} L_t^{-\alpha}\right)^{\frac{\xi}{1-\xi}}$$

The calculation of the steady state from formula 2.3.10 follows below:

$$L_{t+1} = \frac{\gamma}{\rho}(1-\tau_t)\left(\tau_t (AX)^{\alpha} L_t^{-\alpha}\right)^{\frac{\xi}{1-\xi}} (AX)^{\alpha} L_t^{1-\alpha} \tag{6.1.6}$$

In the steady state it holds that $L_{t+1} = L_t = \bar{L}$, we plug this in.

$$\bar{L} = \frac{\gamma}{\rho}(1-\tau_t)\left(\tau_t (AX)^{\alpha} \bar{L}^{-\alpha}\right)^{\frac{\xi}{1-\xi}} (AX)^{\alpha} \bar{L}^{1-\alpha}$$

$$\bar{L}^{\alpha} = \frac{\gamma}{\rho}(1-\tau_t)\left(\tau_t (AX)^{\alpha} \bar{L}^{-\alpha}\right)^{\frac{\xi}{1-\xi}} (AX)^{\alpha}$$

$$\bar{L}^{\alpha} = \frac{\gamma}{\rho}(1-\tau_t)\tau_t^{\frac{\xi}{1-\xi}} (AX)^{\frac{\alpha\xi}{1-\xi}} \bar{L}^{\frac{-\alpha\xi}{1-\xi}} AX^{\alpha}$$

$$\frac{\bar{L}^{\alpha}}{\bar{L}^{\frac{-\alpha\xi}{1-\xi}}} = \frac{\gamma}{\rho}(1-\tau_t)\tau_t^{\frac{\xi}{1-\xi}} (AX)^{\frac{\alpha\xi}{1-\xi}+\alpha}$$

$$\bar{L}^{\frac{\alpha}{1-\sigma}} = \frac{\gamma}{\rho}(1-\tau_t)\tau_t^{\frac{\xi}{1-\xi}} (AX)^{\frac{\alpha}{1-\xi}}$$

$$\bar{L} = \frac{\gamma}{\rho}^{\frac{1-\xi}{\alpha}} (1-\tau_t)^{\frac{1-\xi}{\alpha}} \tau_t^{\frac{\xi(1-\xi)}{\alpha(1-\xi)}} (AX)^{\frac{\alpha(1-\xi)}{\alpha(1-\xi)}}$$

Which simplifies to equation 2.3.12

$$\bar{L}^{*} = AX\left(\frac{\gamma}{\rho}\right)^{\frac{1-\xi}{\alpha}} (1-\tau)^{\frac{1-\xi}{\alpha}} \tau^{\frac{\xi}{\alpha}} \tag{6.1.7}$$

This equation is maximized to obtain the optimal tax rate:

$$\underset{\tau_t}{\text{maximize}} \quad \bar{L} = (AX)\frac{\gamma}{\rho}^{\frac{1-\xi}{\alpha}} (1-\tau_t)^{\frac{1-\xi}{\alpha}} \tau^{\frac{\xi}{\alpha}}$$

$$\frac{\partial}{\partial \tau_t}\left[AX\frac{\gamma}{\rho}^{\frac{1-\xi}{\alpha}} (1-\tau_t)^{\frac{1-\xi}{\alpha}} \tau_t^{\frac{\xi}{\alpha}}\right] = AX\frac{\gamma}{\rho}^{\frac{1-\xi}{\alpha}} \frac{\partial}{\partial \tau_t}\left[(1-\tau_t)^{\frac{1-\xi}{\alpha}} \tau_t^{\frac{\xi}{\alpha}}\right]$$

$$\frac{\partial}{\partial \tau_t} = 0 \rightarrow (AX)\frac{\gamma}{\rho}^{\frac{1-\xi}{\alpha}} \left[\frac{(1-\tau_t)^{\frac{1-\xi}{\alpha}} \tau_t^{-1+\frac{\xi}{\alpha}} \xi}{\alpha} - \frac{(1-\tau_t)^{-1+\frac{1-\xi}{\alpha}} \tau_t^{\frac{\xi}{\alpha}} (1-\xi)}{\alpha}\right] = 0$$

$$(1-\tau_t)^{1-\frac{\xi}{\alpha}}\tau_t^{-1+\frac{\xi}{\alpha}}\xi = (1-\tau_t)^{-1+\frac{1-\xi}{\alpha}}\tau_t^{\frac{\xi}{\alpha}}(1-\xi)$$

$$\frac{(1-\tau)}{\tau} = \frac{(1-\xi)}{\xi} \rightarrow \tau* = \xi$$

6.2 Appendix Chapter 3

Figure 6.2.1: Distribution over time of the first year of full female franchise for countries in sample (N=98)

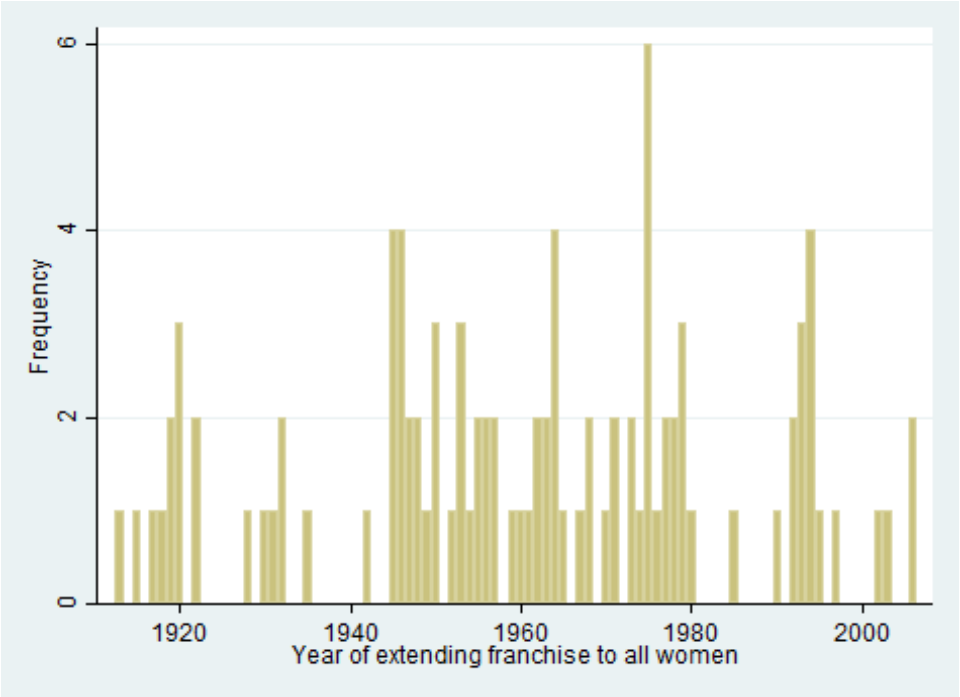


Table 6.1: Descriptive statistics

Variable	Mean	Std. Dev	Min	Max	Source	Time	Description
Female Franchise	1959.96	17.48	1913	1997	Beck	1893-2006	First year of full voting rights for women
Distance	1559.84	1726.22	0.00	8723.39	Gleditsch and Ward	1893-2006	Mean distance to all prior adopters
Close Neighbours	0.25	0.67	0	7	Gleditsch and Ward	1893-2006	Count of prior adopters within 640 km radius
Tradeshare	41.44	37.80	0	100	Barbieri	1893-2006	Percentage of total trade volume with prior adopters
Waryieldlost	0.05	0.27	0	4	Sarkees	1893-2006	Count of wars lost or yielded to prior adopters
Diplomatic Exchange	6.37	10.31	0	84	Bayer	1893-2006	Count of diplomatic exchanges with prior adopters
Colonial Dependency	0.30	0.64	0	4	Correlates of War	1893-2006	Count of colonial relations with prior adopters
Cultural Distance	0.02	0.03	0.00	0.16	Melitz et al.	1893-2006	Mean linguistic proximity to prior adopters
WWI	0.01	0.11	0	1	Bruns	Time-invariant	Dummy for ending of WWI (1919)
WWII	0.01	0.11	0	1	Bruns	Time-invariant	Dummy for ending of WWII (1946)
GDP per Capita	2891.68	3751.06	525	42916	Beck	1893-2006	Annual GDP per capita
Openness	0.39	1.59	0.00	29.38	Bruns	1893-2006	Ratio of Trade Volume to GDP
Ongoing War	0.13	0.34	0	1	Correlates of War	1893-2006	Dummy if a country is involved in a war in that year
Common Law	0.02	0.14	0	1	CIA World Factbook	Time-invariant	Dummy if the legal system is based on common law
Gender Equality Index	0.58	0.14	0.23	0.85	Human Development Report	2008	Index of loss of achievement due to gender inequality
Sharebuddhist	0.03	0.11	0.00	0.56	ARDA	2010	Share of Buddhists in total population
Sharechristian	0.04	0.13	0.00	0.62	ARDA	2010	Share of Christians in total population
Sharehindus	0.00	0.04	0.00	0.76	ARDA	2010	Share of Hindus in total population
Sharemuslims	0.26	0.44	0.00	1.50	ARDA	2010	Share of Muslims in total population
Shareatheist	0.01	0.02	0.00	0.07	ARDA	2010	Share of Atheists in total population
Shareagnostic	0.05	0.06	0.00	0.32	ARDA	2010	Share of Agnostics in total population
Polity	-2.36	5.57	-10	10	Marshall et al.	1893-2006	Composite indicator of full autocracy (-10) to full democracy (+10)
Lagyears	13.81	17.28	0	65	Bruns	1893-2006	Years since extension of franchise to all males
Plough absent	0.02	0.07	0.00	0.96	Bolt	Time-invariant	Dummy for historical absence of plow cultivation
Migrate	-0.93	2.31	-9.52	11.68	CIA World Factbook	2010	Net Migration rate
Regional Fixed effects			0	1	Bruns	Time-invariant	Dummy for each continent

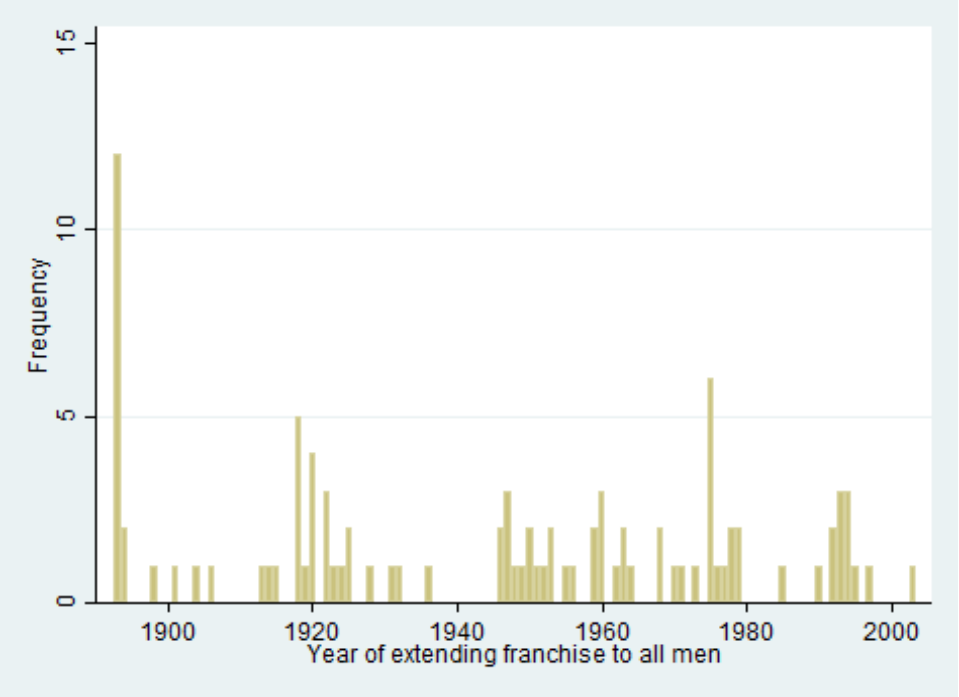
Table 6.2: Countries and adoption year in unrestricted sample (n=101)

Country	Year	Country	Year	Country	Year
Algeria	1962	Guinea Bissau	1975	Sao Tome and Principe	1975
Afghanistan	1964	Haiti	1950	Saudi Arabia	N.A.
Albania	1946	Honduras	1954	Slovakia	1993
Angola	1975	Hungary	1920	Solomon Islands	1978
Argentina	1947	Indonesia	1953	Somalia	1964
Bahrain	2002	Iran	1979	South Africa	1968
Belarus	1994	Iraq	1980	Spain	1931
Belgium	1948	Israel	1948	Sudan	1964
Bhutan	1994	Italy	1945	Sweden	1919
Bolivia	1952	Japan	1945	Switzerland	1971
Bosnia and Herzegovina	1995	Jordan	1974	Syria	1973
Brazil	1985	Kuwait	2006	Tajikistan	1994
Brunei	N.A.	Kyrgyzstan	1993	Tanzania	1963
Bulgaria	1945	Laos	1957	Tunisia	1959
Cambodia	1956	Latvia	1922	United Arab Emirates	2006
Cape Verde	1975	Libya	1963	United Kingdom	1928
Chile	1970	Lithuania	1922	United States	1920
China	1947	Mexico	1953	Uruguay	1932
Colombia	1957	Moldova	1994	Uzbekistan	1992
Comoros	1975	Morocco	1962	Venezuela	1946
Costa Rica	1949	Mozambique	1975	Vietnam	1976
Cuba	1935	Netherlands	1919	Yemen	1990
Cyprus	1960	Nicaragua	1955	Zaire	1967
Czech Republic	1993	Nigeria	1977	Zambia	1964
Denmark	1915	Norway	1913	Zimbabwe	1979
Dominican Republic	1942	Oman	2003		
Ecuador	1978	Pakistan	1971		
Egypt	1956	Panama	1946		
El Salvador	1950	Papua New Guinea	1977		
Eritrea	1997	Paraguay	1961		
Estonia	1920	Peru	1979		
Ethiopia	1955	Philippines	1973		
France	1945	Portugal	1968		
Germany	1918	Qatar	N.A.		
Greece	1953	Romania	1946		
Guatemala	1965	Russia	1917	Mean	1961.13

Table 6.3: Countries and adoption year in restricted sample (n=55)

Country	Year	Country	Year
Afghanistan	1947	Kyrgyzstan	1993
Albania	1946	Mexico	1953
Bahrain	2002	Moldova	1994
Belgium	1948	Netherlands	1919
Brazil	1985	Nicaragua	1955
Cambodia	1956	Norway	1913
Chile	1970	Panama	1946
China	1947	Paraguay	1961
Colombia	1957	Philippines	1973
Costa Rica	1949	Portugal	1968
Cuba	1935	Slovakia	1993
Czech Republic	1993	Spain	1931
Denmark	1915	Sweden	1919
Ecuador	1978	Switzerland	1971
El Salvador	1950	Syria	1973
Greece	1953	Tajikistan	1994
Guatemala	1965	Turkey	1930
Haiti	1950	United Kingdom	1928
Honduras	1954	United States	1920
Hungary	1920	Uruguay	1932
India	1950	Venezuela	1946
Indonesia	1953	Vietnam	1962
Iraq	1980	Algeria	1962
Jordan	1974	Egypt	1956
Kuwait	2006	Lybia	1963
		Mozambique	1976
		South Africa	1968
		Tunisia	1968
		Zaire	1967
		Zimbabwe	1979
		Mean	1960.62

Figure 6.2.2: Distribution over time of the first year of full male franchise for countries in sample (N=98)



6.3 Appendix Chapter 4

Derivation of Equation 4.3.2
Starting form the production function in equation 4.3.2:

$$Y_i(t) = K_i(t)^{\alpha} H_i(t)^{\alpha} (A_i(t)L_i(t))^{1-\alpha-\beta} \tag{6.3.1}$$

Labour and Technology grow at rates n and g:

$$L(t) = L(0)e^{nt} \tag{6.3.2}$$
$$A(t) = A(0)e^{gt} \tag{6.3.3}$$

$A(t)L(t)$ is the number of effective units of labour and grows at rate $n+g$. We define $y \equiv Y/AL$, $k \equiv K/AL$ and $h \equiv H/AL$. The evolution of the economy is determined by the following two equations:

$$\dot{k}(t) = s^k y(t) - (n+g+\delta)k(t) \tag{6.3.4}$$

$$\dot{h}(t) = s^h y(t) - (n + g + \delta)h(t) \tag{6.3.5}$$

where s^k is the fraction of income invested in physical capital and s^h the fraction of income invested in human capital. MRW make the assumption that one unit of consumption can be transformed into one unit of human or physical capital without transaction costs. Furthermore, MRW assume the same depreciation rate δ holds for both types of capital and decreasing returns to all capital, i.e. $\alpha + \beta < 1$. This model gives a prediction for the steady state, and much of the differences in income between countries or regions is due to differences in determinants of the steady state. However, we are interested in growth rates and want to allow for the possibility that economies are out of the steady state. Therefore we use MRW's log-approximation of the speed of convergence around the steady state:

$$\frac{\partial ln(y(t))}{\partial t} = \lambda[ln(y*) - ln(y(t))], \tag{6.3.6}$$

with $\lambda = (n + g) + \delta)(1 - \alpha - \beta)$. Integrating this expression gives us an expression for the derivative of $ln(y(t))$ with respect to t in terms of $y(0)$. When we plug this into equation 6.3.6, we get the following equation:

$$ln(y(t)) = (1 - e^{-\lambda t} ln(y*) + e^{-\lambda t} ln(y(0)), \tag{6.3.7}$$

where $y(0)$ is GDP per effective worker at the beginning of the period under study. By subtracting $y(0)$ from both sides and substituting for the expression found for $y*$, MRW end up with the following specification:

$$\begin{aligned} ln(y(t)) - ln(y(0)) = (1 - e^{-\lambda t})\frac{\alpha}{1 - \alpha - \beta} ln(s^k) + \\ (1 - e^{-\lambda t})\frac{\beta}{1 - \alpha - \beta} ln(s^h) - \\ (1 - e^{-\lambda t})\frac{\alpha + \beta}{1 - \alpha - \beta} ln(n+g + \delta) - (1 - e^{-\lambda t}) ln(y(0)) \end{aligned}$$

This means that in the augmented Solow model we can estimate the parameters of the model by regressing the average growth of income on the income share of physical capital, the income share of human capital, the growth rates of population and technology, the depreciation rate and the initial level of income. Taking technical change g and depreciation δ to be the same across our European regions, we can estimate equation 4.3.2:

$$\frac{lnY_i(T) - lnY_i(T0)}{T - T0} = \alpha_0 + \alpha_1 s_i^k + \alpha_2 s_i^h + \alpha_3 n_i + \alpha_4 lnY_i(T0) + \epsilon_i \quad (6.3.8)$$

Table 6.4: Selecting regimes - OLS residuals without regional controls

TEA	Log-likelihood	Obs	Parameters	AIC	Significance of Entrepreneurship
1 class	-183.643	107	2	**0.632**	No
2 classes	-174.819	107	7	3.398	-, +
3 classes	-168.748	107	11	3.359	-, +, No
TEA_{in}					
1 class	-183.472	107	2	**0.629**	No
2 classes	-177.068	107	7	3.441	+, No
3 classes	-170.735	107	11	3.397	No, No, No
TEA_{hj}					
1 class	-177.588	107	2	**0.519**	+
2 classes	-173.142	107	7	3.367	+, +
3 classes	-166.580	107	11	3.319	+, +, +

Regions in Sample

- **Greece**
- Nisi Aigaiou/Kriti
- Voreia Ellada
- Kentriki Ellada

- **Netherlands**
- Noord Nederland
- Oost-Nederland
- West-Nederland
- Zuid-Nederland

- **Belgium**
- Brussel
- Vlaams Gewest
- Region Wallone

- **France**
- Ile de France
- Parisien Bassin
- Nord-Pas de Calais
- East
- West
- South-West
- Center-East
- Mediterranee

- **Spain**
- Galicia
- Asturias
- Cantabria
- Pais vasco
- Navarra
- La Rioja
- Aragon
- Madrid
- Castilla y León
- Castilla La Mancha
- Extremadura
- Catalunya
- Comm Valenciana
- Baleares
- Andalucia
- Murcia
- Canarias

- **Hungary**
- Central Hungary
- Central Transdanubia
- Western Transdanubia
- Southern Transdanubia
- Northern Hungary
- Northern Great Plain
- Southern Great Plain

- **Italy**
- Nord-Ovest
- Nord-Est

- Centro
- Sud
- Isole

- **United Kingdom**
- Scotland
- North East
- North West
- Yorkshire Humberside
- East Midlands
- West Midlands
- East Anglia
- Greater London
- South East
- South West
- Wales
- Northern Ireland

- **Denmark**
- Copenhagen area
- Sjaelland
- Syddanmark
- Nordjylland
- Midtjylland

- **Sweden**
- Stockholm area
- Östra Mellansverige
- Sydsverige
- Norra Mellansverige
- Mellersta Norrland
- Övre Norrland
- Småland med öarna
- Västsverige

- **Norway**
- Oslo and surroundings
- Hedmark
- Sor-ostlandet
- Agder og Rogaland
- Vestlandet

- **Germany**
- Baden-Württemberg
- Bayern
- Berlin
- Brandenburg
- Bremen
- Hamburg
- Hessen
- Mecklenburg-Vorpommern
- Niedersachsen
- Nordrhein-Westfalen
- Rheinland - Pfalz
- Saarland
- Sachsen
- Sachsen-Anhalt
- Schleswig-Holstein

- Thüringen

- **Portugal**
- Norte (incl Porto)
- Algarve
- Centro
- Lisboa e Vale de Tejo
- Alentejo

- **Ireland**
- Border, Midlands and Western
- Southern and Eastern

- **Finland**
- South (Etela-Suomi)
- West/Middle (Lansi-Suomi)
- North (Ita ja Pohjois)
- Helsinki Area (Uusima)

- **Slovenia**
- Vzhodna
- Zahodna

Table 6.5: Cross-correlation table

Variables	ΔlnY	$lnY(2005)$	s^k	s^h	n	YNG	$SPEC$	$POPDENS$	$NEIGH_n$	$NEIGH_{s^k}$	$NEIGH_{s^h}$	TEAYY	TEA_{hj}	TEA_{in}
ΔlnY	1.0000													
$lnY(2005)$	0.0389	1.0000												
s^k	-0.0334	-0.3589*	1.0000											
s^h	-0.0432	-0.0755	-0.0393	1.0000										
n	0.2507*	0.2409*	-0.1268	0.2207*	1.0000									
YNG	-0.1508	-0.1311	-0.0135*	0.6293*	0.0860	1.0000								
$SPEC$	-0.4737*	0.0635	-0.0614	-0.0051	0.1084	0.2233*	1.0000							
$POPDENS$	0.1118	0.2283*	-0.1643	0.2518*	0.2686*	0.0945	-0.0144	1.0000						
$NEIGH_n$	0.0274	0.1544	-0.1034	-0.0528	0.6793*	-0.0889	0.1106	0.0261	1.0000					
$NEIGH_{s^k}$	0.0501	-0.1542	0.0408	-0.2597*	-0.3473*	-0.2368*	-0.0035	0.1290	-0.4400*	1.0000				
$NEIGH_{s^h}$	-0.2681*	-0.2463*	-0.0208	0.2709*	0.0087	0.1029	0.0325	-0.1643	0.1406	-0.0805	1.0000			
TEA	0.0048	-0.1134	0.0080*	0.0997	0.0714	0.0886	0.1071	0.1600	-0.0699	-0.1200	0.1973*	1.0000		
TEA_{hj}	0.3260*	0.2333*	-0.0088	-0.0244	0.0655	-0.0190	-0.1867	0.3292*	-0.1142	0.0314	-0.2578*	0.4773*	1.0000	
TEA_{in}	0.1159	0.1829	-0.0051	0.0590	0.2654*	0.0464	0.0783	0.1721	-0.1084	0.1240	-0.1046	0.5764*	0.6287*	1.0000

References

Acemoglu, D. (2009). *Introduction to modern economic growth*. MIT press.

Acemoglu, D., S. Johnson, and J. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*, 1369–1401.

Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth 1*, 385–472.

Acemoglu, D. and J. A. Robinson (2012). *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. Crown Publishers, Croydon.

Acs, Z. J. and M. Sanders (2012). Patents, knowledge spillovers, and entrepreneurship. *Small Business Economics 39*(4), 801–817.

Acs, Z. J., L. Szerb, and . Autio, E. (2014). National systems of entrepreneurship: Measurement and policy. *Research Policy Implications 43*, 476–494.

Adner, R. and R. Kapoor (2010). Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations. *Strategic Management Journal 31*, 306–333.

Aghion, P. and P. Howitt (1992). A model of growth through creative destruction. *Econometrica 60*, 323–351.

Aidt, T. S. and B. Dallal (2008). Female voting power: The contribution of women's suffrage to the growth of social spending in western europe (1869–1960). *Public Choice 134*(3-4), 391–417.

Aidt, T. S., J. Dutta, and E. Loukoianova (2006). Democracy comes to europe: Franchise extension and fiscal outcomes 1830–1938. *European Economic Review 50*(2), 249–283.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics 128*, 469–530.

Añón Higón, D. and A. D. Minguela (2014). Womens suffrage and political representation, 1893-2010: Economic development and women empowerment. *Working paper at Centre for Global and Economic History (CGEH) Workshop*, 1–27.

Andersson, M. and M. Henrekson (2015). Local competitiveness fostered through local institutions for entrepreneurship. In D. B. Audretsch, A. N. Link, and M. L. Walshok (Eds.), *The Oxford Handbook of Local Competitiveness*, Chapter 8, pp. 145–190. Oxford, United Kingdom: Oxford University Press.

Andersson, M. and S. Koster (2011). Sources of persistence in regional start-up rates - evidence from Sweden. *Journal of Economic Geography 11*, 179–201.

ARDA (2011). Association of religion data archives. Technical report, www.thearda.com.

Armelagos, G. J. and M. N. Cohen (1984). *Paleopathology at the Origins of Agriculture*. Academic Press.

Ashraf, Q. and O. Galor (2011). Dynamics and stagnation in the Malthusian Epoch. *American Economic Review 101*, 2003–2041.

Ashraf, Q. and O. Galor (2013a). Genetic diversity and the origins of cultural fragmentation. *The American economic review 103*(3), 528–533.

Ashraf, Q. and O. Galor (2013b). The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development. *The American Economic Review 103*(1), 1–46.

Ashraf, Q. and S. Michalopoulos (2015). Climatic fluctuations and the diffusion of agriculture. *The Review of Economics and Statistics 97*, 589–609.

Audretsch, D. (1995). *Innovation and Industry Evolution*. MIT Press, Massachusetts.

Audretsch, D., O. Falck, M. Feldman, and S. Heblich (2012). Local entrepreneurship in context. *Regional Studies 46*, 379–389.

Audretsch, D. B., A. N. Link, and M. L. Walshok (Eds.) (2015). *The Oxford Handbook of Local Competitiveness*. Oxford, United Kingdom: Oxford University Press.

Autio, E. (1998). Evaluation of RTD in regional systems of innovation. *European Planning Studies 6*(2), 131–140.

Bar-Yosef, O. (1998). On the nature of transitions: The Middle to Upper Paleolithic and the Neolithic Revolution. *Cambridge Archeological Journal 8*, 141–163.

Barbieri, K., O. Keshk, and B. Pollins (2008). Correlates of War Project Trade data set codebook. *Codebook Version 2*. Version 3.0.

Barbieri, K. and G. Schneider (1999). Globalization and peace: Assessing new directions in the study of trade and conflict. *Journal of Peace Research 36*(4), 387–404.

Baumol, W. (1990). Entrepreneurship: Productive, unproductive, and destructive. *Journal of political economy*, 893–921.

Bayer, R. (2006). The Correlates of War Diplomatic Exchange data set. The Correlates of War Project. V2006.1.

Beck, T., G. Clarke, A. Groff, P. Keefer, and P. Walsh (2001). New tools in comparative political economy: The Database of Political Institutions. *World Bank Economic Review 15*(1), 165–176. Updated version accessed in 2015.

Becker, G. S. and H. G. Lewis (1974). Interaction between quantity and quality of children. In *Economics of the family: Marriage, children, and human capital*, pp. 81–90. University of Chicago Press.

Becker, G. S., K. M. Murphy, and R. Tamura (1994). Human capital, fertility, and economic growth. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)*, pp. 323–350. The University of Chicago Press.

Beine, M., F. Docquier, and M. Schiff (2013). International migration, transfer of norms and home country fertility. *Canadian Journal of Economics/Revue canadienne d'économique 46*(4), 1406–1430.

Bennett, D. S. and A. C. Stam (2000). Eugene: A conceptual manual. *International Interactions 26*(2), 179–204.

Bernstein, T. P. and X. Lü (2003). *Taxation without representation in contemporary rural China*. Cambridge University Press.

Bertocchi, G. (2011). The enfranchisement of women and the welfare state. *European Economic Review 55*(4), 535–553.

Binmore, K. (1994). *Game Theory and the Social Contract, Vol. I: Playing Fair.* The MIT Press, Cambridge.

Binmore, K. (2011). *Natural Justice*. Oxford University Press.

Birdsell, J. B. (1957). Some population problems involving Pleistocene man. In *Cold Spring Harbor Symposia on Quantitative Biology*, Volume 22, pp. 47–69. Cold Spring Harbor Laboratory Press.

Blanton, R. E. (1975). The cybernetic analysis of human population growth. *Memoirs of the Society for American Archaeology* (30), 116–126.

Boeninger, E. (1991). Governance and development: Issues and constraints. *The World Bank Economic Review 5*, 267 – 288.

Bolt, J. (2012). A new database on the origins of institutional development. University of Groningen working paper.

Borck, R. (2014). Adieu rabenmutter - culture, fertility, female labour supply, the gender wage gap and childcare. *Journal of Population Economics 27*(3), 739–765.

Boserup, E. (1970). *Woman's Role in Economic Development*. London: George Allen and Unwin Ltd.

Bosma, N., A. van Stel, and K. Suddle (2008). The geography of new firm formation: Evidence from independent start-ups and new subsidiaries in the netherlands. *International Entrepreneurship and Management Journal 4*(2), 129–146.

Bosma, N. S. (2009). *The Geography of entrepreneurial activity and regional economic development*. Ph. D. thesis, PhD Series at Urban and Regional research centre Utrecht.

Bosma, N. S. (2013). The global entrepreneurship monitor and its impact on entrepreneurship research. *Foundations and Trends in Entrepreneurship 9*(2), 143–248.

Bowles, S. (2011). Cultivation of cereals by the first farmers was not more productive than foraging. *Proceedings of the National Academy of Sciences 108*(12), 4760–4765.

Boyer, R. and A. Orléan (1992). How do conventions evolve? *Journal of Evolutionary Economics*.

Brakman, S. and C. v. Marrewijk (2008). It's a big world after all: on the economic impact of location and distance. *Cambridge Journal of Regions, Economy and Society*, 411–437.

Braun, D. P. and S. Plog (1982). Evolution of "tribal" social networks: Theory and prehistoric north american evidence. *American Antiquity*, 504–525.

Breschi, S. and F. Lissoni (2003). Mobility and social networks: Localised knowledge spillovers revisited. *CESPRI Working Papers 142*.

Brinks, D. and M. Coppedge (2006). Diffusion is no illusion neighbor: Emulation in the third wave of democracy. *Comparative Political Studies 39*(4), 463–489.

Brosnan, S. F. (2006). Nonhuman species' reactions to inequity and their implications for fairness. *Social Justice Research 19*(2), 153–185.

Cao, X. (2010). Network as channels of policy diffusion: Explaining worldwide changes in capital taxation, 1998-2006. *International Studies Quarterly 54*, 823–854.

Carree, M. and A. Thurik (1998). Small firms and economic growth in Europe. *Atlantic Economic Journal 26*(2), 137–146.

Carree, M. and A. Thurik (2010). The impact of entrepreneurship on economic growth. *Handbook of Entrepreneurship Research*, 557–594.

Chen, X. and W. D. Nordhaus (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences 108*(21), 8589–8594.

Childe, V. G. (1936). *Man makes itself.* Mentor Books, New York.

Christopherson, Susan, G. H. and R. Martin (2008). The world is not flat: putting globalization in its place. *Cambridge Journal of Regions, Economy and Society*, 343–349.

Chwe, M. S.-Y. (2013). *Rational ritual: Culture, coordination, and common knowledge*. Princeton University Press.

CIA (2013-2014). The World Factbook. Washington D.C. https://www.cia.gov/library/publications/the-world-factbook/index.html.

Ciccone, A. (2002). Agglomeration effects in europe. *European Economic Review 46*(2), 213–227.

Clark, J. B. (1918). *Essentials of Economic Theory: As applied to modern problems of industry and public policy*. Macmillan.

Coase, R. H. (1960). Problem of social cost, the. *Journal of Law and Economics 3*, 1–44.

Coase, R. H. (1964). The regulated industries: Discussion. *American Economic Review 54*, 192–197.

Cook, K. (2001). *Trust in society*. Russell Sage Foundation.

Cooke, P. (2002). Regional innovation systems: General findings and some new evidence from biotechnology clusters. *The Journal of Technology Transfer 27*(1), 133–145.

Cosmides, L., J. Tooby, and J. H. Barkow (1992). Introduction: Evolutionary psychology and conceptual integration. In . J. T. e. Jerome Barkow, Leda Cosmides (Ed.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press.

CoW (2002). Colonial dependency and contiguity data, 1816-2002f. Correlates of War 2 Project. Version 3.0 Online: http://correlatesofwar.org.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 187–220.

Damjanovic, T., G. Selvaretnam, et al. (2015). Economic growth and evolution of gender equality. Technical report.

Davis, G. F. (1991). Agents without principles? The spread of the poison pill through the intercorporate network. *Administrative Science Quarterly 4*, 583–613.

Davis, L. E., D. C. North, and C. Smorodin (1971). *Institutional change and American economic growth*. CUP Archive.

De Waal, F. (2009). *Primates and philosophers: How morality evolved*. Princeton University Press.

De Waal, F. B. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol. 59*, 279–300.

Diamond, J. (1997). *Guns, Germs, and Steel: The Fates of Human Societies.* W.W. Norton and Company.

Diamond, J. (2005). *Collapse: How societies choose to fail or succeed*. Penguin.

Dietz, R. D. (2002). The estimation of neighborhood effects in the social sciences: An interdisciplinary approach. *Social science research 31*(4), 539–575.

Docquier, F., E. Lodigiani, H. Rapoport, and M. Schiff (2015). Emigration and democracy. *Journal of Development Economics*.

Dogaru, T., F. Van Oort, and M. Thissen (2011). Agglomeration economies in European regions: Perspectives for objective 1 regions. *Tijdschrift voor economische en sociale geografie 102*(4), 486–494.

Doloreux, D. (2002). What we should know about regional systems of innovation. *Technology in society 24*(3), 243–263.

Dunbar, R. I. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and brain sciences 16*(04), 681–694.

Durham, W. H. (1991). *Coevolution: Genes, culture, and human diversity*. Stanford University Press.

Easterlin, R. A. (1974). Does economic growth improve the human lot? some empirical evidence. *Nations and households in economic growth 89*, 89–125.

Easterly, W. (2001). *The elusive quest for growth: economists' adventures and misadventures in the tropics*. MIT press Cambridge, MA.

Eckersley, R. (2000). The mixed blessings of material progress: Diminishing returns in the pursuit of happiness. *Journal of Happiness Studies 1*(3), 267–292.

Edquist, C. (Ed.) (1997). *Systems of Innovation: Technologies, Institutions and Organizations*. Hove, United Kingdom: Psychology Press.

Eicher, T. S. and S. J. Turnovsky (2003). *Inequality and growth: theory and policy implications*, Volume 1. Mit Press.

Eller, E., J. Hawks, and J. H. Relethford (2009). Local extinction and recolonization, species effective population size, and modern human origins. *Human biology 81*(5), 805–824.

Ensminger, J. (1997). Changing property rights: Reconciling formal and informal rights to land in africa. *The frontiers of the new institutional economics*, 165–196.

Eriksson, R. H., H. K. Hansen, and U. Lindgren (2013). The importance of business climate and people climate on regional performance. *Regional Studies*, 1–21.

Eurostat (2016). Regional economic accounts.

Fafchamps, M. et al. (2004). Market institutions in sub-Saharan Africa: Theory and evidence. *MIT Press Books 1*.

Farole, T., A. Rodríguez-Pose, and M. Storper (2011). Cohesion policy in the european union: Growth, geography, institutions. *JCMS: Journal of Common Market Studies 49*(5), 1089–1111.

Fay, M. (2012). *Inclusive green growth: The pathway to sustainable development*. World Bank Publications.

Fehr, E. and U. Fischbacher (2004). Third-party punishment and social norms. *Evolution and human behavior 25*(2), 63–87.

Felbermayr, G. J. and F. Toubal (2010). Cultural proximity and trade. *European Economic Review 54*(2), 279–293.

Feldman, M. P. (2014). The character of innovative places: Entrepreneurial strategy, economic development, and prosperity. *Small Business Economics 43*(1), 9–20.

Fernandez, R. (2007). Women, work, and culture. Technical report, National Bureau of Economic Research.

Fisman, R. and T. Khanna (1999). Is trust a historical residue? Information flows and trust levels. *Journal of Economic Behavior & Organization 38*(1), 79–92.

Flinn, M. V., D. C. Geary, and C. V. Ward (2005). Ecological dominance, social competition, and coalitionary arms races: Why humans evolved extraordinary intelligence. *Evolution and Human Behavior 26*(1), 10–46.

Florida, R. (2005). The world is spiky globalization has changed the economic playing field, but hasn't leveled it. *Atlantic Monthly 296*(3), 48.

Fordham, B. O. and V. Asal (2007). Billiard balls or snowflakes? Major power prestige and the international diffusion of institutions and practices. *International Studies Quarterly 51*(1), 31–52.

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *oxford review of Economic Policy 21*(3), 416–438.

Frankel, J. A. and D. Romer (1999). Does trade cause growth? *American economic review*, 379–399.

Freeman, C. (1987). *Technology Policy and Economic Performance: Lessons from Japan*. London, United Kingdom: Pinter Publishers.

Freeman, C. (1995). The national system of inovation in historical perspective. *Cambridge Journal of Economics 19*(1), 5–24.

Frenken, K., F. van Oort, and T. Verburg (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies 41*, 685–697.

Friedman, T. (2005). The world is flat. *New York: Farrar, Straus and Giroux*.

Fritsch, M. and P. Mueller (2007). The persistence of regional new business formation-activity over time–assessing the potential of policy promotion programs. *Journal of Evolutionary Economics 17*(3), 299–315.

Fritsch, M. and P. Mueller (2008). Effects of new business formation on regional development over time: The case of Germany. *Regional Studies 30*(8), 15–29.

Fritsch, M. and F. Noseleit (2013a). Indirect employment effects of new business formation across regions: The role of local market conditions. *92*, 361–382.

Fritsch, M. and F. Noseleit (2013b). Investigating the anatomy of the employment effect of new business formation. *37*, 349– 377.

Fritsch, M. and A. Schroeter (2011). Why does the effect of new business formation differ across regions? *36*, 383–400.

Fritsch, M. and M. Wyrwich (2014). The effect of regional entrepreneurship culture on economic development - evidence for Germany. *Papers in Evolutionary Economic Geography (PEEG) 1411, Utrecht University, Section of Economic Geography*.

Fukuyama, F. (2011). *The origins of political order: from prehuman times to the French Revolution*. Macmillan.

Gallup, J. L., J. D. Sachs, and A. D. Mellinger (1999). Geography and economic development. *International regional science review 22*(2), 179–232.

Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.

Galor, O. and O. Moav (2002). Natural selection and the origin of economic growth. *Quarterly Journal of Economics*, 1133–1191.

Galor, O. and O. Moav (2007). The Neolithic origins of contemporary variation in life expectancy.

Galor, O., M. O., and D. Vollrath (2009). Land inequality and the emergence of human capital promoting institutions and the great divergence. *Review of Economic Studies 76*, 143–179.

Galor, O. and D. Weil (1999). From Malthusian stagnation to modern growth. (89), 150–154.

GEM (2016). Global Entrepreneurship Monitor.

Ghosh, P. and D. Ray (1996). Cooperation in community interaction without information flows. *The Review of Economic Studies 63*(3), 491–519.

Ghosh, T., S. J. Anderson, C. D. Elvidge, and P. C. Sutton (2013). Using nighttime satellite imagery as a proxy measure of human well-being. *Sustainability 5*(12), 4988–5019.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of theoretical biology 206*(2), 169–179.

Gintis, H., S. Bowles, R. Boyd, and E. Fehr (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior 24*(3), 153–172.

Gissinger, R. and N. P. Gleditsch (1999). Globalization and conflict: Welfare, distribution, and political unrest. *journal of world-systems research 5*(2), 327–365.

Glaeser, E., S. Rosenthal, and W. Strange (2010). Urban economics and entrepreneurship. *Journal of Urban Economics 67*(1), 1–14.

Glaeser, E. L., W. R. Kerr, and G. A. Ponzetto (2010). Clusters of entrepreneurship. *Journal of Urban Economics 67*, 150–168.

Gleditsch, K. S. (2002). Expanded trade and GDP data, 1946-92. *Journal of Conflict Resolution 46*, 712–724.

Gleditsch, K. S. and M. D. Ward (2001). A minimum-distance database and applications to international studies. *Journal of Peace Research 38*, 739–758.

Gleditsch, K. S. and M. D. Ward (2006). Diffusion and the international context of democratization. *International Organization 60*(04), 911–933.

Gorard, S. (2003). What is Multilevel modelling for? *51*, 46–63.

Granato, J., R. Inglehart, and D. Leblang (1996). The effect of cultural values on economic development: Theory, hypotheses, and some empirical tests. *American journal of political science*, 607–631.

Greene, W. (2007). *LIMDEP, Version 9.0: Reference Guide*. Econometric Software, Inc.

Greif, A. (1989). Reputation and coalitions in medieval trade: Evidence on the Maghribi traders. *The journal of economic history 49*(04), 857–882.

Greif, A. (2005). Commitment, coercion, and markets: The nature and dynamics of institutions supporting exchange. In *Handbook of new institutional economics*, pp. 727–786. Springer.

Grzymala-Busse, A. (2006). Disaggregating temporal effects. In *Filadelfia: Annual Meeting American Political Science Association*.

Guimaraes, B. and K. D. Sheedy (2012). A model of equilibrium institutions.

Hagen, E. H. and P. Hammerstein (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical population biology 69*(3), 339–348.

Hamilton, D. B. (2008). *Evolutionary Economics: A Study of Change in Economic Thought* (3 ed.). Transaction Publishers, New Brunswick, New Jersey.

Hechavarria, D. M. and P. D. Reynolds (2009). Cultural norms & business start-ups: The impact of national values on opportunity and necessity entrepreneurs. *International Entrepreneurship and Management Journal 5*, 417–437.

Helmke, G. (2007). The origins of institutional crises in Latin America: A unified strategic model and test. *Unpublished manuscript*.

Helsley, R. and W. Strange (2011). Entrepreneurs and cities: Complexity, thickness and balance. *Regional Science and Urban Economics 41*, 550–559.

Henrekson, M. and D. Johansson (2010). Gazelles as job creators: A survey and interpretation of the evidence. *Small Business Economics 35*(2), 227–244.

Hibbs, D. A. J. and O. Olsson (2004). Geography, biogeography, and why some countries are rich and others are poor. *Proceedings of the National Academy of Sciences of the United States*, 3715–3720.

Hill, R. A. and R. I. Dunbar (2003). Social network size in humans. *Human nature 14*(1), 53–72.

Hobbes, T. (1651). Leviathan, or, the matter, forme, & power of a common-wealth ecclesiasticall and civill.

Hodgson, G. M. (2006). What are institutions? *Journal of Economic Issues XL*, 1–25.

Hodgson, G. M. (2007). Evolutionary and institutional economics as the new mainstream? *Evolutionary and Institutional Economics Review 4*(1), 7–25.

Hodgson, G. M. (2009). On the institutional foundations of law: the insufficiency of custom and private ordering. *Journal of Economic Issues 43*(1), 143–166.

Hofer, K. E. (2015). Does female suffrage increase public support for government spending? evidence from swiss ballots. *Economics Working Paper Series from University of St. Gallen, School of Economics and Political Science 1502*.

Holling, C. (2001). Understanding the complexity of economic, ecological and social systemsŠ,. *Ecosystems 4*, 390–405.

Hume, D. (1978, First published 1739). *A Treatise of Human Nature (Second Edition)*. Clarendon Press Oxford.

Huntington, S. P. (1991). Democracy's third wave. *Journal of democracy 2*(2), 12–34.

Huxley, S. J. (2003). *Charles Darwin: The Origin of Species*. Penguin BoPenguin.

Hwang, V. W. and G. Horowitt (2012). *The Rainforest: The secret to building the next Silicon Valley*. Regenwald Los Altos Hills, CA.

Ikenberry, G. J. (2009). *After victory: Institutions, strategic restraint, and the rebuilding of order after major wars*. Princeton University Press.

Inglehart, R. (1997). *Modernization and post modernization: Culture, economic and political change in 43 societies.* Princeton University Press, Princeton.

Isphording, I. E. and S. Otten (2013). The costs of Babylon- Linguistic distance in applied economics. *Review of International Economics 21*(2), 354–369.

Itzigsohn, J. (2000). Immigration and the boundaries of citizenship: The institutions of immigrants' political transnationalism. *International Migration Review*, 1126–1154.

Jaffe, A., M. Trajtenberg, and R. Henderson (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics 108*(3), 577–598.

Jenkins, S. P. (2005). Survival analysis. *Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK*.

Johnson, S., J. McMillan, and C. Woodruff (1999). *Entrepreneurs and the ordering of institutional reform: Poland, Romania, Russia, the Slovak Republic and Ukraine compared*. European Bank for Reconstruction and Development.

King, B. G., M. Cornwall, and E. C. Dahlin (2005). Winning woman suffrage one step at a time: Social movements and the logic of the legislative process. *Social Forces 83*(3), 1211–1234.

King, R. G. and R. Levine (1993). Finance, entrepreneurship and growth. *Journal of Monetary economics 32*(3), 513–542.

Klein, P. G. (1998). New institutional economics. *Available at SSRN 115811*.

Kranton, R. E. (1996). The formation of cooperative relationships. *Journal of Law, Economics, and organization 12*(1), 214–233.

Kubiszewski, I., R. Costanza, C. Franco, P. Lawn, J. Talberth, T. Jackson, and C. Aylmer (2013). Beyond GDP: Measuring and achieving global genuine progress. *Ecological Economics 93*, 57–68.

Laffont, J.-J. and D. Martimort (2009). *The theory of incentives: the principal-agent model*. Princeton university press.

Leckie, G., R. Pillinger, J. Jenkins, and J. Rasbash (2010). School, family and neighbourhood: Which is most important to a childŠs education. pp. 67–70.

Leser, E. and K. Menger (1883). Untersuchungen über die methode der sozialwissenschaften, und der politischen Oekonomie insbesondere.

Levine, D. K. and S. Modica (2013). Anti-Malthus: Conflict and the evolution of societies. *Research in Economics 67*(4), 289–306.

Levitsky, S. and M. V. Murillo (2009). Variation in institutional strength.

Levitt, P. (1998). Social remittances: Migration driven local-level forms of cultural diffusion. *International migration review*, 926–948.

Livi-Bacci, M. (2012). *A concise history of world population*. John Wiley & Sons.

Lizzeri, A. and N. Persico (2004). Why did the elites extend the suffrage? Democracy and the scope of government, with an application to Britain's "Age of Reform". *The Quarterly Journal of Economics*, 707–765.

Lucas, R. (1988). On the mechanics of economic development. *Journal of monetary economics 22*(1), 3–42.

Lundvall, B.-A. (Ed.) (1992). *National Systems of Innovation: towards a theory of innovation and interactive learning*. London, United Kingdom: Pinter Publishers.

Malthus, T. (1798). *An essay on the principle of population*. Printed for J. Johnson, in St. Paul's Church-Yard, London.

Mankiw, N., D. Romer, and D. Weil (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics 107*(2), 407–437.

Marshall, M. G., T. R. Gurr, and K. Jaggers (2014). Polity IV Project: Political regime characteristics and transitions, 1800-2013. Technical report, Center for Systemic Peace.

Martin, R. and P. Sunley (2003). Deconstructing clusters: Chaotic concept or policy panacea? *Journal of economic geography 3*(1), 5–35.

Marx, K. (1857). Foundation of the critique of political economy.

Mayshar, J., O. Moav, and Z. Neeman (2011). Transparency, appropriability and the early state. *Working Papers 002-11, International School of Economics at TSU, Tbilisi, Republic of Georgia.*.

McCammon, H. J. and K. E. Campbell (2001). Winning the vote in the West: The political successes of the women's suffrage movements, 1866-1919. *Gender & Society 15*, 55–82.

McFarland, D. (2015). *The Biological Bases of Economic Behaviour: A Concise Introduction*. Palgrave Macmillan.

Melitz, J. and F. Toubal (2014). Native language, spoken language, translation and trade. *Journal of International Economics 93*(2), 351–363.

Menger, C. (1976 (1871)). *Principles of Economics. Translation of Grundstitze der Volkswirtschaftslehre.* New York University Press.

Michelacci, C. (2003). Low returns in R&D due to lack of entrepreneurial skills. *The Econimic Journal 113*(484), 207–225.

Milkman, R. (1987). *Gender at work: The dynamics of job segregation by sex during World War II*. University of Illinois Press.

Nash, J. F. (1950). Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences,*.

Nelson, R. R. (1993). *National innovation systems: a comparative analysis*. Oxford university press.

Nickell, S. (1996). Competition and corporate performance. *Journal of political economy*, 724–746.

Nickell, S., D. Nicolitsas, and N. Dryden (1997). What makes firms perform well? *European Economic Review 41*(3), 783–796.

Norgaard, R. B. (1994). The coevolution of economic and environmental systems and the emergence of unsustainability. *Evolutionary concepts in contemporary economics*, 213–225.

North, D. (1991). Institutions. *The Journal of Economic Perspectives* (5), 97–112.

North, D. and R. Thomas (1973). *The Rise of the Western World.* Cambridge: Cambrige University Press.

North, D. C. (1989). Institutions and economic growth: An historical introduction. *World development 17*(9), 1319–1332.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge university press.

Nunn, N. (2008). The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics 123*, 139–176.

Olsson, O. and D. J. Hibbs (2005). Biogeography and long-run economic development. *European Economic Review 49*, 909–938.

Osberg, L. and A. Sharpe (2002). An index of economic well–being for selected OECD countries. *Review of Income and Wealth 48*(3), 291–316.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Ostrom, E. (2014). Collective action and the evolution of social norms. *Journal of Natural Resources Policy Research 6*(4), 235–252.

Owen, J. M. (2002). The foreign imposition of domestic institutions. *International Organization 56*(02), 375–409.

Paterson, L. and H. Goldstein (1992). New statistical methods for analyzing social structures: An introduction to multilevel models. *17*, 387–393.

Patterson, T. C. (2005). Craft specialization, the reorganization of production relations and state formation. *Journal of Social Archaeology 5*(3), 307–337.

Peic, G. and D. Reiter (2011). Foreign-imposed regime change, state power and civil war onset, 1920–2004. *British Journal of Political Science 41*(03), 453–475.

Peregrine, P. N. (2003). *Atlas of cultural evolution*.

Perkins, R. and E. Neumayer (2009). Transnational spatial dependencies in the geography of non-resident patent filings. *Journal of Economic Geography*, 1–24.

Piketty, T. (2014). *Capital in the Twent-First Century*. Harvard University Press.

Porter, M. E. (2000). Location, competition, and economic development: Local clusters in a global economy. *Economic development quarterly 14*(1), 15–34.

Pritchett, L. and M. Woolcock (2004). Solutions when the solution is the problem: Arraying the disarray in development. *World Development 32*(2), 191–212.

Puga, D. (2002). European regional policies in light of recent location theories. *Journal of economic geography 2*(4), 373–406.

Putnam, R. D. (1993). The prosperous community. *The american prospect 4*(13), 35–42.

Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century the 2006 johan skytte prize lecture. *Scandinavian political studies 30*(2), 137–174.

Putterman, L. and D. Weil (2010). Post-1500 population flows and the long-run determinants of economic growth and inequality. (125), 1627–1682.

Ramirez, F. O., Y. Soysal, and S. Shanahan (1997). The changing logic of political citizenship: Cross-national acquisition of women's suffrage rights, 1890 to 1990. *American Sociological Review*, 735–745.

Renfrew, C., I. Todd, and R. Tringham (1974). Beyond a subsistence economy: The evolution of social organization in prehistoric Europe. *Bulletin of the American Schools of Oriental Research. Supplementary Studies* (20), 69–95.

Reynolds, P., N. Bosma, E. Autio, S. Hunt, N. De Bono, I. Servais, P. Lopez-Garcia, and N. Chin (2005). Global entrepreneurship monitor: Data collection design and implementation 1998–2003. *Small Business Economics 24*(3), 205–231.

Reynolds, P., D. Storey, and P. Westhead (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies 28*(4), 443–456.

Ricardo, D., E. C. K. Gonner, and Q. Li (1819). The principles of political economy and taxation.

Richerson, P. and J. Henrich (2009). Tribal social instincts and the cultural evolution of institutions to solve collective action problems. *Context and the Evolution of Mechanisms for Solving Collective Action Problems Paper*.

Rocha, H. O. (2004). Entrepreneurship and development: The role of clusters. *Small business economics 23*(5), 363–400.

Rodríguez-Pose, A. (2013). Do institutions matter for regional development? *Regional Studies*, 1–14.

Rodríguez-Pose, A. and R. Crescenzi (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. *Regional Studies 42*, 51–67.

Rodríguez-Pose, A. (1999). Convergence or divergence? Types of regional responses to socio-economic change in Western Europe. *Tijdschrift voor Economische en Sociale Geografie 90*(4), 365–378.

Rodrik, D. (2004). Getting institutions right. Technical report, CESifo DICE Report Number 2/2004. Ifo Institute for Economic Research, Munich.

Rodrik, D., A. Subramanian, and F. Trebbi (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. *Journal of economic growth 9*(2), 131–165.

Romer, P. (1986). Increasing returns and long-run growth. *The Journal of Political Economy 6*, 1002–1037.

Rowthorn, R. and P. Seabright (2010). Property rights, warfare and the Neolithic Transition. *TSE Working Paper Series.*.

Sachs, J. (2005). *The end of poverty: economic possibilities for our time*. London: Penguin.

Sachs, J. and P. Malaney (2002). The economic and social burden of malaria. *Nature 415*(6872), 680–685.

Sahlins, M. D. (1974). *Stone age economics*. Transaction Publishers.

Sala-i Martin, X. (1994). Cross-sectional regressions and the empirics of economic growth. *European Economic Review 38*, 739–747.

Sala-i Martin, X. (1996). Regional cohesion: Evidence and theories of regional growth and convergence. *European Economic Review 40*, 1325–1352.

Salehyan, I. and K. S. Gleditsch (2006). Refugees and the spread of civil war. *International Organization 60*, 335–366.

Santarelli, E. and M. Vivarelli (2007). Entrepreneurship and the process of firms' entry, survival and growth. *Industrial and Corporate Change 16*(3), 455–488.

Sarkees, M. R. and F. W. Wayman (2010). *Resort to war: a data guide to inter-state, extra-state, intra-state, and non-state wars, 1816-2007*. Correlates of War Project. V4.0.

Schumpeter, J. (1942). *Capitalism, Socialism and Democracy, (1942)*, Volume 20. Routledge, London.

Schumpeter, J. A. (1912). *The Theory of Economic Development: An Inquiry Into Profits, Capital, Credit, Interest, and the Business Cycle*.

Searle, J. R. (1995). *The construction of social reality*. Simon and Schuster.

Sengupta, N. (1991). *Managing common property: irrigation in India and the Philippines.* Sage Publications India Pvt Ltd.

Shennan, S. (2000). Population, culture history, and the dynamics of culture change1. *Current Anthropology 41*(5), 811–835.

Shubik, M. (1999). *The Theory of Money and Financial Institutions. Volume I*. Massachussets Instititute of Technology.

Silverberg, G., B. Verspagen, et al. (1995). *Evolutionary theorizing on economic growth*. Internat. Inst. for Applied Systems Analysis.

Simmons, B. A., F. Dobbin, and G. Garrett (2006). Introduction: The international diffusion of liberalism. *International Organization 60*(04), 781–810.

Simmons, B. A., F. Dobbin, and G. Garrett (2007). The global diffusion of public policies: Social construction, coercion, competition or learning? *Annual review of sociology 33*, 449–472.

Smith, A. (1776). *The wealth of nations*.

Smith, A. D. (2013). *Nations and nationalism in a global era*. Routledge.

Sng, T.-H. (2014). Size and dynastic decline: The principal-agent problem in late imperial China, 1700–1850. *Explorations in Economic History 54*, 107–127.

Snijders, T. A. and R. J. Bosker (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London etc.: Sage Publishers.

Sobel, R. S. (2008). Testing Baumol: Institutional quality and the productivity of entrepreneurship. *Journal of Business Venturing 23*(6), 641–655.

Sokoloff, K. L. and S. L. Engerman (2002). Factor endowments, inequality, and paths of development among New World economies. *Economica 3*, 41–88.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, 65–94.

Spigel, B. (2015). The relational organization of entrepreneurial ecosystems. *Entrepreneurship Theory and Practice*, DOI: 10.1111/etap.12167.

Spilimbergo, A. (2009). Foreign students and democracy. *American Economic Review 99*(1), 528–43.

Stam, E. (2009). Entrepreneurship. In R. Kitchin and N. Thrift (Eds.), *The International Encyclopedia of Human Geography*.

Stam, E. (2015). Entrepreneurial ecosystems and regional policy: A sympathetic critique. *European Planning Studies 23*(9), 1759–1769.

Stam, E., N. Bosma, A. van Witteloostuijn, J. de Jong, S. Bogaert, N. Edwards, and F. Jaspers (2012). Ambitious entrepreneurship. A review of the academic literature and new directions for public policy. Number 41 in AWT rapport.

Steckel, R. H. and J. C. Rose (2002). *The backbone of history: health and nutrition in the Western Hemisphere*, Volume 2. Cambridge University Press.

Stenholm, P., Z. Acs, and R. Wuebker (2011). Exploring country-level institutional arrangements on the rate and type of entrepreneurial activity. *Journal of Business Venturing*.

Stephens, H. and M. Partridge (2011). Do entrepreneurs enhance economic growth in lagging regions? *Growth and Change 42*(4), 431–465.

Stevens, J. R. and M. D. Hauser (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends in cognitive sciences 8*(2), 60–65.

Stifel, D. and B. Minten (2008). Isolation and agricultural productivity. *Agricultural Economics 39*(1), 1–15.

Strang, D. and J. W. Meyer (1993). Institutional conditions for diffusion. *Theory and society 22*(4), 487–511.

Swallow, B. M. and D. W. Bromley (1995). Institutions, governance and incentives in common property regimes for African rangelands. *Environmental and Resource Economics 6*(2), 99–118.

Szerb, L., Z. J. Acs, R. Ortega-Argilés, and E. Komlosi (2015). The entrepreneurial ecosystem: The regional entrepreneurship and development index. *Mimeo*. Available at SSRN 2642514.

Tabellini, G. (2010). Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association 8*(4), 677–716.

Testart, A., R. G. Forbis, B. Hayden, T. Ingold, S. M. Perlman, D. L. Pokotylo, P. Rowley-Conwy, and D. E. Stuart (1982). The significance of food storage among hunter-gatherers: Residence patterns, population densities, and social inequalities [and comments and reply]. *Current anthropology*, 523–537.

Thissen, M., N. Limtanakool, and H. Hilbers (2011). Road pricing and agglomeration economies: A new methodology to estimate indirect effects applied to the netherlands. *The Annals of Regional Science 47*(3), 543–567.

Tilly, C. (1985). War making and state making as organized crime. In C. Besteman (Ed.), *Violence, a Reader*.

Tinbergen, J. (1962). *Shaping the World Economy*. New York: The Twentieth Century Fund.

Tridimas, G. (2015). War, disenfranchisement and the fall of the ancient Athenian democracy. *European Journal of Political Economy 38*, 102–117.

Turchin, P., T. E. Currie, E. A. Turner, and S. Gavrilets (2013). War, space, and the evolution of Old World complex societies. *Proceedings of the National Academy of Sciences 110*(41), 16384–16389.

Unger, J. M., A. Rauch, M. Frese, and N. Rosenbusch (2011). Human capital and entrepreneurial success: A meta-analytical review. *Journal of Business Venturing 26*(3), 341–358.

Van Bavel, B. (2010). *Manors and Markets: Economy and Society in the Low Countries 500-1600*. Oxford University Press.

Van Den Bergh, J. C. and S. Stagl (2003). Coevolution of economic behaviour and institutions: Towards a theory of institutional change. *Journal of Evolutionary Economics 13*(3), 289–317.

Van Marrewijk, C. Mimeo, 2016.

Van Praag, C. M. and P. H. Versloot (2007). What is the value of entrepreneurship? A review of recent research. *Small Business Economics Journal 29*(4), 351–382.

Weisdorf, J. (2009). Why did the first farmers toil? Human metabolism and the origins of agriculture. *European Review of Economic History 13*(2), 157–172.

Wennekers, S., A. van Stel, R. Thurik, and P. Reynolds (2005). Nascent entrepreneurship and the level of economic development. *Small business economics 24*, 293–309.

Williamson, O. E. (2000). The new institutional economics: Taking stock, looking ahead. *Journal of economic literature*, 595–613.

Wilson, J. A. (1980). Adaptation to uncertainty and small numbers exchange: The New England fresh fish market. *The Bell Journal of Economics*, 491–504.

Witt, U. (1989). The evolution of economic institutions as a propagation process. *Public Choice 62*(2), 155–172.

Zhao, R. and L. Cao (2010). Social change and anomie: A cross-national study. *Social Forces 88*(3), 1209–1229.

Samenvatting

Introductie

Een fundamenteel vraagstuk in de macro-economie is hoe de grote inkomensverschillen tussen landen over de wereld tot stand zijn gekomen. Binnen de economische wetenschap heeft een verschuiving plaatsgevonden van directe verklaringen naar fundamentele verklaringen. De eerste groeimodellen leggen de nadruk op de accumulatie van fysiek kapitaal (Solow, 1956) en later op de accumulatie van menselijk kapitaal (Mankiw et al., 1992; Becker et al., 1994; Frankel and Romer, 1999) en innovatie (Aghion and Howitt, 1992). Deze verklaringen gaan echter niet in op de vraag waarom in bepaalde landen meer innovatie plaatsvindt, of meer geïnvesteerd wordt in menselijk of fysiek kapitaal ten opzichte van andere landen. Hiervoor moet worden teruggegrepen op de fundamentele verklaringen, die door Acemoglu (2009) zijn samengevat als de hypothesen van cultuur, instituties, geluk en geografie. In deze dissertatie focus ik op meerdere facetten van de institutionele hypothese.

Instituties zijn de door mensen ontworpen restricties die structuur aanbrengen in onze politieke, economische en sociale interacties. Deze definitie werd geïntroduceerd door North (1991) en is breed aanvaard in de economische wetenschap. Binnen deze definitie is er een breed scala aan concepten dat onder instituties kan worden geschaard, zoals wetten, overheidsorganisaties, parlementen, monarchieën, een gedeelde taal, taboes en gewoonten of religieuze gebruiken.

De economische wetenschap heeft een consensus bereikt dat instituties zowel remmend als bevorderend kunnen werken voor economische groei, en dat verschillen in de kwaliteit van instituties grotendeels verklarend zijn voor de inkomensverschillen tussen landen over de wereld. Het is ook bekend dat de verschillen in instituties te maken hebben met biologische en geografische condities zoals die in de prehistorie vastlagen, zoals klimaat-

verschillen en de mate waarin de grond geschikt is voor het bedrijven van landbouw. Deze theorie is voorgesteld door Jared Diamond (1997) en later wetenschappelijk getest door onder andere Olsson and Hibbs (2005), Galor et al. (2009) en Ashraf and Michalopoulos (2015).

Minder is bekend over de exacte processen die hebben geleid tot het ontstaan van deze (verschillen in) instituties. In deze dissertatie heb ik getracht deze 'zwarte doos' in drie delen te openen. Ten eerste heb ik gekeken naar het ontstaan van instituties in een geïsoleerde maatschappij als gevolg van de overgang van een nomadische jager-verzamelaarlevensstijl naar een sedentaire agrarische levensstijl en de veranderingen in bevolkingsdichtheid en technologie die daarmee gepaard gaan. Ten tweede heb ik geanalyseerd hoe een institutie zich kan verspreiden naarmate maatschappijen meer met elkaar in contact komen. Tenslotte heb ik gekeken of een kern-periferie patroon in de kwaliteit van instituties, dat uit de theorie voortkomt, in de data zichtbaar gemaakt kan worden door het toetsen van een heterogeen effect van ondernemerschapsactiviteit op economische groei.

Modellen

Hoofdstuk 2 dient als een verdiepende inleiding tot de dissertatie. Ik geef een overzicht van de literatuur aangaande instituties en institutionele verandering. Hier wordt een onderscheid gemaakt tussen formele en informele instituties (North, 1990) organische instituties en instituties die actief gecreëerd zijn (Menger, 1871; Greif, 2005), zichzelf versterkende instituties en instituties die actieve ondersteuning vereisen (Hodgson, 2006), en regelingen of wetten versus de institutionele omgeving of setting (Davis et al., 1971).

De contributie van dit hoofdstuk is het aantonen dat de hogere bevolkingsdichtheid die bereikt werd na de Neolithische Revolutie —de overgang naar landbouw die op verschillende plekken in de wereld vanaf 10.000 voor Christus in gang trad —een mechanisme kan zijn dat leidt tot een hogere vraag naar complexe, ontwikkelde instituties. Jager-verzamelaarstammen zijn door de vereisten van het nomadisch bestaan en de beperkte opslag van voedsel per definitie beperkt in hun groei. Een concreet voorbeeld hiervan is dat een vrouw maar één jong kind tegelijkertijd kan dragen wanneer de stam zich moet verplaatsen. Door de beperkte omvang zijn alle leden van de stam nauw met elkaar verbonden via familiebanden en kunnen transacties veelal op de basis van vertrouwen plaatsvinden.

Zodra de maatschappij sedentair wordt kan een hogere bevolkingsdichtheid bereikt worden, wat impliceert dat er meer onbekenden in de groep zijn en de kans op valsspelen of het ontwijken van verplichtingen toeneemt. De transactiekosten zullen dus toenemen en een vraag naar instituties om transacties tussen vreemdelingen te reguleren ontstaat. Instituties zijn echter een publiek goed, wat inhoudt dat de baten aan niemand ontzegd kunnen worden terwijl ieder individu motieven heeft om niet bij te dragen heeft aan de kosten. Dit betekent dat de vraag naar deze instituties niet altijd leidt tot hun totstandkoming en een duurzame investering van de maatschappij om het publieke goed te behouden. Om te analyseren in welke type maatschappijen en onder welke omstandigheden dit wel bereikt kan worden, voldoet een model met representatieve agenten niet en moet dit gecombineerd worden met inzichten uit de gedragseconomie.

Een maatschappij moet vele hindernissen overwinnen in de ontwikkeling van groeibevorderende instituties. Om dit proces te beschrijven zijn meerzijdige modellen vereist. De samenhang tussen die hindernissen en verschillende modellen die ze beschrijven laat ik zien in een stroomschema dat de verschillende fasen van institutionele ontwikkeling weergeeft. De inerte, en zichzelf versterkende aard van instituties, impliceert daarnaast dat een conventie of institutie die gedurende aanhoudende tijd functioneert in een maatschappij, erg moeilijk is te veranderen of te vervangen. Dit betekent dat instituties stand kunnen houden zelfs al zijn deze door veranderde omstandigheden niet meer optimaal.

Een onvermijdelijk gevolg van de groei in omvang van maatschappijen is meer contact tussen deze maatschappijen en de competitie om schaarse goederen. Dit contact kan vriendelijk zijn en door schaalvoordelen de baten van alle partijen vergroten, maar de geschiedenis leert ons dat oorlog en koloniale betrekkingen ook belangrijke vormen van contact waren zodra natiestaten zich geconstitueerd hadden. Hoofdstuk 3 gaat dieper in op de vraag hoe de verschillende vormen van contact (oorlog, handel, diplomatie en koloniale betrekkingen) kunnen leiden tot de overname van een institutie wanneer één van de partijen deze institutie nog niet geïmplementeerd heeft.

In het bijzonder kijk ik welke factoren de introductie van universeel kiesrecht voor vrouwen voorspellen, in de tijdsperiode 1893 tot 2008. In 1893 voerde Nieuw Zeeland dit recht als eerste in en vandaag de dag is er nog slechts een handvol landen waar vrouwen geen volledig kiesrecht hebben. Het blijkt dat internationaal contact en de invloed van buurlanden belangrijke voorspellers zijn van de invoering van vrouwenkiesrecht; ook als externe schokken en voorspellende factoren binnen het land, zoals de

positie van vrouwen in het land, worden meegenomen. Daarnaast is er een verschil tussen vroege en late adopteerders, waarbij diffusie vaker plaatsvindt naarmate de institutie meer wijdverbreid wordt.

Een hogere kwaliteit van instituties leidt tot meer economische groei, maar meer economische groei gaat vaak ook gepaard met de ontwikkeling van betere instituties. Door dit endogeniteitsprobleem is het causale effect van instituties op economische groei empirisch lastig vast te stellen. Daarnaast zijn de instituties die worden verondersteld economische groei te bevorderen lastig meetbaar en kennen alle voorgedragen instrumenten ook onzuiverheden. Hoofdstuk 4 hanteert daarom een indirecte aanpak via het kanaal van ondernemerschapsactiviteit. Ondernemerschapsactiviteit leidt theoretisch tot meer innovatie en daarom tot hogere economische groei (Michelacci, 2003; Acs and Sanders, 2012).

In de praktijk wordt dit effect echter niet altijd gevonden. Verklaringen omvatten dat het effect van ondernemerschap op groei zich pas na meerdere jaren vertoont (Fritsch and Mueller, 2007; Fritsch and Schroeter, 2011), dat alleen bepaalde soorten ondernemerschap tot hoge groei leiden (Henrekson and Johansson, 2010; Stam et al., 2012) of dat er een modererend effect is van de kwaliteit van instituties op de relatie tussen ondernemersactiviteit en groei (Baumol, 1990). Deze laatste hypothese test ik in een set van 107 Europese regio's op NUTS-2 aggregatieniveau voor de periode 2001 tot 2014. Een modererend effect van (informele) instituties zou betekenen dat de relatie tussen ondernemerschap en groei niet voor alle regio's homogeen is. Er kunnen dan verschillende klassen onderscheiden worden van regio's die verschillen in institutionele kwaliteit, specifiek in hun ondernemerschapscultuur.

Deze heterogene relatie wordt echter niet gevonden in deze analyse, wat niet betekent dat de hypothese van een modererend effect van instituties verworpen moet worden. De periode die bestudeerd is valt samen met de financiële crisis en de problematiek in de Eurozone, wat betekent dat de groeivoet in Europese regio's sterk verlaagd was terwijl ondernemerschapsactiviteit minder direct beïnvloed werd. Daarnaast is het waarschijnlijk dat de verschillen in institutionele regimes zich niet op het politieke aggregatieniveau van NUTS-2 bevinden, maar op lagere aggregatieniveaus. Deze analyse zou dus gerepliceerd kunnen worden in andere steekproeven om meer bewijs te krijgen voor de latente invloed van instituties op de relatie tussen ondernemerschap en economische groei.

Conclusie

Mijn bevindingen impliceren dat de invloed van instituties in de economische wetenschap niet onverminderd is. Er bestaat geen twijfel dat verschillen in ontwikkeling, en vooral het achterblijven van ontwikkelingslanden in termen van welvaart, gelinkt zijn aan verschillen in institutionele kwaliteit. Toch zijn er nog weinig studies die meetbaar maken hoe deze instituties ontstaan, hoe ze zich verspreiden, en welke instituties groeibevorderend dan wel groeiremmend zijn.

Meetproblemen en endogeniteitsproblemen compliceren deze analyses, daarom moet de zoektocht naar indirecte methoden en manieren om verschillen in instituties te instrumenteren, voortgezet worden. Deze studie draagt bij aan het begrip van het ontstaan, de diffusie en de mogelijke latente effecten van instituties op economische groei, maar om directe causale effecten aan te tonen is de combinatie van case-studies en natuurlijke experimenten essentieel.

Instituties zijn van nature inert, en er is veelal een periode van onrust en conflict nodig om drastische veranderingen in de institutionele structuur te bewerkstelligen. Wanneer de institutionele structuur niet meer optimaal is gezien de veranderde omstandigheden, of wanneer de formele instituties niet langer ingebed zijn in de informele instituties van een bevolking, leidt dit tot spanningen. Instituties bieden stabiliteit, maar deze stabiliteit kan ook averechts werken en een rem zijn op innovatie. Er is dus niet één institutionele setting die geschikt zal zijn voor alle landen in alle stadia van ontwikkeling en het zou foutief zijn om de institutionele structuur van ontwikkelde landen te kopiëren naar ontwikkelingslanden. In plaats daarvan is er meer onderzoek nodig naar oorzaken van het voortbestaan van disfunctionele instituties en naar factoren die wel kunnen leiden tot institutionele verandering en het optimale niveau van flexibiliteit.

Curriculum Vitae

Krista Bruns (1988) was born in Heemskerk, Netherlands, where she completed her pre-university education at the Kennemer College in 2006. From 2006 to 2009 she did a Bachelor Programme in Liberal Arts and Sciences at University College Utrecht, where she graduated cum laude. After this she followed a pre-master year with courses in Econometrics at Erasmus University Rotterdam, followed by the Research Master Programme at the Utrecht University School of Economics. In September 2012, she started her PhD at the Utrecht University School of Economics (USE), where she completed this dissertation. As of September 2016, she is working as a policy advisor at the Ministry of Finance. She is currently mainly responsible for analyzing budgetary effects of changes in corporate income taxes.