An attitude diffusion model of the international clustering of political regimes

Johan A. Elkink

PhD thesis

Department of Political Science
Trinity College Dublin

June 2009

Declaration

This thesis has not been submitted as an exercise for a degree at any other university. Except where stated, the work described therein was carried out by me alone.

I give permission for the Library to lend or copy this thesis upon request.

Signed:

21-5-2009

Summary

While pro-active policies to promote democracy abroad by democratic regimes lead to the international clustering of democratic regimes that we observe, the diffusion of attitudes through communication between individual citizens counteracts this trend. The promotion of worldwide democracy should thus focus on the direct influence of individual attitudes towards democracy, on democratic propaganda, rather than on opening channels for communication through increased linkages between countries per se. This thesis provides a theoretical model of the international diffusion of democracy, based on the diffusion of public opinion, that leads to these conclusions.

On the basis of the relatively new and growing literature on the presence of spatial clustering and temporal waves of the spread of democracy in the world, this thesis sets out to make an inventory of the various theoretical explanations that are available to account for these phenomena and to investigate the extent to which a model based on the diffusion of individual attitudes, in combination with a cascading model of revolution, can be a potential explanation of these global and longterm patterns. Almost all existing explanations are entirely based on elite-level explanations of democratization. There is nevertheless no clear *a priori* reason to assume that the geographic clustering we observe cannot have been caused by mass-level attitudes and behavior. The argument is made that even if most transitions to democracy are in the end crucially dependent on decisions and actions by members of the elite, the role of public opinion cannot be ignored. Often elite members make decisions exactly because they are concerned with their popularity among the general population and at other times members of the elite actually lose their political position due to popular elections - increasingly common given the prevalence of 'electoral dictatorships', whereby the power-holders attempt to demonstrate their power to competitors through

demonstrations of electoral fraud. Given that we can assume that mass-level attitudes do indeed matter in processes of democratization, in addition to elite attitudes and behavior, could it then also be that the geographical and temporal clustering as observed in the data over the past two hundred years is indeed an effect of the diffusion of individual attitudes, of public opinion?

The methodological approach in this thesis is explore the *possibility* rather than the *validity* of such an explanation of these global patterns. Instead of positing a small number of key hypotheses and testing these using empirical data, the approach is to study the theoretical possibility of such an explanation through a relatively parsimonious computer simulation. Much akin to the verification of the internal consistency of theoretical models through the application of game-theory, this approach allows us to verify whether, given a large number of agents and given the possibility of non-linear interactive effects, the theoretical model based on mass opinions and cascading revolutions can indeed lead to similar clustering patterns as we observe in the real world.

To clarify the discussion of these patterns, a classification of regime clustering is presented and taken as a guide in the subsequent analysis of the simulation results, as well as in the classification of the various existing theoretical explanations. A distinction is made between spatial clustering, whereby at any given moment in time, democracies are likely to be found in geographically contiguous regions; temporal clustering, whereby democratic transitions and their reversals are likely to occur within short time periods, a kind of *Zeitgeist* that leads to an increase in the number of transitions within a few years; and spatio-temporal clustering, whereby these transitions not only take place one after the other, but also in a geographically contiguous region.

It turns out that the attitudinal model developed in this thesis performs well as a potential explanation for the geographic clustering we observe, but that it is less useful for providing an understanding of the presence of temporal waves. Two mechanisms are implemented in the model in terms of attitudinal diffusion: 1) individuals alter their attitude towards democracy through the individual communication with other citizens, of either the same or a neighboring country; 2) democratic countries have active policies to affect the pro-democratic attitudes of citizens in neighboring countries. From the simulations, under many different parameter configurations, we can learn that the latter is far more influential than the former. The individual diffusion of attitudes is an interactive effect in the sense that it can *slow down* the clustering effect if it is strong, while the actual clustering is generated by the democracy promotion mechanism.

Contents

List of figures

List of tables

Acknowledgments

Writing a dissertation over a period of six years, building on earlier work for my Masters' thesis the two years before that, means that there are so many people that deserve acknowledgments that it is impossible not to forget crucial ones here. Especially anyone who knows the porousness of my brain can imagine the difficulty I have in writing this section. Let me nevertheless make an attempt to show my gratitude to a number of people.

First and foremost I am very grateful for the excellent supervision I received first from Michael Laver and later from Ken Benoit in writing my dissertation. Although it might appear at first sight as an extra hurdle, the transition from one supervisor to another halfway through my dissertation has provided me with an entire extra dimension in supervisory comments, which, I think, has only helped in the formation of this thesis. Both have provided me not only with excellent comments, but also with stimulating examples of how to be successful as an academic.

Of course not only immediate supervisors provide such helpful support in the department where you write your PhD. The large amount of feedback from the very useful Friday seminars has also helped me write my dissertation and in this regard I would like to thank in particular Gail McElroy, Michael Marsh, Robert Thompson, Will Lowe, Eddie Hyland, Alex Baturo, and Slava Mikhailov. Cees van der Eijk also provided useful advice on a visit to the department, pointing me to the work of Elisabeth Noelle-Neumann which has been fundamental to the theoretical foundations of the thesis.

Of course, conference attendance and the accompanying feedback has been important in the formation of my ideas that are now in this thesis. I would like to thank in particular James McCann, Scott Page, Orit Rapaport-Gertzek and Gerald Schneider for their enormously useful comments. All participants of ECPR's Joint Sessions of Workshops both on "Geography, conflict, and cooperation" (Edinburgh, March 2003) and on "The diffusion of policies and institutions" (Nicosia, April 2006), in particular Kristian Gleditsch, Fabrizio Gilardi, Thomas Plümper and Rob Franzese, deserve my grateful acknowledgments. The hike in the Scottish mountains after the workshop in Edinburgh has been crucial, for that is when Kristian Gleditsch asked me whether I was interested in agent-based models of democratic diffusion. Apparently I am. Through summer schools and accompanying presentations I also received substantial feedback and support, especially from James Fowler and Kristin Kanthak.

Although preceding the writing of the actual PhD, the support, advice, and encouragement of the members of the Department of Political Science at Leiden University cannot go unmentioned. It is thanks to Huib Pellikaan that I ended up at Trinity College in the first place, which has been an eye-opening experience. If he had not recommended it, I would never have thought of it. Alas, I never got to the coalition theories he thought I should have been studying. The Masters' thesis supervision by Hans Oversloot and Renske Doorenspleet has helped me lay the foundations on which later this thesis was build. Working as a research assistant in projects by Joop van Holsteyn, Ruud Koole, Rudy Andeweg and Jacques Thomassen has generated the enthusiasm for academic life that has been crucial in deciding to do a PhD. The many things I learned from those people, and during those projects, has been my guide in all academic work, including my teaching, ever since. That is not to say that I never had doubts about my academic career - quite to the contrary! - and I am enormously grateful for the mental support in particular from Joop van Holsteyn, Huib Pellikaan and Herman van Gunsteren. Without them I would not have made it this far.

Of course, not only the many discussions of my work in a formal or institutional context have been important. Several friends have, over the years, been a listening ear or have made encouraging comments with regards my dissertation research. In particular Alex Baturo, Alex Perez, Michel Schilperoord, Tom van der Meer, Sarah de Lange, Shane Mac Giollabhui, Sico van der Meer and Maarten Rozing should be mentioned. Although I often avoided talking about my dissertation, the many conversations we have had have been hugely important to me and to the formation of my dissertation.

It goes without saying that besides direct support for my academic work, personal support from friends and family has been perhaps even more important. Here it will be impossible to acknowledge every single person who has been important to me over the past six years. But I am certainly eternally grateful to my brother Teunis, to Peggy, Ruud, and to Sico, Maarten, Yusuf, Alejandro, Aleesh (two times an Alex in my phone was too confusing!), Zbyszek, Sarah, Tom, Martha, Sander, Slava, Shane, David, Rory, Martin, Eduardo, Tim, Kostya, Mo, Sergey, Eric, Oleg, Victor, Wuu Kuang, Kenneth, Michael, Doc, Anthony, and all members of the Out and About hiking group, Menelaos, Maken, William, and all members of the DLGFF film festival committee, Mike, Teunis H., Anne, Paula, Kees, Rika, my grandparents, Umut, Adina, and all my colleagues and students at UCD, ... and it scares me how many people I must have forgotten to mention that I care so much about!

I should not only thank people. I could not have done my dissertation without the financial support from the Irish Government, first through the Cultural Treaty with the Netherlands and then for years on the basis of my scholarship with the Irish Research Council for the Humanities and Social Sciences. The Institute for International Integration Studies at Trinity provided me with useful desk space and with a brilliant opportunity to meet people in a variety of different disciplines and to attend hugely interesting seminars. My stay for one year at the Institute for Quantitative Social Science and its Harvard-MIT Data Center at Harvard University has been a formative period in my academic career and during that time I wrote a large part of this thesis and learned an enormous amount about quantitative methods in the social sciences.

That writing, however, took in actual fact place in Starbucks. Without their coffee, I don't know how far this thesis would have progressed. My new favorite, Coffee Society in Camden Street, also deserves a mention. But alas, I finally got around to buying my own coffee maker.

This dissertation is dedicated to Jan and Maatje. Not only are they the best parents one could ever wish for and not only have they always been incredibly supportive for whatever I chose to do, they also have the rare quality among parents of being among my closest friends for all these years.

Introduction

The idea that democracy is contagious, that democracy diffuses across the world map, is now well established among policy makers and political scientists alike. The few theoretical explanations of this phenomenon focus exclusively on the political elites. In this thesis an alternative theoretical model is presented in an attempt to explain these patterns on the basis of individual, mass-level attitudes and behavior. The main question to be answered in this thesis is:

> *"How can the observed patterns of regime clustering and global trends towards democratization be explained on the basis of the diffusion of ideas among people and what obstacles to this diffusion exist that explain its continuing character?"*

Recent developments in Ukrainian politics form a superb example of the type of political behavior that is being studied in this thesis. A Western-oriented, liberal candidate Yushchenko won the elections, while the powerful president in office, Kuchma, supported Yanukovich, who is more oriented towards the East. The president had the control over the state apparatus and was able to commit considerable electoral fraud to support his preferred candidate and thus showed behavior similar to that of other presidents in the region, who try to combine a democratic constitution with an authoritarian style of government. The reason Yushchenko managed to win was the considerable and unexpected pressure from protesters in Kiev. Unexpected, because the Ukrainian population was considered to be rather apathetic to the political situation in their country. Stability and economic progress seemed to be higher on their agendas than political reforms. What then explains the sudden large number of protesters and the resulting failure of the regime to preserve its authoritarian status and of the president to elect his own successor?

A factor that cannot be ignored in this election has been the role of international pressure. The United States and the European Union, among others, openly supported the democratic movement within the Ukraine and pressed the Ukrainian president to give in to those pressures and organize fair elections. The Russian president, Putin, openly supported Yanukovich, whose position showed considerable similarities to that of Putin after Yeltsin stepped down, and who supported a more pro-Russian foreign policy. Moreover, several Eastern European organizations for democratization were active in the region and actively tried to persuade the Ukrainian population to fight for their democratic rights. The successes of democratization in other Eastern European countries, and especially in Serbia, thus had a direct effect on the success of the election process in Ukraine. Examples of countries in the region and personal pressure from those involved thus helped in strengthening the pro-democratic sentiments among the Ukrainian population, encouraging them to protest on the streets.

It is precisely this type of international pressure to democratize that is the topic of this thesis. Observing maps of political regimes in the world over time, or when looking at the number of democratic transitions in the world over time, one observes very clear clustering patterns. Particular regions and particular time periods see significantly more transitions to democracy than other regions or time periods. The goal of this thesis is to develop a theoretical framework to explain these international clustering patterns on the basis of the diffusion of attitudes between individual citizens and the attempt by democratic government to promote democratic attitudes among citizens of nearby countries.

In the literature on regime transitions, and more specifically processes of democratization, the focus tends to be on domestic factors that make a transition more or less likely to start or to succeed. Typical factors taken into account are the economic development of the country (Lipset 1959; Cutright 1963), the existence or absence of a strong civil society (Putnam, Leonardi and Nanetti 1993), the possibility of pacts among groups within the elite (Higley and Burton 1989), or the historical legacy or path dependency of the country (Moore 1966; Linz and Stepan 1996). When regime transitions are largely explained by domestic factors, however, it becomes difficult to explain the clear patterns of clustering we can see on a worldwide scale both geographically and over time.

When one looks at a map of Europe in the 19th century, or one in the 20th, or one after the so-called third wave of democratization in the early nineties (Huntington 1991), one can see whole groups of neighboring countries that show similar patterns of regime changes. Transformations towards or away from democratic political systems often show clear regional patterns. In Eastern Europe in the early nineties, a clear snowballing effect of democratization was visible with the collapse of the Soviet Union and its satellite states. When one studies a map denoting regime types in Africa, Sub-Saharan Africa clearly stands out as a more democratic region than most other parts of the continent. Comparing similar maps of Latin America decade after decade, one can see whole groups of countries introducing democratic reforms in one period and again almost simultaneously reverting these in another. To explain these clustering patterns, in space (clusters of democratic adjacent countries) as well as in time (almost simultaneous transitions to or from democracy in a large number of countries), one needs to look at international rather than purely domestic factors.

The observation that democratic regimes cluster geographically has been made repeatedly, in particular of the past twenty years. Although the first major publication highlighting the pattern was published almost twenty years ago (namely Starr 1991), most of the research has been inspired by the developments in Eastern Europe since the collapse of the Soviet Union. Like a set of dominos, one authoritarian regime in the region fell after the other during the first half of the 1990s. Of course, this might have more to do with the collapse of the major power in the region than with anything like democratic diffusion, it still instigated the renewed attention to geographic clustering in the democratization literature. Factors that are pointed out range from war and conflict (Ethier 2003), to survival in the international system (Cederman and Gleditsch 2004), to the diffusion of democracy as an innovation (Modelski and Perry 1991), to conditionality on aid (Kopstein and Reilly 2000; Levitsky and Way 2005), to the dominance of a normative idea (Fukuyama 1989) and to the spuriousness due to the clustering of other, domestic factors of democratization (O'Loughlin et al. 1998; Elkink 2003; Braun and Gilardi 2006; Gleditsch and Ward 2006). In this thesis the various explanations will be discussed and subsequently complemented with a theoretical model which explains the clustering of democratic regime by the diffusion of attitudes towards democracy. For empirical evidence of the geographical clustering of democratic regimes, see Starr (1991); Ward et al. (1996); O'Loughlin et al. (1998); Ward and Gleditsch (1998); Gleditsch and Ward (1997, 2000, 2006); Brinks and Coppedge (2006); Gleditsch (2002); Elkink (2003); Doorenspleet (2001, 2004); Wejnert (2005), and Fordham and Asal (2007).

Not only regional trends are clearly visible, but also a worldwide trend towards more democracy. Although it is debatable whether one can really talk of waves of democracy (Doorenspleet 2001), it is undoubtedly the case that over time a larger and larger proportion of the world population is living under democratic regimes. In 1800 there were practically no democracies, around 1930 about 30 percent of the world's countries had democratic regimes, and now around 45 percent of them do. For as far as waves are visible, with democratic collapses outnumbering transitions to democracy in the thirties and the sixties, the overall trend is still very clearly upwards. Again, such a worldwide trend can hardly be explained with domestic factors only.

An attempt will be made here to study processes of democratization and democratic collapse in a fashion more suitable to explain this geographical and temporal clustering of regime types. The focus in this thesis will be on the attitudes of the masses towards their regime and the related mass behavior, as opposed to more elitist theories of democratization. This is not to argue that those elites are not or less important in the explanation of democratization, but to highlight just one of the many mechanisms operating in the process of democratization and to show how this might be a possible explanation of the clustering that we observe. The model developed in this thesis will be explaining the patterns described here by way of studying the diffusion of attitudes towards democracy between people, within as well as between countries.

If one assumes straightforward diffusion of attitudes among people, without any specific obstacles, the result is fairly predictable. All deviations of the more common public opinion will gradually be fading out through the diffusion process. Perhaps a few neighbors of the deviating person will temporarily take on some of that person's ideas, but over time, the space of public opinion will be flattened out, with all people having the same, average opinion. Such a situation does not seem to be a valid description of the empirical world as we know it. Even after many centuries, not everybody has the same attitude towards democracy. Whereas democracy is considered almost axiomatically as the best type of regime in many Western democracies, many people in other parts of the world hold completely different attitudes towards democracy, often on the basis of religious conviction, traditional values, or the perceived need for stability rather than freedom. Not only does public opinion vary more than such a flattened model would suggest, also the configuration of democratic and non-democratic states in the world is continuously changing. Revolutions and reforms are still taking place in many areas of the world, in both directions. Since this thesis will presume that a crucial explanation of regime changes lies in public opinion, a flattened opinion structure cannot explain such continuous changes.

While many empirical studies confirm the existence of a spatial clustering of democracies, very few theoretical explanations exist. Generally, it is considered almost natural that proximate political regimes will be more similar than more distant regimes, either by simply referring to the so-called first law of geography that "(e)verything is related to everything else, but near things are more related than distant things" (Tobler 1970: 236) or by arguing that nearby states are more likely to be similar and therefor respond to similar circumstances or are more likely to copy each other's behavior (see, e.g., Bunce and Wolchik 2006: 288). To put it in other words, "the number of diffusion studies continues at a high rate while the growth of appropriate theory is at an apparent standstill" (Katz 1999: 145). Katz refers hereby to the study of diffusion across disciplines and not specifically to the study of democratic diffusion, but the sentiment nonetheless applies. This thesis contributes to the theoretical side of the investigation of the international clustering of democracies, both spatially and temporally.

This leads us back to the main question to be answered in this thesis: "How can the observed patterns of regime clustering and global trends towards democratization be explained on the basis of the diffusion of ideas among people and what obstacles to this diffusion exist that explain its continuing character?". Whereas the key question focuses on the substantial model to be developed in this thesis, it should be pointed out that a significant part will be dedicated to the methodological implications and requirements of the type of model that is to be applied to study this research question.

The research presented here is of a theory-building rather than a theory-testing nature. Clear hypotheses to be tested hence do not exist. The diffusion of democracy is studied from the perspective of a continuous, dynamic process of attitude diffusion. Such a model might or might not lead to geographical or temporal clustering and it might or might not lead to an equibrium state of worldwide democracy. Under what conditions any of these outcomes is observed, and the extent to which such a model can explain the global patterns of democracy we observe is the core question of this project.

The study of the diffusion of democracy is not only of academic significance. Many states have a foreign policy of making more states look similar to theirs. The Soviet Union attempted to convert more and more countries worldwide to their version of communism, the European Union pressures states on its borders to liberalize further, and the United States makes serious attempts to democratize countries in the Middle East, even with the use of force. The recent Iraq war even had an explicit purpose of seeding democracy in the region, thus assuming that once Iraq would be a proper democracy, neighboring states would be likely to follow. This demonstrates how not only academics, but also policymakers are interested in the process of regime diffusion and how this can be influenced (see also Bell and Staeheli 2001).

In terms of the relation between public opinion and regime transitions, a certain body of literature will be assumed to be a correct representation of this link and used as a foundation of parts of the model that is being developed, namely the literature on cascading models of revolution (Granovetter 1978; Kuran 1991*a*, 1995). It serves to return to our example of the recent developments in the Ukraine. An interesting point in this development has been the discrepancy between the reputation of the Ukrainian population as being relatively apathetic and the sudden and determined protests in the streets of Kiev. Of course, pressure from outside and campaigns help, but the discrepancy between apathy and protesting in wintry Kiev for days in a row seems too large to be explained by campaigning alone. A combination of two factors is likely to largely explain this phenomenon. The first factor would be that the little change in attitude as a result of the campaigning might have been just that little bit needed to bring people over a threshold from not protesting to protesting. In other words, their attitude was already very close to that of the protesters, but just needed that tiny little push. The second factor is probably even more important. For those people that had pro-democratic attitudes but were just not passed the threshold to protest, their threshold will have become significantly lower once they saw large numbers of people on the street. Suddenly, they had somewhat less to fear from the authorities, as they would not be standing there on their own in the streets, but in a crowd, and suddenly they knew that they were not the odd exception, but that they had the support of many people in their country. Thus more people started to protest, and the more there were, the more those with a slightly higher threshold felt safe enough to go on the streets as well.

In terms of the attitudinal diffusion process itself, use is made of the social judgment theory. Agent-based models of normative diffusion using this theory (Jager and Amblard 2004) have shown that an implementation of this model in a diffusion context can explain the combination of local homogeneity and global polarization in the same system. It helps to explain how in a society where norms diffuse, different groups still persist, why more than one norm on the same topic can manage to survive. If all individuals had the same, average attitude towards democracy, given the cascading model of revolution, nobody would initiate a protest.

Methodologically, the approach taken in this thesis will be an application of computational agent-based modeling. Agent-based modeling is a relatively new addition to the tool set of the social sciences. Applications in this area have been around for some time, for example Schelling (1978)'s model of social segregation or Axelrod (1986)'s model on the evolution of cooperation, but are only gradually becoming more common, largely thanks to the accessibility of sufficient computational power to large numbers of researchers. Although agent-based models do not necessarily require computers - Schelling did his research purely with pen and paper -, without computers researchers in this area would be limited to only the simplest of models.

A computational agent-based model is a model in which the patterns are studied that result from the interaction between large numbers of actors on the basis of a relatively simple set of behavioral assumptions, simulated in a computer environment. For example, some simple assumptions about the likelihood that a fish of a certain size will eat another fish, the likelihood that a fish will reproduce, and the likelihood that a fish will naturally die, can be modeled in a computer simulation, the analysis of which can provide useful insights in the ecology of fish, which are difficult to trace using other methods of modeling or simulation (DeAngelis and Rose 1992). These models are generally based on assumptions of non-linear relations between variables, due to the fact that actors both create or form their environment, while their environment affects their individual behavior. The behavior of an individual agent is thus dependent on that of many other agents previously, which creates complex patterns not easily deductible from the individual rules of behavior of the agents.

A crucial concept in this area is that of emergence. Emergence concerns the phenomenon of global patterns arising out of individual behavior that is not clearly implied in this behavior. An example would be Schelling's model of social segregation, where the fact that individuals would move when more than half of their neighbors are of a different social group than they are, say black instead of white, leads to a global pattern of complete segregation of the two groups, instead of a just slightly clustered pattern as you would expect from the basic rule of behavior. The global segregation thus emerges out of the model, without being explicitly programmed into it (Schelling 1978).

Most explanations of democratic transitions, or in fact most explanations of political phenomena in general, focus on either of two levels of analysis - they either focus on the micro, individual level, explaining the behavior of individuals like voters or influential members of the elite, or they focus on the macro level, studying for example the link between economic development and the chances for democratization. Although both can deliver interesting insights, what is usually clearly missing is the link between the two levels. How, for example, does a good economic performance influence elites or masses that are more inclined to support democracies? Usually the theorist resorts to fairly general explanations, for example on how a better economic development leads to more literacy among the population and hence a population that is more self-aware of the political situation and more able to voice their opinions, which leads them to press more for democratic reforms. Although some theories do indeed discuss the relation between the macro-level patterns we observe and the individual-level explanations of the results, the exact link between the two is considerably under-studied.

Not only is the relation between micro and macro explanations of political behavior under-studied, also the implications of micro behavior for macro patterns are often misunderstood. The effect of fairly straightforward interactions between individual members of a mass of people can have very complicated and often unexpected effects on the mass behavior as a whole. For example the simple behavior of a car driver, slowing down for cars in front of him or her and speeding up when there is a chance can, given different initial speeds of different cars on the road, easily lead to traffic jams. Simply slowing down for people in front does not trivially lead to traffic jams, yet such behavior, given the diversity in speeds, does have this effect. Moreover, placing traffic lights on a road even without any crossroads can reduce the chances of a traffic jam because they have the effect of homogenizing the speeds of the cars - they all start at the same time at similar distances from each other when the light turns green. Thus the intuitively contradicting idea of stopping cars to avoid traffic jams can actually be quite successful. This example illustrates in a relatively simple way how individual behavior can have unexpected macro effects and this link between local, individual behavior and global, macro behavioral patterns thus deserves attention in social science research.

In this thesis the focus in the study of the global phenomenon of clustering of political regimes will be on precisely this relation between individual behavior of members of the masses and the global patterns that we observe. Relatively simple behavior by individual members of a society can lead to country-wide effects like revolutions and changes of political regimes, and those changes of regimes can in turn lead to global patterns of clustering. Furthermore, this relation is circular in the sense that political regimes will affect the behavior of individuals and that of individuals in other, probably especially neighboring countries. Thus a complex interaction arises between individual behavior of citizens and patterns of political regimes.

Although agent-based modeling is relatively new to political science and only gradually getting somewhat more prominent in the discipline, which became visible for example with the first issue in color of the *American Political Science Review*, where the cover article contained many colorful graphics of an interesting simulation on the secession of countries (Lustick, Miodownik and Eidelson 2004), a number of interesting applications have already been presented. The model just mentioned is indeed a good example of this type of modeling. On a grid representing the country as a whole, with positioned on this grid citizens, government representatives, and rebellion leaders, an abstract visualization was created of the process that determines whether or not a region of a country tends to secession. In terms of visualization, the model resembles that which will be presented later in this thesis. With underlying assumptions and rules of behavior well embedded in the existing literature on secessionist tendencies, the model was designed in a way supporting face validity, and the many runs of the model show patterns that lead to interesting new insights in this area of study.

Another example, or rather, set of examples, is provided by Axelrod (1997*a*). In his work *The Complexity of Cooperation* he bundles a series of articles, each of them containing descriptions of different agent-based models which he designed to further our insight in cooperation. When are people more likely to cooperate rather than defect in certain types of situations? How do norms of cooperation spread among a population? Questions like these are analyzed using these models, which are all of a fairly abstract nature, often dealing with well-known game theoretic situations like the Prisoner's Dilemma.

A third prominent example of a small set of agent-based models are those developed by Cederman (1997), on nationalism and the development and dissolution of nations and states. Describing different plausible models and the different insights that can be deduced from these models, Cederman tries to demonstrate the use of this type of modeling in the field of international relations. A less well-known model of the same author is that which he developed together with Kristian Skrede Gleditsch, a student of democratic diffusion, which models how the clustering of democracies can be explained by the cooperation of democracies in their defense in a hostile international environment (Cederman and Gleditsch 2004). For obvious reasons, this model is of particular interest to this thesis and will for that reason be discussed extensively at a later stage.

Of particular importance in any scientific enterprise is the issue of validity of models that are produced. How do we know whether an agent-based model we produce is in fact a meaningful representation of the real phenomenon, however abstract? When are observations we make on the basis of such models indeed relevant for our insights in how what we try to model operates? All publications discussing the models above dedicate some, but very little attention to this issue. In this thesis, particular emphasis will be put on the validation of the models presented, as a demonstration and defense of an empirical approach to agent-based modeling.

The issue of validity can be divided in two separate categories, internal and external validity. The first concerns the internal logical relations between the different elements of the model. A model has to be consistent in how the conclusions are drawn from the initial assumptions made that underly the model. Agent-based modeling can be compared to a thought experiment (Holland 1995), where specific insights are deduced from a small initial set of assumptions or axioms. The quality of a thought experiment is largely dependent on the consistency of the internal logic. The second, external validity, concerns the relation of the model to the empirical reality. To what extent does the model reflect reality?

An agent-based model can be designed simply to demonstrate how a certain pattern can be created with a particular set of individual rules of behavior. This could be similar to a mathematical endeavor where proofs and deductions are studied of models that are not necessarily related to any empirical phenomena. An agent-based model can also be used to study certain design patterns to create flexible, resistant, adaptive, reliable systems for example for file sharing or computer network coordination. Again, in this case the model does not necessarily relate to any existing empirical system, even though the intention is probably to create such a system based on the model. In a scientific research endeavor, however, the goal of a model is in fact to relate as closely as possible, albeit at a highly abstract level, to empirical phenomena. The ultimate goal is to be able to understand better the nature of things or, in our discipline, of social behavioral patterns. Hence, external validity becomes of crucial importance to the researcher. While the internal validity of a model is to a large extent, although certainly not completely or trivially, implicitly guaranteed by the inherent consistency of computer programs, external validity is a complicated issue to verify and requires considerable attention in any scientific application of agent-based modeling.

In this respect, and in several other ways, agent-based modeling is clearly comparable to models developed within the rational choice paradigm. Rational choice offers a framework for the development of theoretical models of social behavior with special emphasis on logical consistency. The focus of most practitioners of rational choice modeling is indeed on this internal consistency or logical coherence of their model. Many of them are satisfied once they reach the point where the abstract version of the behavior they try to explain could indeed be deduced from their initial axioms and assumptions. The next stage, the external validation or the verification of the relation of the abstract model to the empirical reality, is often neglected, as has been fiercely pointed out by Green and Shapiro (1994). The critique of Green and Shapiro can thus similarly be applied to many existing agent-based models, which also focus on deduction from assumptions instead of empirical validation, while attempting to develop scientific models of observed human behavior.

A key point to be argued in this thesis is that external validation should focus not only on the main, global patterns observed in the model, but also on the individual elements that form the model. Often, validation consists of observing that the global patterns of the model match the observed patterns of the real world, for example in the model of Cederman and Gleditsch (2004), a key validating element is the development of the level of clustering, measured using sophisticated statistical measures, which matches the level of clustering observed using real world data, over time. Although this is of course a crucial part of the validation, it should be seen as a necessary rather than a sufficient validating test. Probably many different models can be designed that result in similar clustering patterns, one attempt being presented in this thesis. To choose between the different possible models, or, indeed, between the model at hand and the unknown possible alternative models, one needs to validate not only the global pattern, but also different smaller elements that lead to the final result in the simulation. For example, the model of Cederman and Gleditsch does not only assume a certain mechanism of clustering, but also specific mechanisms that determine when a country moves troops to which borders, under what circumstances a country decides to attack, when a country breaks up in smaller parts, etcetera. In other words, there are many elements to the model which combined deliver the global pattern observed, but where each of these assumptions might be crucial for this result. Whether these elements are indeed crucial for the result is something that can be tested using different versions of the model itself, but whether these elements indeed reflect real empirical behavior is something that can and should be tested externally. Thus not only global patterns, but also local patterns and constituent elements of the model can be formulated in testable hypotheses and analyzed using empirical data.

Chapter 1 will provide the necessary steps to embark on the study of democratic diffusion. It will clarify the concepts of democracy, clustering and diffusion and subsequently demonstrate that these patterns of clustering can indeed be found in empirical data - as the many confirmations in the literature suggest. Using two different data sets it will be demonstrated that democracy does indeed cluster both geographically and in time.

Chapter 2 provides a brief overview of the many different theories that could potentially explain the international clustering we observed empirically in Chapter 1. Some of these theories have been brought forward with the explicit purpose of explaining the geographic clustering patterns, while other theories are only indirectly relate to the topic, but nevertheless form reasonable potential explanations.

In Chapter 3 we will start to gather the building blocks for the main theoretical model presented here. The chapter will start off with references to the general literature on attitudes and attitude-formation in social psychology and then move on to a more extensive look at the theory of the cascading revolution and the social judgment theory.

The next chapter, Chapter 4, we will describe the theoretical model in precise detail. Using clear mathematical language, the model will be presented in detail. In addition, several specific analyses are used to verify particular claims about the subparts of the model that we are using, coming from other existing models. In this vain, we discuss in detail the work by Jager and Amblard (2004) as well as a brief study of the characteristics of the world system, seen as a (social) network.

The brief Chapter 5 will present the results from the various simulation runs and refer back to the main question of the thesis. The subsequent chapters 6 and 7 will discuss extensively the implications of these findings, both in academic and practical policy-making terms. Particular attention there will be paid to suggestions for further research both in substantive and in methodological areas.

Overall, it will be concluded that although the model provides some evidence for how the international diffusion of attitudes could generate the spatial clustering we observe, albeit more through the explicit attempts by democratic governments to influence those opinions abroad than through inter-personal communication, the temporal clustering in the form of waves are not visible in the simulation output. The conclusion would thus be that individual attitudes cannot be ignored to the extent that is common in studies on democratization, but that the model presented here brings us only part of the way to answering the main question of this thesis.

CHAPTER 1

Democracy and diffusion

A spirit of democratization seems to have gone around through much of Eastern Europe during the early 21st century. Several successful and failed attempts at 'stunning elections' (Markoff 1996: 113-4) took place successively in Serbia, Georgia, Ukraine, Belarus, and Kyrgyzstan. Albeit with democratic constitutions, these countries had leaders solidly in power in part by non-democratic means. Opposition parties of various strengths tried to win the elections by a majority substantial enough to make it impossible for the current leader to stay in power. A strong enough leader can camouflage small majorities through electoral fraud, but large majorities are more difficult to cover up. The successes in some cases became examples for other countries in the region. Representatives from opposition groups in one country became active as mobilizers and advisors to democracy groups in other countries. In other words, the various attempts to revert to a democratic order after autocrats tried to control the elections cannot be seen in isolation and are closely connected to each other. Success in one country became an inspiring example in another country.

These attempts to revolutionize the election outcomes and to replace non-democratic leaders with democratic ones showed clear patterns of temporal and geographical interdependence. The timings of these stunning elections seem closely related in the way they happened one after the other in a short time span. One could talk of a cluster in time of revolutionary attempts. In the remainder of this thesis, we will talk of temporal clustering to refer to this pattern of revolutions following one another closely in time. Secondly, these elections took place in a particular part of the world, as if the voters in these countries were particularly influenced by what happened in other countries in the same region or in neighboring countries. Close neighbors are more likely to be similar, thus successful revolutions in a neighboring country are more like to be replicable than revolutions in countries far away. Furthermore, when countries are nearby there is a higher probability of personal contacts, either through family ties, tourism, or trade, than when countries are far apart. Such personal contacts across borders can form channels of ideological diffusion, spurring revolutionary behavior. These elections were thus clustered geographically, or spatially, as well as temporally.

In this chapter these concepts, of temporal and spatial clustering, will be conceptualized and demonstrated with empirical data. Before we can turn to what clustering is, we need to talk about what is clustering, in other words, we need to briefly look at the conceptualization of democracy and democratization. Since there is already an extensive literature on the conceptualization and measurement of democracy, there is no need to present a lengthy discussion on the subject. It suffices to succinctly place the concept of democracy and democratization as used in this thesis within this literature. Furthermore, before one can present empirical data on these phenomena, we need to know how to measure democracy quantitatively. The first half will focus on these conceptualizations and operationalizations, after which the second half will focus on the clustering or diffusion aspects that we can observe.

1.1 Democracy

1.1.1 Conceptualization

The stunning elections phenomenon shows the difficulty of properly defining the concept of democracy. Officially, these countries already had democratic constitutions, where political leaders were elected through elections and where citizens were protected from government coercion. In practice, however, these countries showed instances of electoral fraud, coercion, limitations on the formation of opposition parties, harassment of members of the opposition, and so forth. The border between a democratic and a non-democratic state can thus be difficult to determine exactly and a precise definition of the concept of democracy is often crucial (Bell and Staeheli 2001). Although lengthy arguments exist on the proper definition of democracy, for example between proponents of seeing democracy as a matter of more or less democratic and proponents of seeing a regime as either a democracy or not. To argue that either one approach is the 'correct' one is not a very fruitful way to go, however. Instead, it makes more sense to take a more pragmatic approach and have the definition used depend on the context and aim of the research project (Collier and Adcock 2001: 532).

The remainder of this thesis will use a concept of democracy in line with the famous definition of democracy of Schumpeter, who states that democracy is "that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote" (Schumpeter 1976: 269). This implies a definition of democracy that is procedural in nature, ignoring for example democratic unfairness due to material inequality among citizens. For a deep analysis of the political system of a particular democratic country this concept might well be too shallow, but for a more global analysis of patterns of democratization, the concept provides a clear distinction between democratic and non-democratic countries, without getting bogged down in philosophical questions on the quality of democracy. For the study of a large number of countries where we are interested in transitions to and from democracy, it is more helpful to look at a definition that makes a clear-cut distinction between democracies and non-democracies than one that provides more information on the quality or depth of democracy. For definitions that are more useful when auditing existing democratic regimes, determining the quality of democracy in a particular country, see, for example, Beetham (1994).

Dahl suggests the use of eight key requirements for a country to be classified as democratic: "(1) freedom to form and join organizations; (2) freedom of expression; (3) right to vote; (4) eligibility for public office; (5) right of political leaders to compete for support (... and) votes; (6) alternative sources of information; (7) free and fair elections; (8) institutions for making government policies depend on votes and other expressions of preference" (Dahl 1971: 3). He subsequently groups these requirements in two main dimensions of political regime classification, political competition and political participation. Only countries that score high on both dimensions are considered full democracies, or, as Dahl prefers, polyarchies. The first dimension is that of participation. For a country to be considered a democracy, a large majority of citizens has to be allowed to participate in the election of leaders. A high level of participation does not guarantee democracy, however. In the Soviet Union, turnout at elections was very high, but the choice on the ballot paper irrelevant (Dahl 1971). Except for an 'against all' option (Oversloot, van Holsteyn and van der Berg 2002), one could only select members or sympathizers of the communist party. The second dimension of democracy is therefor the level of competition among the elites. Only countries with true competition among the members of the elite, and high levels of participation in elections can be considered democratic in the modern, representative sense of the word. Switzerland in the nineteenth century would be a good example of a polity where democracy and political competition were well entrenched, while general suffrage was introduced only in 1867 and 1884 (Dahl 1971), and female suffrage only in the 1970s. One way of seeing those dimensions is to consider them the equivalent of the level (competition) and scope (participation) of democracy - to what extent is there democratic competition and who can influence this competition? According to Bollen (1991), this distinction is unclear and the two, level and scope, are really in effect the same thing. More participation leads to more competition and more competition to more participation. See Figure 1.1, by way of defense against Bollen. If the two were measuring the same thing, this plot would show the main density of cases on the diagonal, while in fact most countries are off the diagonal. Among the cases without much in terms of participation, there is a wide variety of levels of competition among the elites.

It should be noted that neither the competition nor the participation scales differentiate very well between different types of authoritarian regimes. Whereas they are good instruments to separate democracies from non-democracies, as well as semi-democracies from true democracies, they will help little to distinguish monarchies from military dictatorships from totalitarian regimes. An exception to this might be communist-type totalitarian regimes where

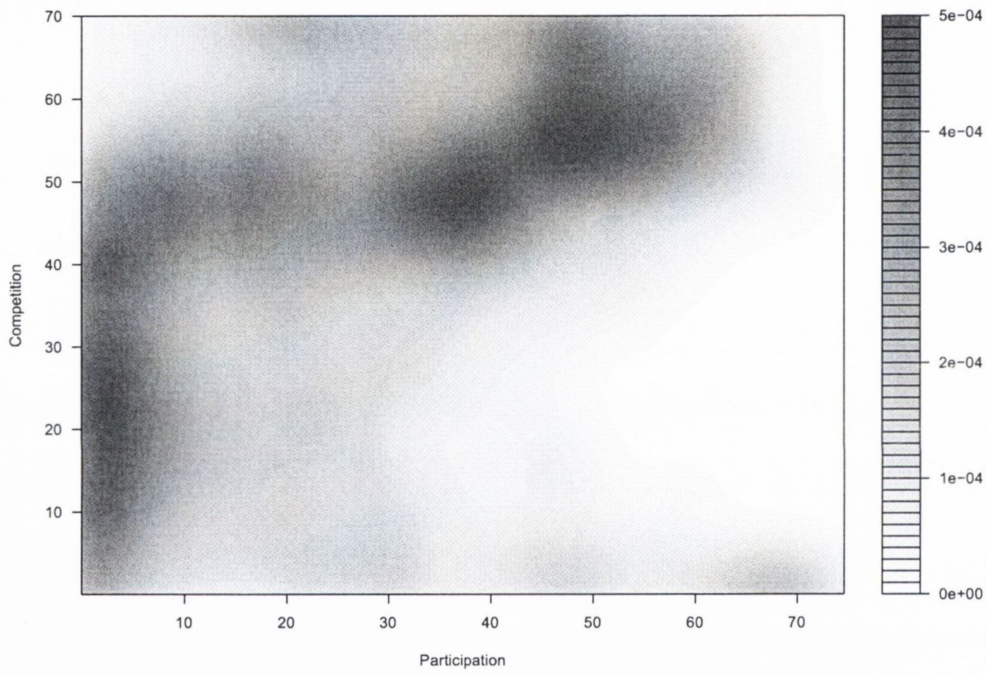

Figure 1.1: Density of regime-years in Dahl's two dimensions of democracy, 1816-2000. Based on measures of democracy by Vanhanen (1997).

elections are still performed, albeit with a very restricted number of candidates. In these cases levels of participation are relatively high compared to other non-democratic regimes. Especially when discussing the different paths that are being followed in the two-dimensional space created by this typology, the lack of information on the non-democratic types of regimes is a disadvantage. The dimensions are clearly dimensions of democracy, rather than the core dimensions of regime classification in general.

On the basis of the above conceptualization of democracy, the concept of democratization simply refers to movements towards more competition, more participation, or both, in a particular country. For empirical research into democratization, the distinction of the two dimensions of democracy is important. Criticizing the measures in the Polity IV data set (Jaggers and Gurr 1995; Marshall and Jaggers 2002), which concentrate on the dimension of political competition, Doorenspleet (2001) suggests to alter the measurement of democracy as common in empirical research by adding the dimension of participation. Only countries that have both high levels of competition and high levels of participation can be considered democratic. When trying to explain changes in levels of democracy, this revised measurement of democracy becomes problematic. Although it is correct to state that regimes with low levels of participation cannot genuinely be considered democratic in the modern sense of the word, the combined measure conflates two conceptually different features of democracy. The extension of universal suffrage in most older democracies took place much later than the establishment of democratic regimes in the first place and would require a different type of explanation in a framework of democratization. Figure 1.2, explained below, shows this important distinction. If extensions in the level of participation and the level of competition were very similar and would occur under sim-

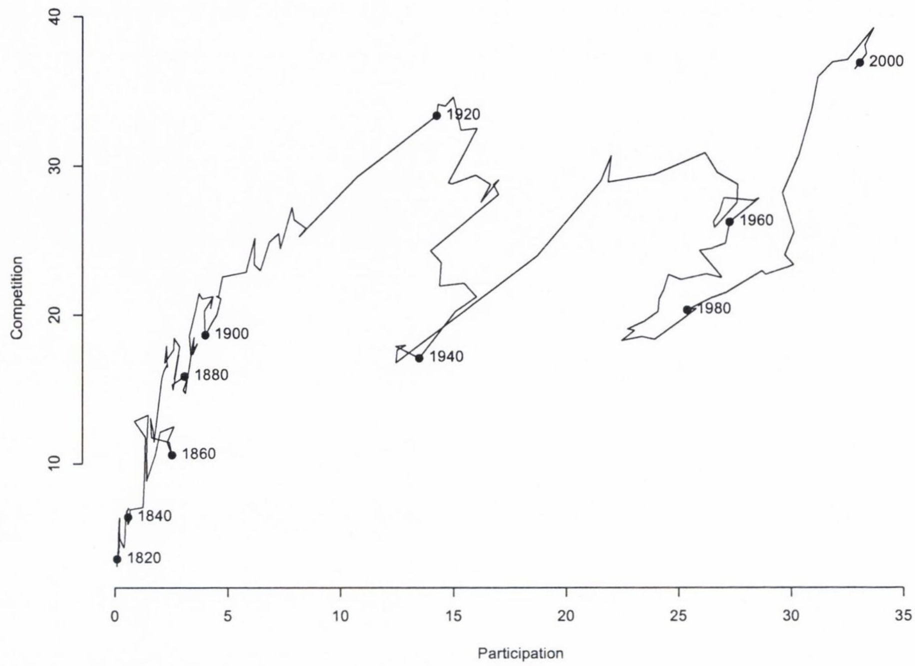

Figure 1.2: Time path through Dahl's two dimensions of democracy, 1816-2000. Based on annual mean values of Vanhanen's democracy scores.

ilar circumstances, one would expect more or less a 45 degree line in this graph. Instead, the slope of the path through time differs considerably over time. In other words, for a theoretical understanding of democratization it is essential to consider transitions that involve the widening of political contestation among the elites as a different phenomenon from the extension of voting rights to larger parts of the population.

In our concept of democratization we are still well within the procedural concept of democracy, leaving aside concerns of the quality of democracy in established countries that are widely considered to be democratic. Modelski and Perry (1991: 24) make a useful distinction between vertical and horizontal democratization, the first referring to the deepening of the quality of democracy within a procedurally democratic country, and the second to the spread of democracy over more countries hitherto undemocratic. This thesis concerns itself entirely with horizontal democratization and ignores issues of vertical democratization. Whether the electoral system of Great Britain is more democratic than that of the Netherlands is an interesting question, but does little to shed light on the diffusion of democracy across the globe. When talking about democratization in this thesis, this refers to relatively sizable movements in Dahl's two dimensions of democracy, on the extension of voting rights to considerable parts of the population or the substantial extension of possibilities for competition among the political elites.

Below use will be made both of the Polity IV dataset (Marshall and Jaggers 2002), which is an often-used measurement of the level of competition in a political regime, and the dataset of Vanhanen (1997), which is an attempt to measure the two dimensions of Dahl independently. In the theoretical model that will be developed further in this thesis, the concept of democracy will be simplified to a dichotomous one. This simplification reduces little of the applicability of the model while making the modeling attempt itself more straightforward. For this reason, in this chapter we will demonstrate the presence of both temporal and geographical clustering of democracies using both continuous and binary measures of democracy, the former using both dimensions of democracy and the latter based on the level of competition as measured in the Polity IV data set.

1.1.2 Measurement

In the research of Vanhanen, the level of competition is measured by "the smaller parties' share of the votes cast in parliamentary or presidential elections, or both", which "is calculated by subtracting the percentage of the votes won by the largest party from 100" (Vanhanen 1997: 34). It is obvious that this indicator is a distant proxy of the complex concept of competition. This type of proxy has two main advantages. One is its simplicity and related objectivity. It is a clear, straightforward measure without much subjective judgment involved - as is the case with almost every other indicator of democracy - and the value of the measure leaves very little doubt in almost any case under study. Secondly, although the indicator is only a proxy, at least it is very clear what exactly it is measuring. It is probably better to have a proxy where you can, thanks to its clarity, estimate or argue how it would differ from the actual concept that one attempts to measure, than to have a proxy where one knows it is not accurate, but where the complexities and subjective judgments involved mean that it is difficult to estimate the unavoidable level of bias. Vanhanen himself combines the two dimensions in one Index of Democracy by multiplying the two dimensions and dividing the outcome by 100, but this Index of Democracy will not be used in this thesis. As argued above, transitions in the two dimensions are very different processes and should not be confounded in one measure of democracy. For the level of participation Vanhanen again uses a very straightforward measure, namely the level of turnout in the elections, which is calculated as "the percentage of the total population who actually voted in the election concerned" (Vanhanen 1997: 34).

Figure 1.1 shows the coverage of the space of political regimes as defined by Dahl and as measured by Vanhanen.[1] It is clear that all corners of the space are populated with countries in particular years. The combination

[1] All code for the empirical plots and analyses presented in this chapter can be found in Appendix B. This density plot ignores the cases where both participation and competition are zero: this point has by far the highest density and renders the remainder of the plot uninterpretable.

of high levels of participation with low levels of competitions, as in many communist countries, is relatively rare. One would almost be tempted to agree with Bollen (1991) that the two dimensions are highly correlated, but although some non-linear relationship between the two exists, this does not suggest that the two measure the same underlying dimension.[2]

It is easy to criticize the measures by Vanhanen. The turnout at elections depends on a lot more than the political rights of citizens. Turnout in Belgium, for example, is high due to compulsory voting rather than because Belgium is so much more democratic than neighboring countries. Turnout in Switzerland is low, probably because people are tired of too many votes and elections, which can hardly be denoted as a sign of a lack of democratic culture in Switzerland. There is no doubt that turnout is affected by many factors and only to a limited extend a good indication of the extend to which the population of a country is allowed to participate in the democratic process (Bollen 1991: 4, 10). A comparison of Vanhanen's measure of participation and Bollen's measure of political suffrage (Bollen, Jackman and Kim 1996) demonstrates the problem Bollen points out. The level of participation varies widely between no participation at all and the maximum limit allowed by law. It is however a well-known fallacy to equate democratic rights by law with *de facto* democratic regimes. The constitution of the Soviet Union as developed under the leadership of Stalin in 1935 is one of the most democratic constitutions ever written. In other words, whereas Vanhanen's measure of participation is biased downwards by ignoring many factors that reduce turnout in democratic countries, Bollen's measure of suffrage is biased upwards by equating law with practice. The great advantage of the measure is that it is possible to collect data consistently, fairly ob-

[2]The correlation coefficient is .53, and clearly significant, but this ignores the panel structure of the data.

jectively, over a long period of time and that the differences between the indicator, turnout at elections, and the concept, scope of the right to participate, are fairly straightforward. It is a preference for a simple, objective, consistent indicator over a more complex, subjective, and therefore inconsistent indicator, even if the latter might be theoretically more accurate. The problem is that in the latter case, it is difficult to know when it is correct and when it is not.

The competition indicator is similarly straightforward and objective, while being only a proxy for what it is supposed to measure. As Bollen points out, it "confound[s] a multi-party system with political democracy" (Bollen 1991: 11). Especially in terms of changes in the level of democracy it might seem awkward that when the largest party increases its seat share from 30 to 40 percent, this would indicate and significant reduction in the level of democracy in that particular country (Bollen 1991: 11). Again, it is a choice for a simple and objective measure over a complex measure like, for example, the measures in the Polity IV dataset (Jaggers and Gurr 1995; Marshall and Jaggers 2002), which in essence are also indicators of the level of competition. When drawing conclusions about particular countries in particular time periods, these measures would be too rough to be sensible. For the investigation of more broad patterns, however, as in this thesis, it makes sense to use simpler indicators. Despite all critique that is possible, it is still reasonable to conclude that a country where the largest party has a very large proportion of the seats in parliament and where only a very small proportion of the populations turns out at elections, is not a democratic country. On average, this indicator can be expected to perform well.

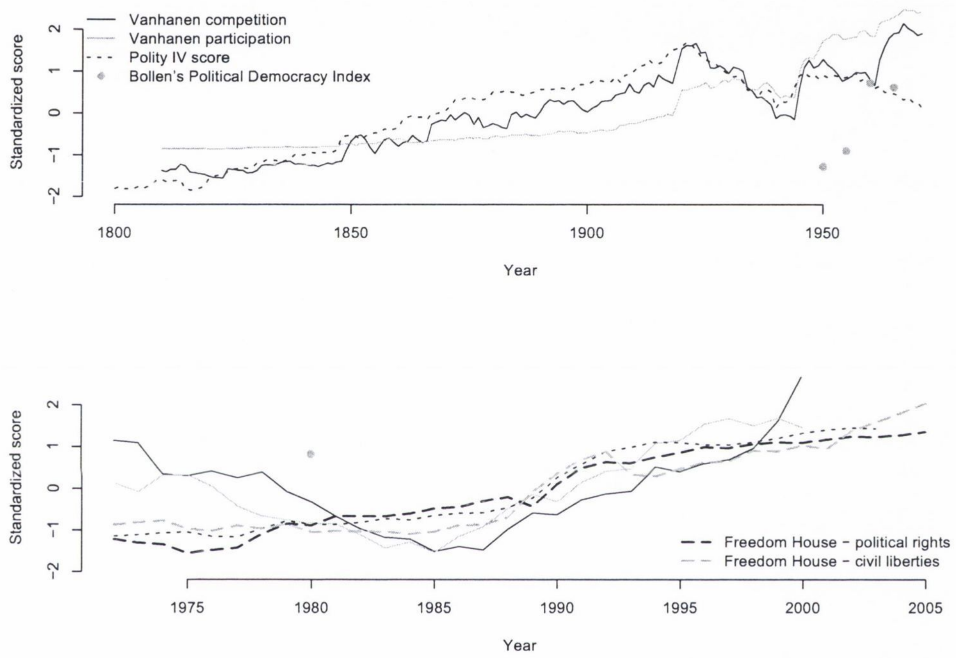

Figure 1.3: Comparison of measures of democracy, 1800-2005. For each measure, the mean within each year is taken. The measures are standardized to make the scales comparable. Separate plots were produced for years before and after 1972, which is the year where the Freedom House data starts.

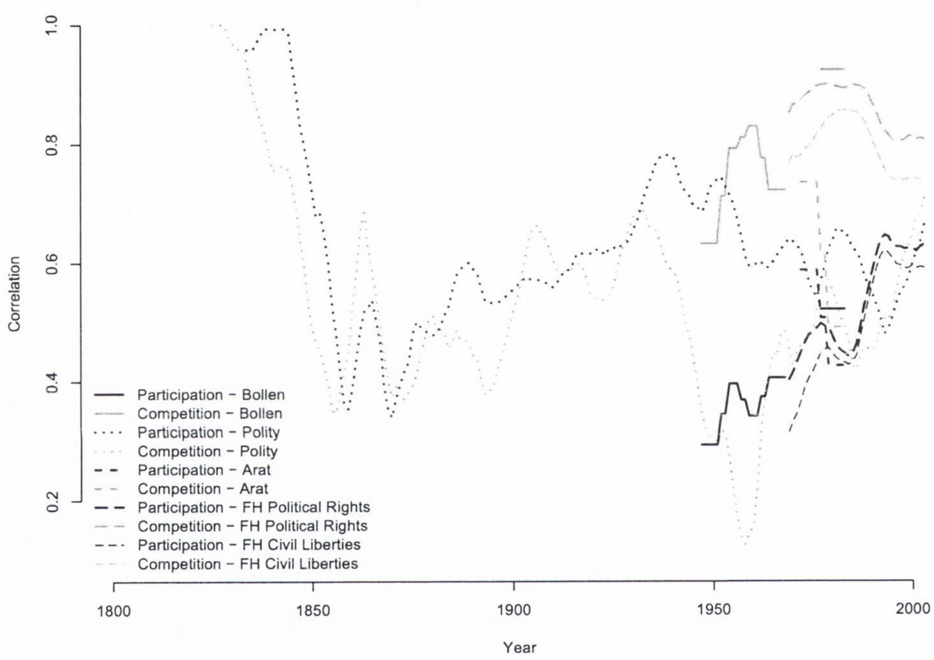

Figure 1.4: Correlations of measures of democracy, 1800-2003. These values represent cross-sectional correlations of measures within each year.

Despite the attractiveness of the Vanhanen indicators, we should confirm this for as far as possible by comparing the measures with other commonly used indicators of democracy. The most frequently used indicators in empirical work are probably those developed by the Polity IV data set (Marshall and Jaggers 2002). The indicators in this data set are primarily concerned with the level of competition among the political elites and measure separately the presence of typical democratic features and more autocratic features to the political system. The eleven point scale for autocracy is subsequently subtracted from the eleven point scale for democracy, generating a twenty-one point overall score. This is the score represented in Figure 1.3, which compares different measures of democracy,[3] as well as Figure 1.4, which depicts the correlations between the various measures, for each year.[4] In terms of the global level of democracy, the indicators by Vanhanen are very similar to the Polity score until the 1960s. After that, the two indicators diverge somewhat more. This divergence is also clearly visible in Figure 1.4, where the correlation between the Polity score and Vanhanen's measure of competition reaches a low point of less than 0.2. In the empirical analysis below, most focus will be on the measures by Vanhanen, but they will be complemented by a dichotomized version of the Polity IV scores, counting all countries with a Polity IV score of 6 or higher as democratic, which is common practice (see, e.g., Fordham and Asal 2007).

Considering his criticism on the Vanhanen indicators, it is interesting to compare the Vanhanen indicators also to Bollen (2001)'s indicator of democ-

[3]It should be noted that these measures are based on different samples - they do not cover the same years, nor do they all sample the same countries - which implies that one should be careful comparing them. In particular, the fact that the Freedom House scores appear so much lower than the other indicators is due to the fact that the standardization of the data was based on a much shorter span of years. The trends are parallel to the other indicators, which is the most important fact that can be distilled from this comparison.

[4]The lines are smoothed over seven years.

racy. The latter would not be very useful for long-term comparative research, since the data only exists for a few years, but as a validation it is still useful. The correlation between Vanhanen's participation score and Bollen's democracy score is quite high - consistently above 0.6 - while the correlation with Vanhanen's competition score is relatively low - around 0.4. Another well-known data set of democracy scores is the one developed by Arat (1991), which is also available for just a few years. The Vanhanen scores correlate fairly well with this indicator, around 0.5. Considering the fact that all these indicators are in theory measuring the same thing, however, the correlations are in fact strikingly low. A clear indication of the complexity of properly measuring democracy in quantitative terms.

It should be pointed out off-hand, however, that even when some indicators seem better measures for democracy, limited availability of data often hampers research using those scales. For example, Coppedge and Reinicke (1991) provide an interesting, unidimensional measure of democracy based on Dahl (1971), but provide results only for 1985 and require information virtually impossible to acquire as far back as the nineteenth century.

There is also a substantial debate in political science as to whether it is more appropriate to measure democracy as a dichotomous variable or as a continuous one. Bollen (1991) argues that since democracy is a matter of degree and thus by definition continuous, it should be measured as such. His conclusion is to a large extent driven by his definition of democracy, however, which defines political democracy as "the extent to which the political power of the elites is minimized and that of the non-elites is maximized" (Bollen 1991: 5). The definition of Schumpeter, as cited above, lends itself much more for a dichotomous interpretation. A regime can be classified as a democracy when its elites acquire power through elections and cannot be classified as such when they do not. In her argument against including regimes without female suffrage as democracies, Doorenspleet (2001) takes a similar stance on this issue and creates a clear threshold which regimes have to cross to be counted as democracies. Dahl's two dimensions of democracy lend themselves well for a continuous approach, where countries can move through the two dimensional space and take any position possible. An arbitrary threshold, separating the corner of high participation and high competition from the rest will only lead to artificial analytical results. This arbitrariness is clearly demonstrated by the refutal of Huntington (1991)'s waves of democracy by Paxton (2000) and Doorenspleet (2001). In both cases, the inclusion of female suffrage as a requirement for being counted as a democracy leads to a measure in which these waves are not visible anymore. This conclusion, however, conflates extensions of competition with extensions of participation, which are two entirely different social developments. The results become arbitrary by the dichotomization on a two-dimensional concept. Figure 1.3 clearly shows the wave-like patterns in the two separate dimensions of democracy. The dichotomization of democracy used later in this chapter avoids this conflation by using the Polity IV data set, which only concerns the competition dimension of democracy.

1.1.3 Diffusion, contagion and clustering

Most definitions of diffusion focus on the spread of ideas between individuals or groups of people. Welsh, for example, defines diffusion as "the process by which institutions, practices, behaviors, or norms are transmitted between individuals and/or between social systems" (Welsh 1984: 3). Rogers defines diffusion (of innovation) as "the process by which an innovation is communicated through certain channels over time among the members of a social system" (Rogers 1995: 5). More specifically within the literature on the diffusion of democratic regimes, Kopstein and Reilly state that "[d]iffusion (...) is a complex process that involves information flows, networks of communication, hierarchies of influence, and receptivity to change" (Kopstein and Reilly 2000: 12). In another useful definition, "[d]iffusion can be defined as a process wherein new ideas, institutions, policies, models or repertoires of behavior spread geographically from a core site to other sites, whether within a given state (...) or across states" (Bunce and Wolchik 2006: 286). This thesis will make use of the latter conceptualization of diffusion, applied specifically to the spread of attitudes towards democracy. In this and the next chapter, the terms diffusion, clustering, and contagion will be used more or less interchangeably, generally referring to the fact that democracies cluster both spatially and temporally on a world-wide scale. In the later chapters of the thesis, however, the model of diffusion that is being developed is one of attitudinal diffusion, whereby the idea that spreads geographically is an attitude to democracy held by individual citizens, which in turn explains the empirically observed patterns of clustering. Clustering will consistently refer to the actual globally observed pattern of democratic regimes, independent of the explanation of this pattern or the type of mechanisms that bring it about.

As an attitudinal diffusion process, this contagion can be seen as the sociological equivalent of epidemics in medicine. Ideas, customs, norms, and as a result institutions spread over a society much like a disease does, by personal contacts between 'victims'. And just like with diseases this leads to clusters of people that are 'infected'. One should be careful not to stretch this analogy between social diffusion and medical epidemics too much, however, as in the adoption of ideas or norms more factors play a role than direct contacts. Media spread information over large areas at once or even on a worldwide scale and there are stark difference between different individuals in the extent to which they infect others. In fact, this international contagious effect is likely to increase over time as individuals are more and more aware of international developments and their own position in this context (Rosenau 1988: 359).

The key empirical pattern that is visible which this thesis attempts to explain is that of clustering of political regimes, in particular democracies. It is clear from the empirical literature on democratic diffusion and that on waves of democracy, that democratization is not an entirely domestic affair. Processes of democratization are affected by international developments. That is to say, the clustering we observe is statistically significantly different from a random distribution of democracies, which strongly suggests that these instances of democracy are not independent. This chapter will demonstrate this empirically, as well as discuss the empirical literature on the subject. To avoid confounding very different processes that lead to clustering, it is useful to distinguish between three different types of empirically observable clustering, despite the fact that the three overlap both in statistical terms as in theoretical explanations. O'Loughlin et al. (1998: 561) use the same classification of types of clustering, or in their terms, of democratic diffusion.

The first type of clustering is temporal in nature. When one studies purely the timing of democratizations across the world, one notices that these appear to happen in waves. Particular periods in time with many countries making a transition to democracy and other periods with a lot of democracies collapsing and return to autocratic regimes. These waves of democracy have been popularized by Huntington (1991). Throughout the remainder of this thesis, this will be interchangeably called waves of democracy or temporal clustering.

The second type of clustering is spatial or geographic in nature. Observing maps of the world at different points in time over the past two centuries, colored according to the presence or absence of a democratic regime in each country, one notices the clustering of democratic regimes. Regions of contiguous countries are all democratic or all autocratic. One can show statistically, as we will in this chapter, that this clustering is significantly different from a random distribution of democracy on a world map (Gleditsch 2002). This type of clustering will be denoted as spatial clustering, or at times geographic clustering.

The third and last type of clustering is a combination of the two, thus denoted spatio-temporal clustering. Spatio-temporal clustering refers to the geographic clustering visible in patterns of democratization, rather than just the presence of democracy. Not do we just have a higher chance of finding a democracy near other democracies, we have a higher chance of finding transitions to democracy near other transitions to democracy, within a relatively short time frame. Whole regions of contiguous countries see their democratic regimes collapse one after the other, or vice versa, make each a transition to democracy, just years or months apart. The most famous example of this is the many transitions to democracy in Eastern Europe in the early 1990, the so-called fourth wave of democratization (Doorenspleet 2001).

Table 1.1 summarizes the three types of clustering and the remainder of this chapter is organized according to this scheme. For each of the different clustering patterns statistical evidence will be presented, as well as references to the empirical literature. This will lay down the empirical observations that this thesis attempts to explain, through the development of a theoretical model and computer simulations. While this chapter concerns

	Spatially random	*Spatially clustered*
Temporally random	Democratization and democracy are randomly distributed in time and space, and can be explained by purely domestic factors.	**Spatial clustering**: Democracies are clustered on the map, but the timing of their transitions is random.
Temporally clustered	**Temporal clustering**: Waves of democracy affect the planet as a whole and chances for democratization are affected by the *Zeitgeist*, the spirit of the times.	**Spatio-temporal clustering**: Democracies are clustered on the map. The chances for democratization are affected by transitions to democracy in neighboring countries.

Table 1.1: Three types of clustering

itself purely with empirical observations, the next chapter will discuss the various existing possible explanations of these patterns, as a contrast to the explanation provided in this thesis.

1.2 Regime clustering

1.2.1 Temporal clustering

The Russian Revolution of 1905 was the first revolution reported daily through telegraphic wire services, stirring revolutionary developments in Iran (1906), Ottoman Empire (1908), Portugal (1910), Mexico (1910-12), and China (1911-12) (Kurzman 1998: 54).

> "Events in Russia have been watched with great attention, and a new spirit would seem to have come over the people. They are tired of their rules, and, taking example of Russia, have come to think that it is possible to have another and better form of

government" (Browne 1910: 120) (as cited in Kurzman 1998: 54)).

> "Partly owing to the globalization of mass communications, protest has become a trans-national phenomenon. In the post-World War II era, there have been two major waves of anti-government protest movements on a global scale. (...) Shortly after the events in China, anti-Communist protests erupted in East Germany, Romania, the Soviet Union, and other Eastern Bloc states, which led to the collapse of one Communist regime after another" (Zhu and Rosen 1993: 234-235).

From reading historical accounts it becomes clear that at times a change of spirit is occurring in the world. A new idea or a new world view quickly gains support and spreads across the globe, or at least large parts of it. Generally referred to as the *Zeitgeist*, the spirit of the times, particular eras experience particular common ideologies and values uncommon in other periods. Whereas it was once commonly accepted that certain countries were superior to others and demonstrated their power by claiming large overseas territories, which provided cheap labor and great economic benefits, after the Second World War colonialism suddenly had a strongly negative connotation, and countries had to rush to get rid of their colonies, sometimes rather reluctantly. The common attitude towards colonization changed dramatically - a new *Zeitgeist* came about. A similar pattern can be observed in relation to female suffrage during the early twentieth century. While it was once commonly accepted that women should concentrate on private affairs, while the men concern themselves with public issues, including participating in politics, it is now practically inconceivable that women would be excluded from the vote. Global attitudes towards the gender roles have changed dramatically thanks to emancipation and feminism movements.

Both examples are closely related to democratization. According to modern concepts of democracy it seems inconceivable to consider countries that refused the female vote as full democracies (Paxton 2000; Doorenspleet 2001), even if nobody at the time would have had any doubts classifying the regime as such. The extension of female suffrage can clearly be seen as a temporal wave of democratization, where various countries in a short period of time follow each other in their transition to democracy. A global wave of democratizations takes place.[5] The *Zeitgeist* can move in all directions, of course, and other ideologies can similarly take hold of the minds of many people in a short period of time. Communism and fascism are both examples of non-democratic ideologies that captured the minds of many people around the world. These changes in ideologies also affected worldwide changes in political regimes and created a reversal of many democratic regimes. The wave of democratization was compensated by a wave of reversals to new types of dictatorships.[6]

Based on the average levels of competition and participation around the world, Figure 1.2 shows the path through time of political regimes around the world, in Dahl's two dimensions of democracy. The waves of democracy are clearly visible in the graph, with peaks of democracy in the early 1920s, the 1950s, and now. Stimulated by the developments of the time, political science research in democratizations follows the pattern closely, and explains the recent increased attention to waves and diffusion of democracy. The long jumps in levels of participation around 1918 and 1919 are clearly visible

[5] As discussed above, this would be a global wave of democratization on the participation dimension of democracy, quite clearly distinguishable from a wave of democratization on the competition dimension.

[6] A "reversal" of democratization, but of course on the other dimension, competition.

and reflect the extension of female suffrage in many democratic countries. The steepness of the line in the 1990s shows the extension of competition, allowing opposition parties to participate fully in national politics, in countries where levels of participation were already high, namely the chain of democratizations after the collapse of the Soviet Union.

Crescenzi and Enterline (1999) demonstrate using a vector auto-regression model that the number of democracies positively affects democratization and that there is strong support for a snowballing effect of democracy, but that these effects are very much time and geographical region dependent. Using data on transitions towards and away from more freedom for the years 1974 to 1987, Starr (1991) shows that transitions to democracy on a world-wide scale are not randomly distributed over time, implying that some kind of contagion-effect is likely to be present. The likelihood of a transition changes when another event took place in the same year, thus they are not independent of each other (Starr 1991: 366). Seeing these international tendencies to various regime transitions in terms of waves has become common since Huntington's *The Third Wave of Democracy* (Huntington 1991; Markoff 1996; Doorenspleet 2000, 2001). Following this logic, we are now awaiting the next wave of breakdowns of democracies, following the many successful democratizations in the last two decades (Starr 1991).

1.2.2 Spatial clustering

From the dynamic pattern of temporal clustering, we now turn to the static pattern of spatial clustering. When observing world maps at any point in time, ignoring the temporal dynamics that created the map, one observes clusters or geographical zones of democracy and other clusters of dictatorships. Western Europe is a zone that has long known democratic regimes, while Africa is a continent that has shown large patches of dictatorship over most of the past decades. Exceptions always exist, such as Botswana in Africa or contemporary Belarus in Europe, but the clusters are nonetheless clearly visible.

Although the next chapter discusses various theoretical explanations for this type of clustering in more detail, it is useful to think briefly about the kind of processes one could imagine to explain this type of clustering. The emphasis here is on the static pattern of clustering - not democratizations that stimulate democratizations in a neighboring region, but patterns of clustering visible after a long period of changes and stability. Although the dynamic equivalent of this pattern, spatio-temporal clustering, would also show spatial clustering at any one point in time, the expected mechanisms that bring about this clustering are very different. The empirical observation that a country has a higher chance of being a democracy when surrounded by other democracies would lead one to believe that the probability of a transition to democracy increases when a neighboring state democratizes. This is not necessarily the case, however.

One possible explanation that does not involve this pattern of a snowballing of democratization within a contiguous set of states is one that focuses on the chances of survival for new democracies. In this perspective, the probability that a particular country makes a transition to democracy is independent of its international context, but the chances for its survival do depend on the geographical context. Embedded in the theory of the democratic peace, which assumes that democracies do not attack each other, the assumption is here that new democracies in a democratic environment have less chance of ending up in military conflict, and democracies will help each other in times of military threat, thus increasing the stability of the new regime (Cederman and Gleditsch 2004). This would lead to a spatial clustering of democracies, but not of transitions to democracy (spatio-temporal clustering).

Another possible explanation would focus more on the prerequisites of democracy in socio-economic terms (Lipset 1959) or in geographical terms - or *fortuna*, as Machiavelli (1531) calls it. Geographically contiguous countries often share similar geographical features that are either advantageous or disadvantageous for their economic development or modernization. Countries in oil-rich regions are operating in an entirely different context economically speaking than do countries in mineral-poor countries with extensive periods of drought. Various theories that focus on domestic causes of democratization can thus implicitly explain patterns of spatial clustering as well, without suggesting that transitions in one country affect processes in neighboring countries. In the next chapter this is discussed under the heading of spurious diffusion.

The human brain is trained in seeing patterns, even where none exist: most people will have the experience of dreamily staring at the clouds and seeing all kinds of faces or other patterns in the clouds. The same holds when looking at a geographical map of political regimes: a casual human observer will find patterns even if the map came about purely by chance. To validate the observation of spatial patterns, we need a statistical test to determine whether the clustering is significantly different from what we would observe if democracies were distributed randomly in space.

To be able to measure the level of clustering, the first step has to be to define a matrix defining which units in the dataset are connecting to which other units. In the remainder of this section a square contiguity matrix W_t

of dimension $n \times n$ is used, which is defined as follows:

$$w_{ijt} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected at time } t \\ 0 & \text{otherwise,} \end{cases} \qquad \text{where} \quad w_{iit} = 0 \quad \forall \quad i, t \tag{1.1}$$

Subsequently this matrix is standardized for each year such that the rows all add up to one. Without this standardization, the fact that different units have different numbers of neighbors would lead to heteroscedasticity in any estimators using W (Tiefelsdorf 2000). The standardized matrix will be denoted as $\tilde{W}_t$.[7] For this section the contiguity data has been based on the Correlates of War project, using the Direct Contiguity dataset (version 3.0), which includes land borders and overseas neighbors for different distances for countries worldwide between 1816 and 2000 (Stinnett et al. 2002). In this analysis only land borders have been taken into account. Taking sea borders into account, the Netherlands border on France, but not on Luxembourg, which does not make any sense from the perspective of a theory of diffusion.

For the global level of clustering, that is, the extent to which overall the regimes are clustered together, as opposed to a local measure of the amount of clustering at a specific location, the most common indicator is Moran's I (Moran 1948, 1950; Anselin 1988; le Gallo 2000; Gleditsch 2002). The

[7] Note that multiplication of a variable x with this matrix, $\tilde{W}x$, is the equivalent of creating a vector with for each unit i the average value on x of all units directly contiguous to i, excluding the value on x for unit i itself.

formula for Moran's I, as presented by le Gallo (2000), is as follows:[8]

$$I_t = \frac{\sum_i \sum_j \tilde{w}_{ijt}(x_{it} - \overline{x}_t)(x_{jt} - \overline{x}_t)}{\sum_i (x_{it} - \overline{x}_t)^2} \cdot \frac{N_t}{\sum_i \sum_j \tilde{w}_{ijt}} \tag{1.2}$$

$$= \frac{\sum_i \sum_j \tilde{w}_{ijt}(x_{it} - \overline{x}_t)(x_{jt} - \overline{x}_t)}{\sum_i (x_{it} - \overline{x}_t)^2}, \tag{1.3}$$

$$\text{where} \quad \overline{x}_t = \frac{1}{N_t} \sum_i x_{it} \tag{1.4}$$

With $\tilde{w}_{ijt}$ being the standardized value of the contiguity matrix W_t, x_{it} being the respective score for unit i in year t, and N_t being the number of observations for year t. Plotting these scores over time, one gets an impression of the level of global clustering over time. In a situation where contiguous units tend to be different from each other, one of the two having a value on x below the average and the other a value above average, Moran's I will be negative. Moran's I is positive in the opposite case, where contiguous units tend to be similar, either both above or both under the average. When Moran's I does not significantly differ from its expected value, $-1/(n-1)$ (Griffith and Arlinghaus 1995: 27), the pattern does not differ from a random distribution of units, neither in the sense of being clustered nor in the sense of being unclustered in a structured way. Thus, if we have a checkered pattern of democracies being surrounded by autocracies and vice versa, we will see a negative score on Moran's I, while if we have a clearly clustered pattern of democracies surrounded by democracies and autocracies by autocracies, we will see a positive score. With an overall pattern where all countries are either democratic or autocratic, thus all are in a sense together one large cluster, Moran's I will be positive and highly significant.

Under the normality assumption, the standard error of this statistic can

[8]Due to the standardization of W, $\sum_i \sum_j \tilde{w}_{ijt} = N_t$, assuming there are no "islands" in the dataset.

be calculated as follows (Gleditsch 2002; Cliff and Ord 1973):

$$E(\sigma^2_{I_t}) = \frac{N_t^2 S_1 - N_t S_2 + 3S_0^2}{S_0^2(N_t^2 - 1)} \tag{1.5}$$

where

$$S_0 = \sum_i \sum_j (w_{ijt}+w_{jit}), S_1 = \frac{1}{2}\sum_i \sum_j (w_{ijt}+w_{jit})^2, S_2 = \sum_i \sum_j (\tilde{w}_{ijt}+\tilde{w}_{jit})^2 \tag{1.6}$$

Since in the analysis presented here w_{ijt} can only have the values 1 or 0, and W_t is a symmetrical matrix, this can be reduced to

$$S_0 = S_1 = 2\sum_i \sum_j w_{ijt}, S_2 = \sum_i \sum_j (\tilde{w}_{ijt} + \tilde{w}_{jit})^2 \tag{1.7}$$

In Figure 1.5 we can see that the level of clustering is relatively erratic until the last years of the 19th century, but positive and significant ever since. The slightly increasing confidence in our estimate - the smaller grey band of uncertainty - can be explained by the increase in the number of countries over time, thus the larger number of cases. The initial high level of clustering on the participation dimension can be ascribed to the lack of participation worldwide. This is thus the example of the whole world being one cluster, with low levels of participation in all countries. The following negative and volatile level of clustering has to be ascribed to the first countries adopting more participatory regimes, which is apparently occurring in a spatially random pattern, breaking the original clustering of low participation and for some years creating a more checkered pattern. In later years, the level of participation increases, as can be seen in Figure 1.3, and starts to cluster around the original adopters of more participatory regimes.

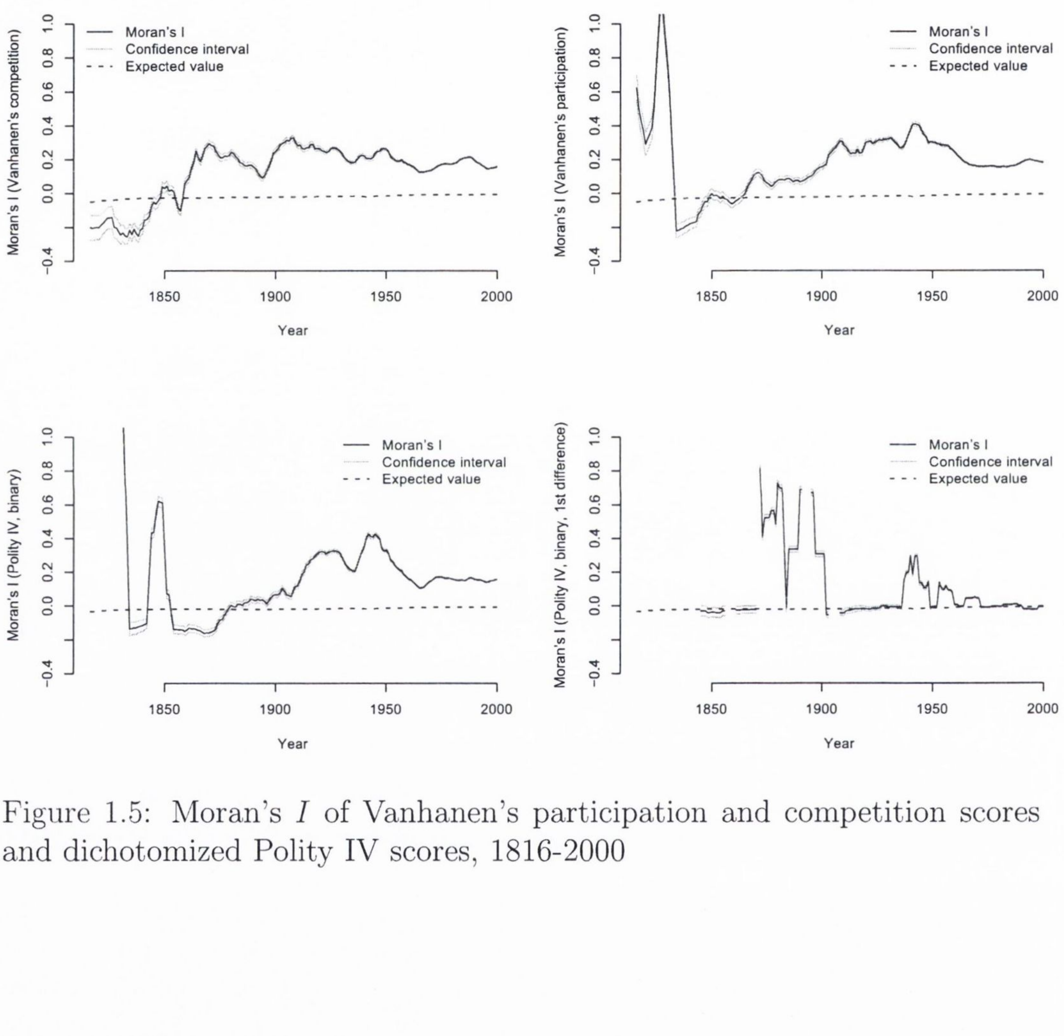

Figure 1.5: Moran's I of Vanhanen's participation and competition scores and dichotomized Polity IV scores, 1816-2000

In addition to the scores by Vanhanen, we can clearly see similar patterns of spatial clustering on the Polity IV scores. The measure becomes significantly different from the expected value given the number of countries in the world and consistently positive as Vanhanen's scores indicate, adding further confidence in the result. The last plot shows the level of spatial clustering in the Polity IV scores, but as first differences. This means that we are looking at the extent to which changes in these scores are geographically clustered, as opposed to the scores themselves. See for a further discussion on this idea the next section.

As opposed to focusing on the global level of clustering, one can measure diffusion effects at a local level, studying the extent to which a country is surrounded by similar countries, or studying the extent to which surrounding regimes explain the type of regime in a particular country. Kopstein and Reilly (2000) show how neighbor effects are stronger explanations of regime type in Eastern Europe in the nineties than are more common explanations as pre-communist experience with bureaucracies, the result of the first democratic elections, and the economic situation in the country. Starr also concludes from his study of neighbor effects that countries that underwent 'treatment', that is, where a neighbor showed a transition towards democracy, have a higher probability of experiencing such a transition themselves (Starr 1991: 373-377). One could also formulate a similar test in terms of a regression context: "Since 1815, the probability that a randomly chosen country will be a democracy is about 0.75 if the majority of its neighbors are democracies, but only 0.14 if the majority of its neighbors are non-democracies" (Gleditsch and Ward 2006: 916). This observation has been confirmed repeatedly in quantitative empirical research (Starr 1991; Ward et al. 1996; O'Loughlin et al. 1998; Ward and Gleditsch 1998; Gleditsch and Ward 1997, 2000, 2006;

Brinks and Coppedge 2006; Gleditsch 2002; Elkink 2003; Doorenspleet 2001, 2004; Wejnert 2005; Fordham and Asal 2007).

1.2.3 Spatio-temporal clustering

> Poland - 10 years, Hungary - 10 months, East Germany - 10 weeks, Czechoslovakia - 10 days (a banner in Prague) (Kuran 1995: 274)

Spatio-temporal clustering is the pattern that in effect generates both patterns described above, simultaneously. If it is true that democracy can spread like a contagious disease, that one country which democratizes increases the chances for a democratization in a neighboring country, then we would expect to see both waves of democratization and spatial clustering. The waves of democratization would often be very localized - the world-wide number of democracies increases because a particular group of countries makes a transition - and the spatial clustering would be temporal in nature. Not based on survival and unlikely to be based on the slow process of modernization, but instead based on direct influence of one transition to democracy on another.

The quote above refers to the most typical example of this process and probably the one that triggered a lot of the research in waves and diffusion of democracy, namely the many transitions to democracy in Eastern Europe in the early 1990s. The transitions to democracy followed one another sequentially, faster and faster. Various quite distinct theoretical explanations can account for this pattern, which are presented in detail in the next chapter, but all share that transitions to democracy are sensitive to changes in the international context of a country. Successful transitions in one country function as an example to political leaders in another (Starr 1991); communication between individual citizens between neighboring countries change their perceptions about democracy (Rosenau 1988: 359); an international regional cooperation offers great benefits to membership, but requires a democratic constitution for membership (Levitsky and Way 2003); an occupying foreign country leaves an entire region to itself - all explanations that refer to the international context in explaining regime changes.

Empirically speaking, we are thus facing the task of demonstrating that the chances of transitions to democracy are affected by transitions in neighboring countries - or, in fact, by the sheer presence of democracy in neighboring countries. To distinguish between the three patterns of clustering, we will study an empirical model whereby the probability of democratization depends on the level of democracy in the previous time period, on the global level of democracy in the previous time period, and on the neighboring level of democracy. The statistical approach taken is in line with Gleditsch and Ward (2006), who apply a first-order Markov chain model. The concept of this model is relatively straightforward. Figure 1.6 provides a graphical visualization of the model. The arrows can be interpreted as transition probabilities from $t-1$ to t. Thus, the probability of a democracy to survive in year t is P_1, while it will collapse into an autocracy with probability P_2. Given the laws of probability, $P_1 + P_2 = 1$. Similarly, for an autocracy to survive at time t the probability is P_4 and there is a P_3 chance that the autocracy will make a transition to democracy, $P_3 + P_4 = 1$. Since P_2 and P_4 are entirely determined by P_1 and P_3, respectively, we will only estimate the probabilities P_1 and P_3. The latter can be interpreted as the probability that a particular country is a democracy at time t, given that it is a democracy or an autocracy at time $t-1$, respectively. In other words, the two probabilities we will estimate is the probability of survival of a democratic

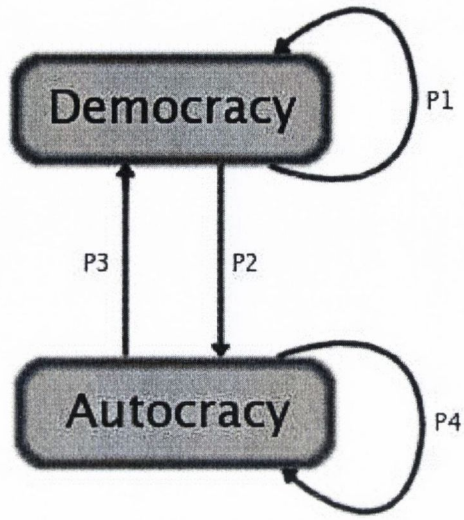

Figure 1.6: First-order Markov chain model of democratization

regime and that of a transition to democracy. In statistical terms, this can easily be implemented in one statistical model by interacting all independent variables with the lagged (dichotomous) dependent variable.

If we consider $\mu_t(\theta)$ to be the systematic part of the model (King 1998), parameterized by θ, we would be estimating the following model:

$$\mu_t(\theta) = \theta_0 + \theta_1 G y_{t-1} + \theta_2 \tilde{W} y_{t-1} + \theta_3 \tilde{W} \Delta y_{t-1} + X_{t-1}\theta_4, \qquad (1.8)$$

where y_t measures the level of democracy at time t; G is a matrix with zeros on the diagonal and $\frac{1}{N_t-1}\sum_{j,j\neq i} C_{jt}$ in all other cells, where N_t is the total number of countries at time t; $\Delta y_t = y_t - y_{t-1}$ is the first difference of the level of democracy; X_t is a set of control variables; and the θ's are regression coefficients, with θ_4 being a vector rather than a scalar.[9] In this model,

[9]Note that $\tilde{W} C_t$ is the equivalent of taking the average competition score of all adjacent

θ_1 will indicate the effect of global waves of democratization, of temporal clustering; θ_2 the effect of spatial clustering; and θ_3 the effect of local waves of democratization, thus spatio-temporal clustering.

We will look separately at models using a dichotomized version of the Polity IV democracy score and the equivalent models using the competition and participation scores of Vanhanen. The first is implemented using the Markov chain model described above, the second with a more straightforward linear model. The Markov chain model is implemented as follows:

$$Pr(y_{it} = 1) = F[\mu_{it}(\alpha) + y_{i,t-1}\mu_{it}(\beta) + \gamma_i], \tag{1.9}$$

where $\gamma_i \sim N(0, \sigma_{\gamma_i})$ is the random effect to account for country-specific idiosyncracies. Here α contains the coefficients explaining the transitions from autocracy to democracy, $\alpha + \beta$ the coefficients explaining the survival probabilities of democracy. The F function can be any function mapping a linear prediction to a probability between 0 and 1, typically the cumulative normal distribution (probit model) or $F(x) = e^x/(1 + e^x)$ (logistic model). The latter is used in the analysis presented here. For the linear model, the approach is more straightforward and results in $y_{it} = \phi y_{i,t-1} + \mu_{it}(\beta) + \gamma_i + \varepsilon_{it}$, where $\varepsilon \sim N(0, \sigma^2)$ and γ_i as in the previous specification.

The Markov chain model can be seen as a replication attempt of the model presented by Gleditsch and Ward (2006). Their model is based on the same principle and the control variables in X are taken from their replication data set.[10] To control for variation in domestic factors not covered by the few variables in X, country random effects were added to the models presented

countries for each country, since each row of $\tilde{W}$ adds up to one.

[10]The Gleditsch and Ward (2006) replication archive can be downloaded from http://privatewww.essex.ac.uk/ ksg/data/gwio2006.zip.

here. Gleditsch and Ward do not include such effects in their model, which explains the fairly substantial differences between the results in their work and those presented here. The control variables are the energy consumption per capita, which closely relates to the level of economic development, but for which data is available for earlier years than for more common economic indicators such as gross domestic product per capita (Gleditsch and Ward 2006: 927); a dummy for whether a country is involved in a civil war; and the number of years of continuous peace on the territory of the country, rescaled to centuries.

Table 1.2 presents the various random effects models, the first four columns covering the Markov chain model using Polity IV data and the latter four columns separate models for the two dimensions of democracy. Overall, most diffusion effects, whether spatial ($\tilde{W}y_{t-1}$), temporal (Gy_{t-1}), or spatio-temporal ($\tilde{W}\Delta y_{t-1}$), show significant results. There clearly are significant patterns that are worth explaining and the remainder of this thesis will set out to do so. The fact that these results are statistically significant, in line with the findings in a large number of other articles, does not imply that these effects are very strong or the main explanation of democratization. Domestic factors still appear to have a bigger impact than does the international context and the effects of diffusion are relatively marginal. Assuming that there are no transitions in the immediate neighborhood, that about 40% of the world countries are democratic, and ignoring the control variables (using the first column of Table 1.2), the coefficient suggests an increase from a 1% to an 8% chance of a transition to democracy. We can thus say that the international context matters and is worthy of further study, but not that this is the only or most crucial factor in transitions to or the survival of democracies.

	Polity IV, dichotomized				Vanhanen			
	Aut. (α)	Dem. ($\alpha + \beta$)	Aut. (α)	Dem. ($\alpha + \beta$)	Competition		Participation	
(Intercept)	-6.48 *	3.36 *	-5.51 *	4.79 *	-0.25	0.62	-0.06	1.00 *
	(0.39)	(0.47)	(0.51)	(0.58)	(0.36)	(0.53)	(0.18)	(0.31)
y_{t-1}					0.87 *	0.87 *	0.90 *	0.90 *
					(0.01)	(0.01)	(0.00)	(0.01)
Gy_{t-1} (temporal)	4.82 *	-1.77	2.91 *	-1.82	0.07 *	0.06 *	0.05 *	0.03 *
	(1.12)	(1.45)	(1.40)	(1.76)	(0.01)	(0.02)	(0.01)	(0.01)
$\tilde{W}y_{t-1}$ (spatial)	2.11 *	2.01 *	1.60 *	0.62	0.06 *	0.05 *	0.06 *	0.06 *
	(0.35)	(0.44)	(0.35)	(0.48)	(0.01)	(0.01)	(0.01)	(0.01)
$\tilde{W}\Delta y_{t-1}$ (spatio-temporal)	1.84 *	-1.71	1.82 *	-0.97	0.01	0.01	0.01	0.02
	(0.70)	(1.35)	(0.69)	(1.47)	(0.02)	(0.02)	(0.02)	(0.02)
Log of energy consumption			0.22 *	0.79 *		0.55 *		0.36 *
per capita			(0.10)	(0.13)		(0.10)		(0.07)
Civil war			-0.11	-0.46		0.17		0.08
			(0.47)	(0.47)		(0.42)		(0.31)
Centuries of peace			0.75 *	-0.04		1.19 *		-0.19
at territory			(0.37)	(0.54)		(0.47)		(0.33)
N	9076		7094		9105	7114	9105	7114

Table 1.2: Democratic transition and consolidation explained by spatial lags. The first two columns are the results of one logistic random effects model, without control variables; the second two columns represent the same model, with control variables added; the last four columns contain four different models, all linear random effects models. The first and third columns contains all autocracies at $t-1$ (hence probability of transition to democracy) and the second and fourth all democracies at $t-1$ (hence probability of survival as democracy).

When taking the dichotomized Polity IV measure as the dependent variable the effect of temporal clustering (Gy_{t-1}) on transitions to democracy is particularly strong. It is slightly reduced by the addition of some important domestic factors, but the effect is still clear and significant. On the survival of democracy, however, the effect entirely disappears to an insignificant and even negative one. The number of democracies in the world does not have an effect on the chance of survival for a democratic regime, but it does substantially increase the probability of a transition to democracy. Looking at the participation and competition dimensions separately and studying democracy as a continuous variable rather than a black-and-white one, the international presence of democracy has a positive effect on both the level of competition and the level of participation in a country.

The effect of a presence of democracy in the immediate neighbourhood of a country ($\tilde{W}y_{t-1}$) has a very similar effect on the chances for transitions to democracy, albeit slightly weaker. It should be noted, however, that the independent variable here takes a wider range of values: some countries are certainly all surrounded by democracies, while at no point in time was more than 50% of the countries in the world democratic. The effect size for spatial clustering is thus remarkably similar to that of temporal clustering. Democracies also have a higher probability of survival as democracy in the presence of other democracies in the immediate neighbourhood, although this effect vanishes when taking economic and war history into account. On the levels of competition and participation, the effect is positive as well - high levels of competition in the neighbourhood leads to high levels of competition in the country itself, and equivalently for levels of participation.

Finally, the effect of spatio-temporal clustering ($\tilde{W}\Delta y_{t-1}$) on transitions to and the survival of democracy follows a pattern identical to temporal clustering. The more countries in the immediate neighbourhood make a transition to democracy in the previous year, the more likely a transition takes place in the country itself, but on survival there is no clear effect. Other countries in the immediate neighbourhood making a transition to democracy has no effect on the survival of democracy itself. The, statistically insignificant, effect is even negative, suggesting that more transitions in the neighbourhood destabilize not only autocratic regimes, but also democratic ones. When we study the levels of competition and participation in a country, changes in neighbouring countries have no effect at all. The spatio-temporal clustering patterns disappear when disentangling different aspects of democracy and when taking also gradual changes into account.

A number of features of these regression results stand out. The first striking result is that the patterns for transitions from autocracies to democracies show much clearer clustering patterns than do the survival probabilities of existing democracies. This result would appear to contradict the explanation by Cederman and Gleditsch (2004) that the geographical clustering we observe is related to the increased chances of survival for new democracies when surrounded by other democratic regimes. It is not the survival, but the transitions themselves that show the spatial and temporal clustering patterns. The second feature that stands out from these results is that when democracy is measured along the two dimensions of democracy separately, using Vanhanen's proxy measures, instead of using a dichotomized version of the Polity IV data as in Gleditsch and Ward (2006) and Fordham and Asal (2007), then the spatio-temporal pattern disappears. The level of competition or participation is dependent on the international and local context, but less on transitions in the local neighborhood.

1.3 Conclusion

Before embarking on a study to explain the world-wide clustering of democratic regimes one has to start by defining what is meant by democracy and by clustering and by demonstrating that this clustering does indeed occur. This chapter has done just that. We established that democracy can, for the purposes of a large-N empirical study, best be defined on a minimalist, procedural basis, and that this is often done with the work of Schumpeter (1976) or Dahl (1971) in mind. The latter leads to a conception of democracy which makes a distinction between the amount of competition among the political elites and the amount of participation among the general population in influencing this competition. Using the empirical data sets of Vanhanen (1997) and Marshall and Jaggers (2002) we established that we can indeed clearly observe international patterns of clustering of democracy.

For the purposes of the empirical demonstration and also as a guide in classifying the various different explanations of democratic clustering discussed in the next chapter, we made a distinction between three different types of clustering. Temporal clustering refers to the way democratic transitions are clustered in time. During some periods in time the number of transitions either to or from democracy in the world is much larger than in other periods. Waves of democracy are clearly visible - four waves in the past two centuries can be identified (Doorenspleet 2001). Spatial clustering refers to the geographical clustering patterns, the groups of democracies we see when we publish a map of political regimes at any point in time. Particular regions of the Earth are filled with democratic regimes, while in other areas they are very rare (Starr 1991). Finally, spatio-temporal clustering refers to the occurrance of democratic transitions, or their reversal, in one area in a short time span. One regime tumbles and the other neighboring regime follows shortly after. The most typical example of this is the collapse of the many communist regimes in Eastern Europe, one after the other, in the nineties (Kuran 1995).

Both in several published empirical analyses of democratization and in a brief empirical analysis presented above it is confirmed that the geographical clustering of political regimes happens to a significant effect, and cannot simply be explained away by the geographical clustering of other domestic factors that typically explain democratization. Although it is clear that these patterns exist, there are still many ways by which one could explain such patterns. The next chapter will present a brief overview of these various possible explanations, after which the remainder of this thesis will focus on a model of democratic clustering through the international diffusion of individual attitudes.

CHAPTER 2

Explanations of democratic diffusion

The theoretical perspective of this thesis is a model of the diffusion of attitudes among individuals, and its effect on processes of democratization. There is thus a strong bottom-up view of democratization, where it is mass attitudes more than elite strategy that determines the outcome of political transitions. It is assumed that elites cannot survive without sufficient support from the masses - at least when these masses are themselves sufficiently aware of the public opinion and dare to protest against the regime. The next chapters will get into more detail into the theoretical foundations of this perspective.

This bottom-up approach is not the only possible perspective on the diffusion of democracy. Other theories have either been presented as an explanation of the diffusion of democracy, or can at least be assumed to have an effect even when not explicitly put forward in this field of study. To put the model presented in this thesis in proper perspective it will be useful to first discuss these existing theories and to contrast the approach taken here.

It should be pointed out that although this thesis will present a model that is different from most of the explanations presented below, there is no claim that this model necessarily contradicts any of the below perspectives. Regime transitions are incredibly complex processes where many factors - international, domestic, cultural, historical, economic, strategic, and many other - play a role. No theory will ever be all encompassing in describing a phenomenon as multidimensional as this. We can only hope to identify and understand some of those mechanisms that form part of the explanation. The attempt of this thesis is to find one of those mechanisms - that of a diffusion of attitudes and a cascading revolution through public protest - which is proposed as one possible explanation of the clustering of democracies we described in the previous chapter. If we can establish through simulations that this model is a possible explanation of these patterns, and through empirical studies that this model is plausible as a description of actual patterns of behavior, we are one step further towards understanding democratic transitions. This thesis contains the first of those two steps.

To give some structure to the discussion below, it is useful to get an idea of the broad distinctions between the various theories. The literature on democratic diffusion is based on two different disciplines in political science, and they have different basic assumptions. The structure for this chapter will be a mixture of the common classification of theories of democratization and the classification of the literature on policy diffusion as presented by Braun and Gilardi (2006). The categories of theories used below are in no sense mutually exclusive and unlikely to be exhaustive. Categorization simply helps to structure the debate and make it more clear - it does not replace the finer details of the various theories and many theoretical works in the area have a nuanced perspective combining aspects of various of the categories below.

The literature on the diffusion of democracy is embedded in two distinct disciplines within political science. On the one hand it is clearly related to the international relations literature and most of the authors on the subject are indeed specialists in that area. It is the international relations factor that distinguishes the literature on democratic diffusion - in a broad sense - from the mainstream literature on democratization. Within the field of international relations, authors have a very specific perspective on democratization and focus on the position of democratizing countries within the international system. They are usually less concerned with explaining domestic politics, even though this is not entirely ignored, and have thus almost by default a more elitist perspective on democratization. Processes of democratization are put into the context of international support for political development or in the context of the typical subject of international relations: war. Kant established the thesis that democracies do not fight each other - sometimes considered one of the few actual laws in political science - which has recently been related to studies of democratization (Gleditsch 2002).

The second area is of course the large literature in comparative politics on transitions to democracy and the consolidation of new democracies. The diffusion of democracy is a logical hybrid between the two, since it concerns processes of democratization in an international context. The literature on democratization and democratic consolidation - two very different processes, but often mentioned as if different sides of the same story - has exploded in recent decades and is too large to summarize in this chapter. Only some parts of this literature shed light directly on the international factors to democratization and the diffusion of democracy and only those that do will be considered here.

In the literature on democratization, one distinction is often emphasized, which is that between theories that focus primarily on elite behavior, and those that focus on public opinion and revolutions of the masses. The debate among historians and philosophers whether it is the leaders or the followers that are crucial in explaining developments in human society is as old as those disciplines (Tolstoy 1993) and is reflected in this debate among comparative political scientists.[1] Although this thesis does not attempt to take a strong position on this debate, when locating the model in the wider literature on democratization, it clearly relates to mass-based models of democratization. The model is a bottom-up description of revolutionary behavior and practically assumes away all elite level strategic behavior. Elites are assumed to be dependent on publicly exposed support from the masses.

In this chapter possible explanations of the diffusion and clustering of political regimes, mostly from the perspective of democratization, will be discussed first from the perspective of international relations and peace studies. Subsequently, the focus will shift to similarly elitist theories of diffusion that are more directed to domestic policies and based in the comparative politics literature. Finally mass based, bottom-up theories of democratic diffusion will be discussed, making for a smooth transition to the discussion of the key model of this thesis.

[1]In the agent-based modeling literature there is a strong emphasis on bottom-up explanations of social phenomena as opposed to top-down explanations. This distinction is related to, but not the same, as the one between elite and mass based theories of social transitions. There might thus be a slight bias in perspective induced by the method of research. Ultimately, the distinction between the two perspectives appears to be a discussion of whether the glass is half full or half empty and is a matter of taste rather than one being more correct than the other. Elites matter, as does public opinion.

2.1 Domestic factors in democratization

Because the prime aim of this thesis is to provide a possible theoretical understanding of the international clustering of democratic regimes, the focus will be entirely on the international factors that play a role in processes of democratization. Since the spatial and temporal clustering patterns of political regimes are international phenomena, only international factors in democratization can explain these patterns. If processes of democratization in different countries were independent of each other, no such patterns would arise.[2] For readers familiar with the literature on democratization, this chapter might therefore read as a somewhat awkward overview of theories - most common explanations of democratization are hardly mentioned below. The usual focus in studies of democratization is, after all, on domestic factors - economic development, class structure, colonial past, development of civil society, etcetera.

The development or modernization thesis argues that the chances for democratization are determined strongly by the level of economic development in a country. A certain minimum level of development is necessary for popular protest to come about and to be effective. Anti-regime mobilization is costly and individuals who are primarily concerned with their own food, health, and safety have little opportunity to influence the political regime. This original idea is generally related to Lipset (1959) and has been confirmed in many subsequent studies (Doorenspleet 2004: 312). As a slight alteration to this hypothesis, it is often argued that in the very rich countries, for example the Middle East oil states, democratization is unlikely, since the au-

[2]The exception here would be a pattern of spurious clustering, whereby some domestic factor that influences processes of democratization clusters, while the democratization processes themselves are internationally independent of each other. This is discussed in §2.3.4.

tocratic regime can provide all the goods and services to keep the population satisfied. There would thus be a curvilinear relationship between economic development and democracy. Although the economic development thesis is a purely domestic one, it is also the most likely candidate for a spurious explanation of democratic clustering (see §2.3.4), whereby economic development itself is geographically clustered (see, e.g., Hak 1993) and in turn explains the level of democratization in a particular cluster of countries.

Closely related to the modernization thesis is the idea that the class structure of a country is relevant in the explanation of transitions to or the consolidation of democracy. Thanks to its need for peace and freedom to pursue its economic interests and its ability to moderate conflict, Lipset (1959) argues that it is primarily the size of the middle class in society that matters for democratization. Others argued that it is the size of the working class that matters, since they are the most interested in extended voting rights: economic development would increase the political power of the working class, which could in turn push for more democratic rights (Doorenspleet 2004: 315). Except in its relation to economic development, class structure is presumably a purely domestic factor, which cannot explain the international and temporal world-wide clustering of democratic regimes.

Another theory closely related to the modernization thesis is one that focuses on the position of a country in the world system. The argument here is that the middle class in countries that are in the poorer (semi)periphery of the world system play a different role in domestic politics than those in the core countries. The idea is that the elites in the countries in the periphery form an alliance with the elites in core countries to provide them with the resources available in the periphery, which reduces the political influence of the lower and middle classes in those countries (Doorenspleet 2004: 316).

Here democratization is thus clearly linked to an international factor, namely the position in the world system of countries. The international pattern here is one of economic dependency, however, rather than geographical location. Although the theory is international in nature, it is unlikely to help in the understanding of the international clustering patterns of democratization - the links through which this mechanism operate are not geographical.

Besides these theories whereby it is the overal structure of a society that would determine democratization, there are various actor-oriented theories of democratization (Doorenspleet 2004: 310). In these approaches, it is the make-up of the political elite or the actions of members of the political elite that are the prime determinants of transitions to and consolidation of democracy. For example, Burton and Higley (1987) discuss the necessity for political elites to establish pacts before democratization can take place. Different fractions of the elite need to compromise for sufficient political stability to be created for democracy to be established.

Acemoglu and Robinson (2001) provide an actor-oriented theory of democratization that supports the correlations found between economic development and democratization. Their model focuses on the redistribution through the state of resources between the poor and the rich. This redistribution occurs primarily through tax rates, which are set by the government. In an autocratic regime it is the rich that set the rate, while in a democracy it is the median voter, the poor. The rich can avoid a revolution by establishing a democracy, which solidifies their promise of future fair redistribution of resources. In a very unequal society, the costs of such a fair redistribution will be too high for the elite, while in a more equal society, the fear for revolution will outweigh these costs and democracy will be established. This approach is compatible with various internationally focused theories of democratization, including the one presented in this thesis. The level of inequality could be determined by international clustering patterns in economic development (§2.3.4), the elites might learn from elites in similar countries to determine their chances of success (§2.3.1), or the poor might learn about their chances from the poor in neighbouring countries (this thesis).

2.2 Realpolitik, or the international relations approach

From the international relations perspective I will discuss a number of theories that relate the diffusion of democracy or of political regimes to the military or strategic circumstances of countries in an international context. The connection between democratization and international conflict is a fairly recent one in the academic literature and has only received serious attention in the last decade or so. Most of this work relates to the idea of a democratic peace and the extent to which international conflict increases or reduces chances of democratization, but a somewhat exceptional approach has been put forward in the work of Cederman and Gleditsch (2004), which does not attempt to explain transitions to democracy as such in a context of international conflict, but the survival of different types of regimes. The first theory to be discussed is a much older one, however, and one originating from outside academia, but certainly one that is closely related to the idea of a diffusion of political regimes, namely the famous domino theory.

2.2.1 Domino theory

At a press conference on April 7 of 1954, the U.S. President Dwight D. Eisenhower commented on a question about the strategic importance of Indochina, contemporary Vietnam, to the 'free world': "[Y]ou have the possibility that many human beings pass under a dictatorship that is inimical to the free world. (...) [Y]ou have broader considerations that might follow what you would call the 'falling domino' principle. You have a row of dominoes set up, you knock over the first one, and what will happen to the last one is the certainty that it will go over very quickly. So you could have a beginning of a disintegration that would have the most profound influences" (*The President's News Conference of April 7, 1954* 1954). This statement was the invention of what became widely known as the domino theory, the idea that if communism is not stopped for example in Vietnam, there is a serious risk that further countries in the region will fall to communism as well. That Eisenhower really had this geographical aspect in mind can be seen in a remark he made a few seconds later in the same press conference: "But when we come to the possible sequence of events, the loss of Indochina, of Burma, of Thailand, of the Peninsula, and Indonesia following, now you begin to talk about areas that not only multiply the disadvantages that you would suffer through loss of materials, sources of materials, but now you are talking really about millions and millions and millions of people. Finally, the geographical position achieved thereby does many things. It turns the so-called island defensive chain of Japan, Formosa, of the Philippines and to the southward; it moves in to threaten Australia and New Zealand" (*The President's News Conference of April 7, 1954* 1954).

The domino theory acquired prominence in the debates on the expansion of communist regimes after the Second World War, but is similarly applied to the spread of other types of regimes. The democratization of Eastern and Central Europe followed a pattern very similar to that which Eisenhower feared in 1954, albeit in the opposite direction. As dominoes the different regimes tumbled one after the order, ever faster. This idea of democratic dominoes (Starr 1991) seems to be a prominent feature in contemporary rhetoric in relation to processes of democratization, especially in the more journalistic literature. Developments in Serbia, where Milosevic's regime was brought down, in Ukraine, where Kuchma's preferred successor did not manage to win the elections, and Georgia, where Shevardnadze was forced to resign after protests following his rigged re-election, are all seen as related, as one group of protesters following the example, and advice, of other previous protesters in geographically proximate, although not bordering, countries. When the United States attempts to argue its case for attacking Iraq in the second Iraq war, it submits that by placing a beacon of democracy in the Middle East, neighboring populations will soon be inspired and regimes in the region will find it more and more difficult to maintain their autocratic regimes (Reynolds 2003). Elections for the Palestinian Authority and the retraction of Syrian troops from Lebanon were quickly pointed out as early signs of such developments. The domino theory has even been applied to the spread of fundamentalist Islam (Staten 1996).

This theory of falling dominoes applies to different types of regimes; it has been pointed out in relation to both the spread of communist regimes, as well as that of liberal democratic ones. This finding is important to keep in mind. Most of the research in this area operates under the nomer of democratic diffusion, thus emphasizing only one of the two mechanisms, and is placed in the context of studies into democratic transition and consolidation. The clustering that is observed on a world-wide scale, however, can be explained by the geographical diffusion of democratic regimes, but equally well by the clustering of other regime types. This text will, as most of the relevant literature, often fall in the trap of focusing implicitly on the spread of democratic regimes more than other similar contagion processes, simply because this is what happens in the literature this text fits into and refers to, but it should be kept in mind that this focus might well be artificial and misleading, simply caused by the rising optimism about democratization since the nineties.

Strategically countries have a very special interest in their geographic neighbors that is substantially different from their interest in other countries, for the simple reason that it is often much easier to move military units to a neighboring country than it is to move them to a farther country, where one is more dependent on long distance transport and a solid navy or air force. Also refugee flows are usually felt most by neighboring countries of regions in turmoil. For this reason, international cooperation is much more often encountered between neighboring countries or many countries within a specific region than it is between other combinations of countries. Regional cooperation leads to international pressure among political elites to implement similar regimes in neighboring countries, ranging from policy details like water management to the fundamental organization of the political system.

Although the domino theory has a clear military-strategic perspective on the geographical clustering of regimes, there is also an ideological under-tone to the theory. One might wonder to what extent Eisenhower was worried about military intervention by communist states in neighboring countries or to what extent he was more worried about the spread of the communist idea. Would the people succumb to the communist ideas through propaganda by neighboring countries or would they be forced into a communist regime through violence. The former would be very close to the fundamental model of this thesis, where democracy promotion is a key mechanism. The model is partly driven by the fact that democratic regimes support democracy in foreign provinces close to the democratic capital, which is quite similar to propaganda by communist organizations across borders. The more military perspective, however, is unrelated to the model presented here.

In terms of global patterns of democratization - the growth, temporal and geographic clustering of political regimes - the two explanations might be difficult to distinguish. Both lead to similar types of clustering, geographically and most likely also temporally. In terms of accompanying developments with the transitions, perhaps a clearer distinction can be made, where one version goes combined with a lot more violence than the other. An interesting control variable in an empirical validation of the model presented in this thesis would thus be the presence of domestic and cross-border violent interchanges, in particular interacted with the type of political regime of the opposing country. Perhaps the picture will become murkier when domestic violence is taken into account, as it might be difficult to distinguish aggressive actions by citizens inspired by neighbors from similar actions by foreign individuals. For example, one could imagine revolutionaries from one country to move over to help the democratic opposition in a neighboring country, through violent means (which would correspond to the domino theory), as opposed to similar revolutionaries constraining themselves to propaganda towards this neighboring opposition (which would correspond to the model of this thesis).

2.2.2 Democratization by force and decolonization

When discussing democratization in an international context one very straightforward explanation can not be ignored: the direct implementation of democratic regimes in occupied territory by Western actors. The two most well-known examples are West-Germany and Japan, which were occupied at the end of the Second World War and where democratic regimes were established under the supervision of the allied forces. These regimes have survived remarkably well during the decades since and have become beacons of peace and democracy at the international stage. The (re-)establishment of democracy in most West-European countries that were occupied by the Germans can be similarly considered to be a democratization process induced by military occupation (by the Allied Forces) (Ethier 2003: 100). The much more recent attempt to establish democracies in Afghanistan and Iraq, where it is far less certain whether it will be similarly successful, can also be counted in this category. Another more recent example is Bosnia-Herzegovina, which has been ruled for years by the Office of the High Representative (Jacoby 2006: 630). The general pattern here is that democratic countries occupy territories abroad and then initiate processes of democratization to affect the political regime after the occupation.

> "Warfare has usually been associated with boundary changes, annexations, limitations on the autonomy of the defeated states, the levying of tribute, reparations and the like. In (...) the last centuries it has resulted in changes in internal political organization. Thus the advancing Napoleonic armies carried revolutionary constitutions and legal codes along with them. And the armies of the 'Holy Alliance' re-established legitimate constitutions as they advanced. The (...) Allied occupation of Germany after the First

World War was intended not only to guarantee reparations, but to ensure an acceptable German regime. During and after the Second World War the advances of Nazi armies, Soviet armies and Western armies were all associated with political changes of the most significant sort. The eastern European countries were sovietized; the Nazi-occupied territories were 'nazified'; in the post-war period the German and Japanese constitutions were in considerable part drafted by American authorities." (Almond 1989: 248)

Decolonization can, to some extent, be considered to be in the same category (Whitehead 1996).[3] The most striking example is the Commonwealth of all the former British colonies, which all exhibit strong democratic traditions ever since the decolonization in the forties and fifties. The heritage of other colonial empires is generally less positive and certainly not all processes of decolonization are combined with transitions towards democracy, especially not in the cases where the colonial power was more resistant to change, as in for example the former Dutch colonies and the former French ones (Wejnert 2005: 56).

Democratization by force or through decolonization will generally have a stronger effect on temporal rather than spatial clustering.[4] Germany and Japan were individual cases where the geographical location had little effect on their democratization, while the many colonies that survived as independent democratic states are far spread and not often geographically contiguous. Most contiguous regions during the colonial empires were submerged

[3]A similar categorization of democratization by external pressure has been labeled control (Whitehead 1996; Ethier 2003) or substitution (Jacoby 2006).

[4]See Strang (1991) for an analysis of decolonization as a temporal and spatial diffusion of sovereignty. In fact, he finds that decolonizations in the same geographic region have a stronger effect than decolonizations in the same colonial empire (Strang 1991: 344).

under one colonial authority and as a result of path dependency later established as single entities, despite their heterogeneous pre-colonial past. The many political units of India became part of one country after the decolonization process, with the exception of the region of Pakistan and Bangladesh. The many prince- and chiefdoms of Indonesia became one political unit after independence from the Dutch. The democratization of those units did thus not particularly contribute to a clustering pattern of democracies, since to the extent that they were geographically contiguous, they were merged into one political unit.

In a temporal sense, however, this clustering can certainly be observed. Many processes of decolonization occurred in a short wave after the Second World War, thus in temporally contiguous time periods. The peak of democratization processes in the 1950s, directly after the war, can to a large extent be explained by the sudden change in the *Zeitgeist*, as a result of the increased prominence of the United States on the world stage, towards colonies. Whereas before the war colonies were a commonly accepted feature of the international configuration of states, after the war it came to be seen as unethical and unacceptable, and the old empires came quickly under far more pressure to leave their colonies. The involvement of military personnel from the colonies in both the world wars also made it more difficult to deny them their right of self-determination.

The idea that decolonization or occupation affect democratization and affect the waves of democracy can hardly be denied. This is one mechanism that at least co-exists with other processes of diffusion, rather than forming an alternative explanation for the same patterns. Any empirical validation of a diffusion process of democratization will thus have to control for those cases where the cause is obviously one of force. For as far as democratization is concerned, these would include the democratization of Western Europe after the Second World War, the democratization of the former British empire, and perhaps even the democratization of Eastern Europe in the early 1990, which can be seen as a process of decolonization (McFaul 2002: fn. 4). The latter is debatable, however.[5]

Although this perspective is very different from the model presented here and has to co-exist, there might still be some overlap between the two. For example, one could see the effect of decolonization on democratization as an effect of the propaganda of the colonial power towards the colonies. The process of decolonization in these cases can thus be seen as a somewhat special case of democracy promotion in line with our model. The idea here would be that the colonial powers affected the attitudes towards democracy among the population which has an effect on the popular support for democracy in the newly decolonized countries. Or, alternatively, the initial values of the attitude towards democracy is high if these are considered to be new countries. The agent-based model presented below does not take into account the establishment of new countries - the configuration of countries is assumed to be constant - and democracy promotion only appears between neighboring countries - while colonies generally do not border on the colonial powers - but the general idea is not necessarily incompatible.

2.2.3 Clustering for survival

Methodologically, the study into democratic diffusion that comes closest to the one presented here is the agent-based model by Cederman and Gleditsch (2004). Their model embeds the diffusion of democracy in the international

[5]The key difference here would perhaps be that while the colonial powers were generally democracies, this was not the case with the Soviet Union. The democratization after decolonization can thus not be seen as learned behavior from the former colonial capital.

relations and war studies literatures. Their model is an extension of the model on war and state formation by Cederman (1997, 2002). Based on empirical observations and existing theory on the "democratic peace", they explicitly assume a priori that countries have a higher probability of democratizing when surrounded by democracies and that countries that are democratic do not attach each other. The geographic clustering that is observed in this model is thus of little surprise. In most of their base simulation runs, however, they find that very few democracies survive on the long term.

To arrive at trends more similar to those observed in the empirical data, they add the assumption that contiguous democracies assist each other when attacked by an outsider. Geographic clusters of democratic countries, again by design, thus operate in a similar fashion to large single countries, at least in terms of their military defense. Under this configuration, democracies have a much stronger chance of survival when adjacent to other democracies. Since these clusters survive better, they also increase the chances for democratization in neighboring countries and thus the end result is an overall increase in the number of democracies and a strong clustering of democratic regimes. Their model thus suggests that adaptation to survive a hostile environment is an important possible explanation for the geographic clustering of democratic regimes.

It should be pointed out that although their cooperation mechanism explains the survival of clusters of democracy, which in turn affect neighboring countries towards democratization, this effect on neighboring countries is not actually explained by the model, but rather assumed a priori. The diffusion mechanism of democracy in their base model does not take off because democracies do not survive long enough to affect their neighbors and the increased changes of survival thanks to cooperation resolves this problem.

The fact that the diffusion mechanism exists in the first place, however, is assumed rather than modeled. The model presented in this research focuses more on the latter mechanism and is thus not incompatible with the model of Cederman and Gleditsch. The two mechanisms would reinforce each other and might thus well co-exist.

The survival theory is probably the strongest possible explanation that explains geographic clustering without assuming spatio-temporal or temporal mechanisms.[6] Cederman and Gleditsch (2004) assume the latter for their model, but even without these assumptions and an entirely domestic view of democratization, their model would likely lead to a clustered pattern of democracy, albeit at a slower pace than their current model. If one assumes forward-looking political leaders, however, one would assume that small countries might also democratize in order to be more similar to neighboring countries and to be able to count on their protection in the face of larger neighbors. In this sense, the fact that neighboring countries are democratic would affect the pay-offs of turning democratic oneself, one of Simmons and Elkins (2004)'s two main requirements for a policy diffusion to occur.

Although the two models of geographic clustering of democracies - through diffusion and through survival - are not incompatible, it will still be necessary to distinguish them empirically to ascertain that both are in fact operating in practice. Although the temporal clustering cannot be explained by Cederman and Gleditsch's model, the geographical clustering could well be explained by the processes of war, state-formation, and survival. The key here would be to clearly distinguish diffusion that leads to democratization from democratization that occurs independent of international factors but ends up in a geographically clustered fashion. Additionally, one should include control

[6]See, however, §1.2.3, where it is demonstrated that it is the transitions to rather than the survival of democracies that show the clearest clustering patterns.

variables related to international conflict to see whether the transition to democracy alters the chances for survival.

2.3 Elite perspectives on democratic diffusion

Leaving the military strategic perspective on the diffusion of democracy behind for now, I will turn to explanations derived from the elitist perspective on domestic processes of democratization. Since these theories still concern international patterns of democratization, they continue to focus on international aspects, but without the focus on war and conflict. To focus on elite actions in the study of democratization is perhaps more common than to concentrate on mass public opinion (McFaul 2002). The idea of seeing diffusion of democracy as a diffusion of innovation similar to that often studied in business studies is clearly an idea of elites accepting the benefits of an innovation in political authority structures. The implicit assumption here is that democracy is better than autocracy and it is only a matter of time for all countries to democratize. Another explanation to be highlighted in this section is more closely embedded in the field of international relations again, but this time with a more peaceful character: the effect of conditional political aid. In recent years it has become more and more fashionable for democracies to attach conditions of democratic governance to their development aid for foreign countries. Both of these factors are mostly at an elite level, with elites observing developments abroad as an example, or elites having to negotiate with financial aid providers. In both cases, however, an effect on public opinion can certainly not be excluded. The distinction between the elite perspective and the mass perspective is thus in this context not very sharp.

2.3.1 Diffusion of innovation

Modelski and Perry (1991) study the global trend of the increasing number of democracies from the perspective of that of technological innovations. They argue that democratization can be seen as an improvement of the way authority structures are organized in a manner comparable to that of technological improvements in other areas. Countries thus will learn from each other and apply similar changes. In the literature on innovations, usually a distinction is made between innovators, early adopters, late adopters, and laggards (Rogers 1995; Modelski and Perry 1991), which together forms an S-shaped curve whereby first only a few actors innovate, then the speed of the spread of the innovation rapidly increases, while after half of the population adopted the innovation, the speed reduces again to the few actors that innovate very late. Modelski and Perry (1991, 2002) demonstrate how the percentage of the world population that is living under a democratic regime follows a pattern over time very similar to this curve, with the situation at their time of writing being that of the development being at its fifty percent turning point. Despite the fact that it is somewhat awkward to demonstrate an S-shaped curve with only the bottom half (and to subsequently extrapolate two centuries into the future), the thesis itself is an interesting one.[7]

Although less based on its similarity with the diffusion of technology, Starr (1991) emphasizes similarly the existence of a model or prototype and

[7] In effect, what Modelski and Perry (1991, 2002) do is explaining the fraction of democracies, as a fraction of the world population, aggregated to decades, by the time elapsed since the start of their data collection. They do this using the Fisher-Pry model, $\frac{F}{1-F} = e^{2\alpha(t-t_0)}$, where F is the fraction of the world population living under a democratic regime (Modelski and Perry 2002: 367), which is *de facto* a logit link function. Using such a function makes sense when the dependent variable is a fraction (Papke and Wooldridge 1996). They thus simply observe an overall pattern, ignoring a lot of variation by aggregating to decades, and without any control variables, which they then extrapolate to a further two centuries.

its related demonstration effects in his discussion of the international diffusion of democracies. These demonstration effects can be internal to the country, thus generating a positive or negative reinforcement effect (Starr 1991: 360), or they can be external or global in character. Or, as Rosenau puts it: "citizens and leaders in all parts of the world are increasingly able to comprehend where they and their collectivities fit - and should fit - in the processes of global politics" (Rosenau 1988: 359). One could speak of an international culture of norms and values that affect the internal structures of national states. States become "entities embedded within a worldwide cultural framework that influences their constitution and activity through exposure to world standards and principles of political citizenship" (Ramirez, Soysal and Shanahan 1997: 737).

Fordham and Asal (2007) demonstrate how, although less significantly than the effect of geographic proximity, the institutions and practices of major powers in the international system provide such an example for other countries in the world. More minor states are inclined to copy the practices and institutions from apparently successful major powers, although more visible for female suffrage and the practice of jailing or killing political opponents than for democracy in general, which are the three dependent variables the authors study. One would expect the prestige of institutions of major powers to have such a demonstration effect because of their heightened visibility - the serious implications of such a state's actions make that they are observed more closely; their stronger influence on international media and information flows; and their explicit attempt to emphasize their superiority and invincibility to maintain their international power status (Fordham and Asal 2007: 32-33).

An interesting question is why democracy is seen as similar to a technological innovation by these authors (except for Fordham and Asal). The argument put forward by Modelski and Perry is closely related to both the survival thesis (see §2.2.3) and the effect of public opinion towards democracy on the political regime:

> "The principal advantage of democracy lies in its capacity for enhancing cooperation and managing conflict. Cooperative societies and societies that handle conflict successfully are more effective and productive than those that do not. Members of democratic societies are, as Rudolph Rummel has shown, much less likely to suffer from violence and politicide (killing by government), nor do they fight wars with each other. And, according to Amartya Sen, they are also much less likely to experience hunger. Contemporary experience also shows, as in the case of China or the Soviet Union, that political repression and social violence go hand in hand with environmental degradation. No wonder that people increasingly prefer to live in democracies, and persistently reveal their preference with their feet, in their choices of countries of immigration. No wonder then that the example of democracy is contagious, and that it spreads and snowballs." (Modelski and Perry 2002: 367)

In his vision of democracy as an innovation, Starr (1991) primarily emphasizes the changes in the options available to elites in designing the political system: "The changing levels of democracy in the system, the region, or in neighbors will be seen as changing the 'menu' of states, that being the overall incentive structures within which decision makes and peoples consider foreign and domestic policy" (Starr 1991: 361). In this sense, this idea fits well into the literature of policy diffusion, seeing the constitutional arrangement as a type of policy. For example, Simmons and Elkins (2004) make a distinction between two key factors in diffusion. The first factor is the effect international developments have on the pay-offs of a particular regime, for example in terms of the chances of survival for a democracy (see §2.2.3). The second factor is the information available about the consequences of a particular policy chance, which is clearly in line with Starr's 'menu of choice'. Similarly, Braun and Gilardi (2006) describe learning as one mechanism of policy diffusion, similarly concentrating on the changes in information available to policy makers due to experiences in other countries.[8]

The role of geographic borders in this perspective is of relatively little importance, with innovation spreading over large distances (Starr 1991: 360). In the literature on democratization as innovation, contingency of countries is of only minor importance. Innovation theories might help in explaining the global trends towards more democracies, but they do little in terms of understanding either the spatial clustering or the waves - which somewhat contradict the S-shaped curve of regular innovations - of democratization.

As is clear from the quote above, Modelski and Perry do not necessarily see the diffusion of democracy as an elite business, but rather as regular citizens changing their opinions towards democracy on the basis of observing experiences abroad. This kind of demonstration effect is not directly incorporated in the model below. The changes in opinion due to democratic propaganda and those due to cross-border communication between individual citizens are key to the model, but not explicitly any demonstration effect. The two are very similar, however, and perhaps indistinguishable. The difference is one of intent on the side of democratic regimes. Imagine a citizen

[8]See also the categories of inspiration (voluntary action by domestic actors to emulate foreigners) and coalition (cooperation between foreigners to be emulated and domestic actors) of Jacoby (2006).

of the Soviet Union listening to Radio Free Europe. One could argue that this is a form of propaganda, of democracy promotion, by the West, which is explicitly modeled here, or one can see this as this citizen observing the experiences with democracy abroad, which would be a demonstration effect. The difference is primarily in the intention of the author of the news that is being read, which cannot as such be measured. Most likely, both mechanisms will co-exists, but they are empirically indistinguishable. Whether the diffusion takes place only in contiguous regions or on a world-wide scale also does not allow one to distinguish between the two theories. Both communication and demonstration can take place over long distances, especially with modern communication technology.[9]

2.3.2 Democratization and conditionality

In the literature on democracy promotion strategies, three types of promotion are generally distinguished: control, conditionality, and incentives (Whitehead 1996). Control is the equivalent of what was referred to as democratization by force above (see §2.2.2). Conditionality refers to aid donors setting requirements related to liberalization and democratization of the recipient political regime before applying a certain reward. Incentives refer to giving a reward a priori, with the purpose of stimulating changes in the recipient country. These incentives are often presented as conditions, thus forming a sort of pseudo-conditions, where the actual threat of losing the reward is in fact very low, despite the strong language used (Ethier 2003). In some cases the foreign aid directed at promoting democracy does so in a more direct sense, however, for example "through technical assistance focusing on elec-

[9]Starr (1991: 362) makes a distinction between two types of diffusion, emulative and infectious, in essence referring to the same two perspectives.

toral processes, the strengthening of legislatures and judiciaries as checks on executive power, and the promotion of civil society organizations, including a free press" (Knack 2004: 251). In this section the focus is on conditionality with the aim to alter the behavior of the political elites, rather than for example the promotion of civil society organizations. The latter gets closer to what has been labeled 'broadcasting' in the model developed in this thesis below.

The most prominent example of the effect of conditionality are the accession politics of the European Union. Organizations like the European Union or, to a lesser extent, the African Union have requirements concerning the level of democracy within a country before membership to the union is considered. The effect of this conditionality on democratization in Eastern Europe is generally accepted to be strong (Kopstein and Reilly 2000; Ethier 2003; Levitsky and Way 2005; Jacoby 2006). Eastern Europe is of course the prime example of this mechanism. The democratizations in that area have been the main trigger of the renewed debates on democratization and on the literature on waves of democracy and the related democratic diffusion (Doorenspleet 2001). A perspective of decolonization (see §2.2.2) or one of political conditionality might do more to explain the transformations in Eastern Europe than one of a diffusion of attitudes.

The effect of conditionality is likely to be dependent on both the leverage a country has on another country and the strength of the ties between countries. The leverage is affected by the power of the recipient country, the presence of additional competing goals (other than democratization) within the donor country's policies towards the recipient, and the presence of alternative providers for the recipient, other than the democratic donor (Levitsky and Way 2005: 21-22). Levitsky and Way (2005: 22-23) distinguish five types of linkages that affect the effect of conditionality, namely economic, geopolitical, social, communication and transnational civil society linkages. They subsequently state that all these linkages are generally geographical in nature - proximate countries have closer ties than distant ones.

With the typical example of Eastern Europe in mind, conditionality of membership of a regional cooperative organization is one of the strongest competing explanations of geographic and temporal clustering. The European Union is undeniably also a geographically concentrated organization - which is highlighted by the debates on whether or not to allow Turkey to enter the union, which tends to at least partly focus on what the 'natural borders' of the European Union are. The most prominent example of the democratic diffusion literature might, ironically, be the weakest.

Conditionality might be a mechanism that coexists with diffusion, but the mechanisms are very different and the explanations focus on an entirely different aspect of the process. Where diffusion focuses on the establishment of new ideas and norms, in a relatively voluntary fashion, either among the masses or the elite, the conditionality explanation focuses on the application of democratic institutions for utilitarian reasons, external to the democratization itself. Empirically, the two might be difficult to distinguish, since one can never measure to what extent politicians who support democratization are genuinely convinced of these new ideals or are simply calculating the benefits of their changed attitude.

During the past two decades, foreign development aid has been more and more linked to democratization and what is generally referred to as 'good governance'.[10] The foreign aid program of the United States government, USAID, is the strongest example of such an approach to aid, as are those

[10]More specifically, this became part of the foreign aid discourse around 1986-1989 (Gibbon 1993: 52).

of the key international players in this field, the World Bank and the International Monetary Fund. The incentives are generally very weak, however. According to Ethier (2003: 107), except for a few African states, no recipient states were punished after refusing to carry out democratic reforms. Leandro, Schafer and Frontini (1999) provide a similar negative evaluation of both the compliance rates and the subsequent punishments to reduced aid flows. The effect of such aid on democratization appears to be negligible (Ethier 2003; Knack 2004)[11] and one could even argue theoretically how foreign aid reduces the reliance of political elites on taxes, which in turn reduces the chances for democratization (Knack 2004: 253). Most recipients of such conditional aid only partially implement the required policies, usually focusing only on the short-term policies rather than the more long-term institutional changes that are required for increase transparency, accountability, and good governance (Leandro, Schafer and Frontini 1999: 287) and thus for democratization. Sanctions have been particularly weak in cases where modest democratic transformations were implemented, but where those were combined with electoral fraud and media manipulation to strengthen the power of the elites (Levitsky and Way 2005: 22).

Although this type of democracy promotion is definitely an international factor to democratization, its effect on temporal or geographic clustering is likely to be small. Not only because its overall effect on democratization appears to be negligible, but also because foreign aid is often a very long term relationship and many donors donate to countries across the globe. Cultural

[11]Opposite findings also exists, see, e.g., Kalyvitis and Vlachaki (2006). Their research appears fairly robust, but although they perform many tests on the robustness of their results, they entirely ignore both temporal and spatial autocorrelation of democratization, and their instruments to avoid endogeneity bias are very unlikely to be unrelated to democratization - as they assume. Knack (2004) uses slightly better instruments, but their independence from the level of democracy is also highly debatable.

proximity cannot entirely be ignored here, however, as former colonial capitals still tend to have stronger aid relationships with their former colonies than average and ties in general appear to matter (Levitsky and Way 2005).

Democracy promotion takes either the form of influencing elite behavior or it has a grass-roots orientation. This grass-roots orientation of democracy promotion is closely related to the model presented here and indeed, one of the key mechanisms of the model is similar to this activity. This section focuses rather on the elite level of this mechanism, which is perhaps the strongest competing explanation of democratic clustering, both in space and time. There is of course no contradiction between the idea that democratic norms spread between populations of different countries while at the same time political elites try to affect policies of liberalization and democratization. To distinguish the two empirically, one will have to look at trade and aid flows between countries and attempt to measure the diffusion of democracy at a local rather than a global level.

2.3.3 The end of history?

One of the categories of diffusion distinguished by Braun and Gilardi (2006) is what they label 'taken-for-grantedness'. For most policy choices, various options are clearly available. However, in some cases, as Braun and Gilardi argue, certain policies are simply seen as the only viable option and in an almost axiomatic sense considered to be the only effective policy. Examples would be female suffrage or the abolition of slavery - slavery is rarely considered a viable policy option these days (Braun and Gilardi 2006: 311).

The examples of Braun and Gilardi already demonstrate how closely this idea can be related to democratization. In the 1990s one could often hear arguments about how liberalism and democracy are taking over the world, especially after the end of the Cold War, generally based on the famous *The End of History?* thesis by Fukuyama (1989). Fukuyama's argument is that history is fundamentally based on ideas and ideology - the exact opposite of the Marxist materialist view, despite the fact that both Marx and Fukuyama base their visions on the work by Hegel - and that all ideologies competing with liberal democracy have failed. Communism and fascism are no longer viable options and even those regimes that are still claiming to be communist are *de facto* implementing more and more policies that demonstrate the fundamentally market-oriented attitudes of the political leaders.

In the figures showing democracy over time in the previous chapter, one can not only see the waves of democracy, but also its persistent increase over time. During some periods it grows more than during other periods and in some areas more than in other, but overall the level of democracy is increasing in the world. This would fit with the thesis of Fukuyama that democracy is gradually becoming the only available option, is starting to become taken-for-granted by political leaders, if not entirely, then at least in their *de facto* policies. This perspective tells us little about the causes of the geographic or temporal clustering of the level of democracy, but it does suggest a global pattern of democratization, a certain lack of interdependence of observations, and an ideology based explanation of democratization. One could argue that this vision of the democratization of the world is entirely compatible with the model presented below, and simply refers to the possible final state in which all countries are democratic. The model here is based on a diffusion of norms, which is practically the equivalent of saying that a particular ideology is become more and more prevalent. Unlike the other theories presented in this chapter, the taken-for-grantedness idea is not so much a competing explanation of democratic diffusion, but rather a specific state the model presented below could end up in.

In general, but in particular with this taken-for-grantedness perspective, one has to be careful not to lose sight of the important aspects of the definition of democracy. Although the claim to be democratic is becoming more and more widespread worldwide, also the prevalence of regimes that are democratic in name but in fact under the solid control of a single leader is clearly increasing. There appears to be a common trend in recent years for presidents of countries with democratic constitutions to alter the constitution to allow for longer or indefinite term limits. Often combined with electoral fraud, this ensures that specific leaders can stay in power for long periods of time, while electoral accountability is waning (Zakaria 1997; Levitsky and Way 2005). Here we are in the grey area between Braun and Gilardi (2006)'s taken-for-grantedness and their symbolic imitation. The latter refers to the adoption of particular policies because it has a positive effect in a symbolic sense, even when it is known to be ineffective. A dictator under a democratic constitution professes to be democratic because of the legitimacy and international standing that comes with this label, while in effect the democraticness of the regime is highly doubtful.

2.3.4 Clustering as a spurious effect

The fact that democracies show international patterns like clustering and snowballing effects is often used as evidence that international factors must play an important role in the transitions to democracy, since purely local factors cannot explain the international patterns. It is, however, not necessary for transitions themselves to be affected by the international context for this kind of pattern to occur. Internal factors of democracy themselves might be geographically or temporally clustered (O'Loughlin et al. 1998; Braun and

Gilardi 2006; Gleditsch and Ward 2006: 923). An obvious example of this could be the economic performance of a country, which is significantly affected by spillover effects from neighboring economies (Hak 1993) and which has frequently been related to processes of democratization (Lipset 1959, 1960; Cutright 1963; Burkhart and Lewis-Beck 1994; Przeworski et al. 2000). Neighboring countries will experience some similarity in economic development due to the diffusion of technological innovation, to mutual trade, and to similarity in natural resources that make them dependent on similar international markets. Some preliminary work testing the extent to which democratic diffusion can be explained by controlling for economic diffusion suggests that there is still a diffusion effect to be explained (Elkink 2003). In general, most studies on democratic diffusion include controls on economic developments, which do not cancel out the effect of spatial proximity (e.g. Doorenspleet 2001; Wejnert 2005).

A clear example is the relation between oil and democracy. In most analyses of the relation between wealth and democracy, an exception is made for the oil states, because they are not democratic, but very wealthy (Doorenspleet 2001). A common, geographically clustered factor (oil) affects their levels of democracy probably in a similar fashion. It is not that the (lack of) processes of democratization are affecting each other across borders, but it is their common geological characteristics that clusters geographically and that affects their political regimes. In this case it provides authoritarian regimes with sufficient income to provide for their citizens in such abundance that calls for democratization are significantly reduced.

The key difference from the other explanations of geographical clustering is that while "a large number of actors choose similar policies, (...) individual choices are independent" (Braun and Gilardi 2006: 305). Spurious diffusion is here described as a common reaction to similar pressures, whereby the reaction itself is an independent, rational choice. "Spurious diffusion makes the implicit assumption that some problems have an inherent 'rational' solution. If such a solution does not exist, in effect, it is highly unlikely that many actors would come up with similar solutions independently. This assumption is questionable but defensible" (Braun and Gilardi 2006: 305). The strong connection between spuriousness and rationality as posed by Braun and Gilardi is not obvious - i.e., why would psychological or emotional responses not be natural and therefore similar across different elites? -, but their emphasis on the independence of the individual decisions is a crucial characteristic of spurious diffusion.[12]

One could also imagine an interactive effect between the diffusion of democracy and internal characteristics that cluster geographically (O'Loughlin et al. 1998: 550). Bergesen (1992) analyzes the wave of democratizations in Latin America and Eastern Europe in the 1980 and early 1990s as an effect of their comparable position in the world economy. According to his analysis, countries in the semi-periphery of Wallerstein's world system react in a particular way to changing pressures in the world economy. These reactions are partly synchronized through demonstration effects among countries from this same segment of the world economy. The effect of the world economy, of Kondratieff's economic cycles, affect neighboring countries similarly, while this effect is exacerbated by demonstration effects, the actual democratic diffusion, between these countries. In other words, democratic diffusion occurs

[12]Spurious diffusion is also known as hierarchical diffusion (see, e.g., O'Loughlin et al. 1998: 552). Simmons and Elkins (2004: 172) suspect that the diffusion of policies towards financial and economic liberalization is largely a spurious relation, but they have a different concept of spuriousness, where they consider this diffusion to be a combination of changes in pay-offs for particular policies due to policy changes abroad (e.g., a race to the bottom of tax rates) (see §2.2.3) and changes in the information available concerning particular policies (see §2.3.1).

within groups of countries that experience similar other clustered domestic circumstances, more so than between different such groups.

Spurious democratic diffusion is an important alternative explanation for the patterns explained in this thesis, both in a spatial and in a temporal sense, and thus something one has to carefully control for when doing empirical analyses in processes of diffusion. It is in no sense incompatible with any of the other theories presented and Elkink (2003) indeed suggests that both economic and democratic diffusion co-exist and relate to each other, but it is easy to overestimate the level of democratic diffusion due to such alternative clustering factors.

2.4 Diffusion through the masses

One of the key categories of Braun and Gilardi (2006)'s classification of policy diffusion mechanisms is what they call common norms. They obviously refer to common norms among policy-makers, but one could relate this straightforwardly to the presence of common norms among citizens of different countries. The model presented in this paper focuses exclusively on the role of the masses in the process of democratization. Although it would equally well apply to scenarios where the masses explicitly topple the regime, like the Jacobians in the French Revolution, as to where the regime democratizes when it senses the lack of the necessary public support, it only relates to transitions where the perceived majority public opinion matters.

Many of the theories discussed in this chapter implicitly refer, for some possible cases, to mass level behavior. The domino theory of Eisenhower did not only worry about communist countries attacking neighboring democracies, but also about grass roots organizations infiltrating neighboring societies. The decolonization literature or the democratization after occupation are not only related to forcing the elites to behave in a certain way, but also to instilling a particular political culture. The idea of democracies having a demonstration effect, creating a diffusion of innovation type mechanism can apply to citizens as well, as Rosenau's quote suggests. And the end of history thesis is about how ideas drive politics, which surely includes norms held among the general population.

Yet although all these theories acknowledge implicitly some mass level effect, they all focus primarily on the elite level in explaining transitions. This is a common approach to take in the democratization literature (McFaul 2002). As mentioned in the introduction, this thesis is by no means an attempt to deny the importance of elite behavior in regime transitions. The elites are likely to play a crucial role. The fact that elites are important, however, does not imply that public opinion can be ignored (Welzel 2006), nor that the geographical and temporal clustering of democracies can only be explained through elite based theories. It is reasonable to assume that a regime transition stands a much higher chance of success in an environment favorable to the new regime than in a hostile environment. In fact, the model presented below presents elite based regime transitions as a 'random error', where at times regimes make a transition without general public support and have some chance of survival. A random error thus in the sense that it is not the mechanism that is key to the model, but it is nevertheless acknowledged as a missing variable, as a co-existing alternative explanation of regime change. The next chapter will discuss in more detail the mass based perspective on regime changes.

2.5 Conclusion

In §1.1.3 we made a distinction between three types of clustering of democratic regimes: temporal, spatial, and spatio-temporal clustering. To what extent do the various theories presented here contribute to understanding these various types of clustering? Table 2.1 gives and overview of which are likely to contribute where.

	Temporal	Spatial	Spatio-temporal
Domino theory	-	+	+
Force / decolonization	+	-	-
Survival	-	-	+
Diffusion of innovation	+	(+)	-
Conditionality	+	+	+
Incentives	-	-	-
End of history	+	-	-
Spurious diffusion	+	+	+
Diffusion of norms	+	+	+

Table 2.1: Alternative explanations of international factors in democratization related to three types of clustering

It is clear from the table and the discussion above that the strongest alternative explanations are political conditionality and spurious diffusion. The former exemplified by the fact that the prime example of democratic diffusion, the transformation of Eastern Europe, is the one most strongly explained by conditionality, and the latter because one always has to be careful controlling for spurious effects. Many aspects of society can be assumed to diffuse across borders, including economics and culture, which in turn are common explanations in the democratization literature.

The next chapter will in more detail discuss the relation between norm diffusion and revolutions. First the key mechanisms of the model that relate norm diffusion to revolutions will be discussed and then some attention will be paid to mass based theories of democratization.

CHAPTER 3

Attitudes and revolution

Chapter 2 listed various possible explanations for the empirically observed clustering of democratic regimes. The different theories are based on different assumptions about what matters most in democratization processes. They differ on whether the behavior of elites or mass public opinion matter more; whether realist factors like the military-strategic position in the region or ideological factors matter more; whether international or domestic factors are more important; and whether the material well-being or the ideological position of citizens matter more. Although sometimes explicitly compared, in many cases these are assumptions made prior to the research and arguments that one or the other of those factors have recently been overemphasized at the expense of another factor are almost the standard introduction to many of the articles in the democratization literature. Without claiming that any of those assumptions is more plausible than another, the assumption of this thesis is that mass public opinion, ideology, and international factors matter - perhaps not exclusively, but to a relevant extent.

In this perspective, the clustering of regimes is explained by a process of attitudinal diffusion, both orchestrated by (democratic) governments and spontaneously between individuals. Democratic governments establish democracy promotion projects to convince individuals abroad of the positive qualities of democracy. For example, Western nations supported the running of *Radio Free Europe*, which distributed pro-democratic news in the former Soviet Union and its satellite states, through radio and written news broadcasts. Similarly, departments of state in various countries attempt to stimulate grassroots organizations in potential democracies to influence the opinions of individuals at the mass levels of society, both to stimulate the growth of a civil society and to influence popular opinion towards liberalism and democracy.

Through various channels, individuals communicate with other individuals abroad. They travel abroad and interact with citizens there; they trade with foreigners; they use the Internet or phone to communicate with others; they read foreign media; etcetera. To a large extent, one's beliefs and attitudes are formed through interaction with others, who might have more or less similar attitudes. Ideas are formed through communication and interaction with others. Through such communication, individuals change their attitudes and beliefs about many things, including democracy. The assumption of this thesis is that these attitudes towards democracy will affect their behavior, and that their behavior will affect the political regime under which they live.

This chapter will discuss the theoretical assumptions and foundations of this model. We will first briefly turn to the sociological and social psychological literature on attitudes and attitudinal change. A brief discussion of this fundamental aspect of attitudinal diffusion is a crucial step in the theoretical defense of the model. We will then turn to how attitudes affect behavior and in particular how this relates to anti-regime behavior. Under what circumstances does one prefer to hide attitudes contrary to those held by the political regime and when does one publicly denounce the regime? The model of the spiral of silence, preference falsification, or cascading revolutions - three different names of essentially the same mechanism - will shed some light on this aspect of the model. This will gradually bring us back to home territory, the political science literature, again, as we will turn to a brief overview of the role of mass public opinion in the democratization literature.

3.1 Diffusion through attitudinal change

3.1.1 Attitudes

The main model in this dissertation is based on the international diffusion of attitudes, the relation between these attitudes and behavior, and the effect of this behavior on political regime transitions. Before we can discuss any of these elements of the model, we have to pay some attention to what we actually understand by this term "attitude" and what the common interpretation of the term is. In the social-psychological literature, attitudes are generally defined as "a learned, global evaluation of an object (person, place, or issue) that influences thought and action" (Perloff 2003: 39), or, somewhat more explicitly but along the same lines:

> "When we talk about attitudes, we are talking about what a person has learned in the process of becoming a member of a family, a member of a group, and of society that makes him react to his social world in a consistent and characteristic way, instead

> of a transitory and haphazard way. We are talking about the fact that he is no longer neutral in sizing up the world around him; he is attracted or repelled, for or against, favorable or unfavorable" (Sherif 1967: 2).

Attitudes are thus learned and can be changed. An individual can change attitudes on the basis of communication with others, of seeing advertisements, of direct experiences with an object. Whereas an activist of a government can attempt to persuade individuals to accept the government's view on a particular type of regime, one might also change one's attitude simply by observing other individuals, their attitudes, and their experiences. Attitudes are partly generated through one's culture and education, but they are not fixed or genetically set (Perloff 2003: 36-41) - a feature that is of course crucial in a model of attitudinal diffusion.

Furthermore, attitudes concern someone's evaluation of a particular object. The object of interest in this thesis is that of the concept of democracy. A political regime is a broad concept and attitudes towards a regime will be a complex of attitudes towards a variety of objects more or less closely related to the government.[1] Imagine a citizen X in a country with a democratic constitution, a high level of corruption among public officials, and substantial electoral fraud. Citizen X once tried to petition a local councillor and could not approach him without paying a bribe. Friends tell her that the current head of state won the elections only through substantial fraud, but she knows that her salary was paid much more consistently since he has been in power.

[1]Furthermore, one might well argue that the meaning of the object itself is debatable and that, before one can analyze the dynamics of attitude formation towards an object, one has to analyze the dynamics of the changing meanings the object acquires in society. One thus ends up in a more discursive type of analysis (see, e.g., Shapiro, Bonham and Heradstveit 1988) of attitude formation towards democracy. Although admittedly crucial for a full understanding of the process of attitude formation, this particular aspect is left aside in this thesis.

The leader of the opposition, who sometimes briefly appears on television, is an unappealing man, with a distrust-full look in his eyes, who reminds her of an uncle she long fell out with. Her mother instilled a distrust in politics in general and taught her to concentrate on family and friends. Through an Internet chat channel she often frequents, she hears about the personal wealth of online friends just a few hundred miles away, who live in a democratic country known for its long-standing democratic traditions. Her friends often ask her how she survives in this awful dictatorship - they only hear about the corruption and fraud on their local news. All these various different experiences, communications, and intersecting attitudes and beliefs will affect her attitude towards the concept of democracy or the possibilities or desirability for democracy in her country. This complexity of the object, the attitude towards the object, and the many intertwining stimuli that affect these attitudes, will be largely ignored in this thesis. For the sake of argument, a straightforward, latent, one-dimensional attitude towards democracy will be assumed. Although affected in many ways, it is assumed that there is this latent one-dimensional scale in one's head that summarizes to a large extent the attitude towards democracy. This is an obvious oversimplification, but, keeping Occam's Razor in mind, it will be assumed that this simplification does not harm the main argument of the thesis.

Different theories exist within the field of social psychology on how individuals organize such sets of attitudes. According to the prominent expectancy-value approach, attitudes are "a multiplicative combination of (a) strength of beliefs that an object has certain attributes and (b) evaluations of these attributes" (Perloff 2003: 46). The attitudes towards democracy are thus a sum of all the beliefs concerning the applicability of particular attributes - democracy leads to wealth; democrats are likeable people; democracy leads to corruption; etcetera - and the evaluations of these attributes (Ajzen and Fishbein 1977). The symbolic approach similarly looks at the various intersecting attitudes, but in a more associative fashion - the associations that come up when thinking about 'democracy' - and their emotional loading. The focus is thus more on emotions, affections, and symbols, rather than on beliefs and evaluations (Perloff 2003: 47-48). More heart than head in the organization of attitudes. Although for the creation of a policy to stimulate international democratization (see §7.2) it will be crucial which of these two mechanisms are more important in attitudes towards democracy, for the model presented in this thesis there would be no discernible difference. It is assumed that individuals have an attitude towards democracy that can to a large extent be summarized by a one-dimensional attitude and that this attitude can change through inter-personal communication and international propaganda - how this individual internally organizes this attitude or whether it is more emotional and symbolic or more rational does not affect the mechanisms of the model as described below.

Although most of this thesis talks about individual attitudes towards democracy, the way it affects the political regime in a country is based on more mass based attitudes - on the attitudes of many members of a population rather than just one. This is exactly the distinction between elite and mass based explanations of democratization. We could thus instead speak of the role of public opinion, rather than of individual attitudes, even if in the agent-based models we model citizens individually. It should be pointed out that the term "public opinion" is used somewhat loosely in this thesis. The concept is highly controversial and many different interpretations exist (Price 1992; Noelle-Neumann 1993; Herbst 1993).[2] Conceptualizations vary

[2]See Price (1992) and Noelle-Neumann (1993) for extensive historical overviews of the use of the term "public opinion".

from interpreting public opinion as the opinions of public members of the elite, as an underlying consensus of norms and values in a society, as the majority opinion, as the aggregate of opinions measured in a sample survey, or as a fiction projected by the ruling elite (Herbst 1993: 44-46). In the remainder of this thesis, public opinion refers to the attitudes of members of the general public, either part of the elite or not, and either in the majority or not. It thus neither refers specifically to a consensus underlying the social structure, nor to the overall majority opinion. The distinction between public and mass opinions (Price 1992: 26-29) is also entirely ignored, using both terms interchangeably. Similarly, the terms opinion and attitude are conflated, as is common in many works in the area (Price 1992: 46-47).

3.1.2 Communication

Given that individuals have particular attitudes towards democracy, we are now concerned with how these attitudes change, in particular through interpersonal communication and through democratic propaganda, and how these attitudes translate into behavior. In this section we will discuss the mechanism through which attitudes change; in the remainder of this chapter we will concern ourselves with the relation to actual behavior. The theory used in this paper to explain the changes in attitudes themselves through communication is the social judgment theory (Sherif and Hovland 1961). The basic premise of this theory is that "[w]hereas the quality of arguments may determine the extent to which one is being persuaded by another person, often people respond quite simple by favoring positions close to their own, and rejecting more distant positions" (Jager and Amblard 2004: 295). When confronted with the attitudes of another person, an individual thus adjusts his or her own attitude depending on the difference in opinion. When the advocated position is close to that of the receiving individual it is said to be within the *latitude of acceptance* and the individual is likely to change attitude somewhat towards the advocated position. On the other hand, when confronted with a position entirely different from one's own, within the *latitude of rejection*, the individual will emphasize the difference and move slightly away from the advocated position. In between there is a *latitude of non-commitment* where the individual is not affected by the advocated position. Generally, individuals with an extreme attitude towards a particular object will have relatively large latitudes of rejection (Perloff 2003: 60).

The concepts of latitude of acception and of rejection are closely related to those of *contrast* and *assimilation*. These terms refer to the psychological tendency to distort learned facts according to existing beliefs. One can interpret the same fact in multiple ways to accommodate it in terms of one's own views before observing the fact, for example by degrading an observation to "a mere exception" or by interpreting it as "typical". The Irish tend to interpret a rainy day as "typical" and a sunny day as "exceptional", regardless of the actual weather. Similarly, when judging a message from another individual or through some broadcast, one is likely to overestimate the similarity to one's own attitude when the message is agreeable and to overestimate the difference when the message is more disagreeable (Perloff 2003: 60-61).

The final key concept in the social judgment theory is that of *ego-involvement*. Although this concept has probably had the most significant impact on the field of social psychology (Perloff 2003: 61), it is of somewhat less interest to this thesis. Ego-involvement refers to the extent to which the individual is attached to a particular attitude, which affects the latitude of rejection - highly involved individuals are less likely to change their attitude -, the likelihood of contrasting - highly involved individuals are more likely to contrast others attitudes to their own - and the extent to which persuasion has to be consistent with already held beliefs to be effective (Perloff 2003: 61-62). High ego-involvement also leads to selective perception, where the world is perceived in such a way as to confirm existing beliefs and attitudes (Perloff 2003: 62) and, more strongly, disconfirm those that do not. The extent to which an argument disagrees with existing beliefs affects the extent to which one searches for arguments - more disagreements means a more serious search for arguments - and those arguments that disconfirm the argument are emphasized. This process, in turn, is stronger the more one is emotionally involved (Edwards and Smith 1996). A study by Lord, Ross and Lepper demonstrates how subjects interpret neutral or contradicting evidence in such a way that it confirms their prior beliefs. They adjust their evaluation of the quality, biasedness, logic, and conclusions all on the basis of their prior attitude towards the object under study. And, alas, this holds for laymen and scientists alike (Lord, Ross and Lepper 1979).

Edwards and Smith (1996) provide a brief overview of the finer differences between the various theories about the relation between prior beliefs and the evaluation of arguments. We will leave this argument for the social psychologists. What is interesting to us is that most of those theories suggest a polarization of attitudes where arguments in favor of prior beliefs are evaluated positively and those that do not are dismissed.[3] Thus when communicating with fellow citizens, people are likely to interpret whatever is being communicated towards their own attitudes - they either strengthen their own beliefs on the basis of communication with someone who holds similar beliefs or they discredit whatever argument is brought to the fore

[3]But note Miller et al. (1993), who point out that although there is a clear sign of a change in the *perceived* attitude by the respondent, there is less evidence for actual changes in attitude, let alone subsequent behavior.

that contradicts their prior beliefs.

The face validity of the social judgment theory[4] can be demonstrated with an example close to most readers. Imagine the quantitative political scientist arguing vehemently before a fellow political scientist defending his or her quantitative approach. If the listener is already inclined to do quantitative analyses, he or she is likely to be easily convinced by the arguments and to strengthen the believe that quantitative methods are indeed the most valid approach to political research. If the listener is an area specialist, however, used to apply more in-depth comparative methods and generally opposed to quantifying human behavior, the arguments will sound unconvincing. Moreover, they are likely to highlight the attitudes that were already objectionable to the listener, who will subsequently become even more opposed to quantitative approaches to social science research. The effect of the arguments are thus likely to be dependent on the initial distance between the advocated position and that of the listener or judging agent.[5]

One complication of the social judgment theory is that it is difficult if not impossible to determine empirically whether the latitudes of acceptance and rejection indeed precede the change in attitude and are a cause of the direction of the change, or whether it is merely a fact that might not precede the actual attitude change. Furthermore, it is difficult to determine whether it is the actual attitude that has changed, as the social judgment theory poses, or the description of the attitude, the way it is presented (Miller et al. 1993). Both complications make it difficult to empirically validate the theory

[4]The social judgment theory is closely related to, but not the same as, homophily: "Homophily - the principle that "likes attract" - is a prominent explanation for the persistence of cultural diversity. More precisely, homophily is the tendency of people with similar traits (including physical, cultural, and attitudinal characteristics) to interact with one another more than with people with dissimilar traits" (Centola et al. 2007: 905-906).

[5]The author thanks Scott Page for suggesting this example.

(Petty and Cacioppo 1996: 109-122).

3.1.3 Universality and locality in norm diffusion

The literature in social psychology on the formation of attitudes through confrontation with others' attitudes takes a very microscopic, individual view on the diffusion of attitudes. Another separate branch in the literature concerns itself more with the global patterns of norm diffusion and relates to how fashions spread, how norms spread or die out, and whether or not activists - individuals that play a particularly stimulating role in this process - are necessary for norm diffusion to occur or not. For the context of the diffusion of democracy it is important to keep in mind the clear distinction between institutional learning (Rohrschneider 1996: 424), where individuals acquire certain norms that are congruent with the institutional environment the interact in, and norm diffusion, which, in the context of democratic diffusion, "conjectures plausibly that citizens in previously authoritarian systems have developed a preference for a political order that guarantees basic political liberties" (Rohrschneider 1996: 425).

One can think of the diffusion of norms as an attitude that originates in a particular locale and that subsequently spreads out over a large number of individuals, eventually becoming a universally accepted attitude. Or alternatively, one can think of this process as a number of different norms, probably in part contradicting, that compete for acceptance. Perhaps some will remain as co-existing norms for different groups of people, while other norms will simply 'lose' and be forgotten. Acharya distinguishes two different waves of scholarship in norm diffusion. The "first wave scholarship on normative change" (Acharya 2004: 242) concerns itself primarily with "cosmopolitan" or "universal" norms, where the key actors are transnational actors. These actors can be either individuals, "moral entrepreneurs", or social movements. The focus of this research is on conversion rather than the contestation of different norms, whereby the refusal of the norm is seen as illegitimate or immoral (Acharya 2004: 242). Norms are thus considered to be intrinsically good or bad. Of course, the democratic diffusion literature shows many of these same symptoms, whereby democracy is seen as an obvious good and the predicted outcome of the whole diffusion process is a world-wide occurance of democracy (see, e.g., Modelski and Perry 1991).

The second wave of scholarship, according to Acharya, focuses instead more on the local dynamics of international norm diffusion. It looks at the fit between international norms and existing local ones, both in terms of culture and in terms of organizational structure. The entire focus is here thus on the localization of international norms (Acharya 2004: 242-243). Here one sees the international diffusing norm as competing with existing norms and a universal acceptance of the same norm is not necessarily expected from the outset. Acharya goes on in a more constructivist vain to study this process of localization: "In constructivist perspectives on socialization, norm diffusion is viewed as the result of adaptive behavior in which local practices are made consistent with an external idea. Localization, by contrast, describes a process in which external ideas are simultaneously adapted to meet local practices" (Acharya 2004: 251). He discusses how local individuals adapt international norms to fit their own environment, as well as actively seek international norms when local ones are failing. This local adaptation to international norms then explains the diversity in implementations of the same international norm.

Without reference to local cultures and norms in the process of norm diffusion, we run into a paradox. When diffusion is unconstrained and without any barriers or limitations attitudes simply spread immediately over a group of people and after a certain amount of time, these attitudes will always become homogeneously spread among this group of people. All attitudes will average each other out and one blend, constant equilibrium of attitudes will arise. New ideas might so once in a while randomly pop up, but they will immediately be assimilated into the mass opinion, slightly influencing the mass opinion and being heavily influenced by this opinion, with in the end a negligible effect on the general distribution of attitudes in the group. If poor cases make neighboring cases poorer and rich cases make neighboring cases richer, after some time all cases will eventually be somewhere near the average level of wealth. If all countries influence their neighbors towards regimes similar to their own, eventually all countries will have a similar regime. Most literature on innovation is indeed based on such patterns of diffusion, whereby first a few cases change, then the change occurs faster and faster, and then there are some late adopters that only gradually take on the new contagious state, but eventually all cases will have changed. This generates the well-known S-curve of diffusion (Rogers 1995; Ayres 1999; Modelski and Perry 1991).

This pattern, however, does not match well with what we observe empirically, either in terms of democratic diffusion or in most other examples of norm diffusion. Thus, if diffusion is to be used as an explanation of the kind of clustering, trends, and continuous dynamics we observe in the international distribution of democracies, then this diffusion has to be constrained by one or more factors. A clustered pattern is a pattern where the states of individual cases are often similar to that of adjacent cases, but whereby the overall state is not homogeneous. A system that is not clustered is one where the state of individual cases in independent of that of neighbors. A system where all cases are in the same state could either be described as consisting of one big cluster, or as not having clusters at all. In this case the term clustering is used to distinguish exactly between the cases where the diffusion leads to an overall homogeneous state and those where the diffusion leads to clusters of adjacent cases that have similar states, while other clusters exist in the same system with different systems. Diffusion thus leads to clustering, but only temporarily, while over time this clustering disappears. One should thus be careful not to equate the empirical observation of spatial and temporal clustering with diffusion.

The most likely explanation of such clustering, or limited diffusion, would be that some factors, for example geographical borders of culturally homogeneous regions, or geographical features that inhibit cross-border communication, are reducing the extent to which diffusion can take place over some borders. "Linguistic, cultural, psychological, religious, and ideological differences often serve as barriers to these information flows, leading to a differentiated political mosaic across the globe" (O'Loughlin et al. 1998: 552). In this case one needs not only to explain why political regimes diffuse, but also what limits this diffusion and creates the observed levels of clustering.

The second explanation would be that, while the diffusion started from different points in the system, the process is still ongoing and will eventually lead to an overall homogeneous state. This is the interpretation of Modelski and Perry (1991, 2002) who suggest that the diffusion of democratic regimes is currently halfway, at the point where the diffusion process is fastest and where about half of the world population lives in countries that have adopted the new innovation. They extrapolate from this finding that in 2113 ninety percent of the world population will enjoy democracy (Modelski and Perry 2002: 370). Although it seems doubtful that one can infer an S-shaped curve on the basis of just half the time period, this interpretation fits well in the general literature on the diffusion of innovations. Another reason to cast doubt on this extrapolation is the fact that the number of countries, as well as the size of the world population, is continuously changing, which is not common in the framework of diffusion (O'Loughlin et al. 1998: 553).

Axelrod (1997*a*) developed an agent-based model to explain something very close to that sought after in this paper, namely the dissemination of culture, using a model where cultures locally converge, yet globally maintain polarized clusters. This is close to the dissemination of pro-democratic attitudes which lead to a relatively stable distribution of regimes, but leaves a variety of different regimes, democratic and autocratic, despite the ongoing diffusion mechanism. Axelrod's model contains a number of agents located on a lattice, whereby agents have a certain probability that they will communicate with neighboring agents and exchange part of their cultural information. This chance of communication is directly proportional to the amount of culture the two neighboring agents already share, with communication being impossible when the two agents differ on each cultural element. In his model, culture is abstractly described as a finite set of features with each a finite set of possible values, whereby the proportion of identical features between two agents determines the probability that the two agents will communicate. When they communicate, one feature of those that are still different will be selected and shared, so that after communication, two agents are always more likely to communicate again.

Axelrod's model leads to a dynamic whereby very soon in the simulation a few large cultural regions arise where internal differences are much smaller than those between agents from different regions. Or more precisely, large zones arise where cultures are similar enough to make communication possible, while cross-border communication is not possible due to cultural differences being too stark. This situation is relatively stable for part of the simulation and then collapses into a situation where those regions become completely homogeneous and all communication between regions is impossible, and communication within regions fruitless. Thus, an equilibrium arises whereby there is local clustering - in the sense of completely homogeneous regions - and where there is global polarization - no communication or similarity between agents of different regions.

For the purposes of this paper the most serious problem with Axelrod's otherwise very useful model is the fact that the simulation ends in an eventually entirely stable equilibrium. This equilibrium is not observed in real world data - neither cultural traits nor the distribution of political regimes is ever entirely stable, unless, of course, we are still awaiting the future stabilization, but this seems a counterintuitive prediction. The world and attitudes always change. Furthermore, it seems counterintuitive to assume that with certain cultural differences communication become entirely impossible. Especially when we will take only one cultural trait into account, pro-democraticness, this is an unacceptable assumption. The question thus is how a similar model can be developed, where similar clustering occurs, but without the final stable equilibrium, with a world that will continually be in flux. Axelrod in fact already makes suggestions for this approach, introducing the idea of cultural drift, whereby cultural traits in agents can spontaneously change, making previously blocked communication possible. This leads to substantial changes in the model's behavior and leads to new interactions between different model parameters, thus, "it is not trivial to determine how the introduction of cultural drift affects cultural change in the present model of social influence" (Axelrod 1997*a*: 222).[6]

[6]A quick thought experiment would suggest that the entire dynamic of the model collapses - we are back in a situation that leads to full homogeneity, with short deviations

Axelrod himself acknowledges that his model could well be extended in various ways, including some that relate more closely to Acharya (2004)'s idea of local cultures. He mentions the inclusion of terrain effects, of geographic differences, of status, of broadcasting and education, and of cultural divergence (Axelrod 1997*a*: 221).

The model presented in this thesis uses the social judgment theory approach of Jager and Amblard (2004) rather than the proportional likelihood of communication approach of Axelrod (1997*a*) to generate the effect of local convergence combined with global polarization. Both models are attempts to demonstrate that such a pattern is possible even in the absence of norm entrepreneurs - individuals who play a crucial role in the diffusion of attitudes. In §4.3.1 the intuition behind this model will be demonstrated using a small agent-based model and in the main model, it will be incorporated as one of the fundamental mechanisms, in addition to the cascading model of revolution to implement the relationship between attitude and behavior. It is to this relationship that we will now turn.

3.2 From attitude to behavior

Once an individual acquired a particular attitude towards democracy, how does this manifest itself in action? In particular, when does this lead to action that undermines the political regime and leads to a transition to or away from democracy? Besides the social judgment theory, the second pillar of the model developed in this dissertation is the so-called cascading model of revolutions. In this chapter this theory will be discussed in the context of the theory of the spiral of silence of Noelle-Neumann (1993) and the closely

due to the cultural drift, which will immediately be absorbed again in the overall consensus. A more in-depth analysis is left for future research.

related concept of preference falsification by Kuran (1995). Whereas Noelle-Neumann writes in relation to regular opinion polling in Western democracies, Kuran puts a very similar theory forward as an explanation of the surprising pace at which the revolutions in Eastern Europe took place, thus linking the theory of preference falsification to the theory of cascading revolutions as introduced by Granovetter (1978).

The key attitude studied in this thesis is that towards democracy as a political regime. Citizens are assumed to have a certain evaluative position in relation to their current political regime and they are more or less supportive of a regime change towards or away from democracy. As Ajzen and Fishbein (1977) point out, however, it is important to be very specific about the target of the attitude and the target of the related behavior. Research where the target of the attitude was only indirectly or in generalized terms related to the target of the studied behavior showed only weak or ambiguous correlations between attitude and behavior. Research, on the other hand, where the target of both was very specific and more or less identical showed strong correlations. "Even when it can be shown that an action has evaluative implications for the target, the most appropriate predictor of the single-act criterion is the attitude toward the action rather than the attitude toward the target" (Ajzen and Fishbein 1977: 891). For example, "students were likely to cheat on a test if cheating on that test was potentially useful or desirable, irrespective of their attitudes toward cheating in general" (Ajzen and Fishbein 1977: 894). In the model developed below, the focus is on the relation between a general attitude towards democracy and publicly opposing the regime. In other words, not as specific and identical as Ajzen and Fishbein suggest. In defense of the model, two ways of looking at this appear to be possible: (1) the ascribed behavior is at a similarly abstract level as the measured behavior - as discussed above, the attitude is considered to be a multifaceted combination of evaluations of various aspects of the regime and the public protesting behavior is considered to be a general category of any type of publicly visible actions that are clearly against the current political regime and (2) the way the attitude towards democracy has been implemented using the cascading model of revolutions (see below), one could argue that the actual implementation is an attitude towards participating in public protest against the current regime, rather than as a general attitude towards democracy. The model is general enough to incorporate either interpretation. With this caveat in mind, we will now turn to the preference falsification theory, which is at the core of the modeled relationship between attitude and behavior.

3.2.1 Preference falsification

In *The Spiral of Silence*, Noelle-Neumann (1993) develops a theory of the fear of isolation that leads individuals to hide their preferences when they are incongruent with the perceived majority view. During the campaign for the German parliamentary elections of 1965, the two major parties, the Christian-democratic CDU/CSU and the social-democratic SPD, stayed very close to each other in the opinion polls. Everybody was assuming that the election results would be very tight, but instead, the CDU/CSU won with a clear majority. The opinion polls had thus given a misleading picture of the political climate of the campaign, resulting in a very surprising outcome of the elections. Another question in the same opinion polls was, retrospectively, far more informative. The question ran: "Of course nobody can know, but what do you think: who is going to win the election?" For this question, the parties were very close nine months before the elections, but had gradually moved widely apart, in favor of the CDU/CSU (Noelle-Neumann 1993: 2-3). In other words, the perceived majority by respondents was far more revealing than the demonstrated majority by these same respondents in the polls.

There are two well known theories that explain this phenomenon where the perceived majority gets more votes than the actual amount of support. Noelle-Neumann's approach emphasizes the fear of isolation individuals experience that leads them to support the majority view, while the bandwagon theory emphasizes a more rational calculation that leads individuals to support the winner or at least not waste effort to support an obvious loser (Pierce 1940). As Noelle-Neumann points out, this debate reflects a common distinction between European scholars emphasizing psychological explanations of visible behavior and American scholars emphasizing more rational explanations, the distinction being so influential that American students simply walked out of her classes when she suggested that also their opinions are partly formed by fear of isolation (Noelle-Neumann 1993).

Noelle-Neumann spends most of her book describing both the empirical foundation for her claim and the historical legacy of her perception of public opinion formation in philosophical writings before the 20th century. In a series of opinion polls in Germany from the early 1960s onwards, respondents were tested on the following scenario: "Assume you are faced with a five-hour train ride and someone in your compartment begins to talk very favorably about Chancellor Brandt. Would you like to enter conversation with this person so as to get to know his or her point of view more closely, or wouldn't you think it worth your while?" Of the respondents that elsewhere in the same survey stated that they agreed with Brandt, 50 percent said that they would join in the conversation, while only 35 percent of those who disagreed with Brandt would participate (Noelle-Neumann 1993: 22-23).[7] Thus not only when it concerns decisions that have a lasting effect, like national elections, but also when in a small social group, people have a tendency to hide their opinion when they perceive not to be in the majority. A large number of similar questions, related to small group discussions or other expressions of opinion like campaigning or helping some people more than others or flattening someone's tires, gave similar results - respondents are more likely to express their opinion when they expect it will be well-received.

To see the power of the social environment in opinion formation, the experiments by Asch and Milgram are telling. In one of the experiments, participants were shown two lines and asked whether they think they have the same length. The participants would sit in a group of people most of whom would be collaborators of the investigators. When the respondents were asked separately, they would usually give the correct answer, but if first all other participants were asked, and they were given purposively the wrong answer, the respondent would often follow suit and give the wrong answer as well. In other words, despite the fact that the lines were obviously of different lengths, a respondent would follow the majority 'opinion'. A bandwagon style explanation that the respondent would want to be part of the winning group makes no sense in this context, while a fear of social isolation does.[8]

One aspect of this theory that should be emphasized is its forward looking nature. Individuals adapt their opinions, or at least their expressed opinions, to their perception of the majority opinion, so as to avoid social isolation

[7]The source of the survey is the Allensbach Archives, survey 2086/I+II. The number of respondents was 1011 who agreed with Brandt and 502 who disagreed.

[8]To confuse matters, Kuran (1995) refers to his own model as a 'latent bandwagon' model. In this section we maintain a clear difference between attempts to strategically support groups that are more likely to win (bandwagon) and the more natural tendency to avoid social exclusion either out of fear (spiral of silence) or with rational reputation costs in mind (preference falsification).

in the future. The theory is about adjusting opinions before elections take place or before joining a conversation, rather than expressing one opinion and then changing due to, for example, persuasion. When respondents in election surveys are asked to recall their vote in a previous election, the resulting vote distribution tends to show stronger support for the winners of the election than was actually visible in the elections. This demonstrates an adjustment of opinion retrospectively, whereby a supporter for the minority party recalls supporting the majority party. The spiral of silence, however, does not refer to this mechanism, but rather to the prospective version, where individuals fear a future isolation and adjust accordingly.

The general rational-choice literature on protest behavior primarily deals with the free-rider problem. If the participation of a large number of individuals is sufficient to bring about major political change, it would be irrational for any additional individual to waste the effort to participate as well. Even without this additional participation, and hence cost, the individual would get what it wants, regime change. One proposal for the solution of this paradox in this literature is the idea of a "collective rationality", whereby individuals adhere to a fiction of the necessity of full participation ("it would be bad not to participate, because if everybody would abstain, ...") which convinces them to participate. Finkel, Muller and Opp (1989) provide an extensive account of this approach. The basic twist made in the article is that the rational choice theories are compared to what individuals say to be rational or desirable in surveys. In other words, they demonstrate how individuals reason in a way that could be labeled collective rationality, even if it is not always rational from an individual perspective. Or, similarly but argued from a slightly different angle, they take a rational-choice approach but accept that rational decisions can be based on possible incorrect subjective information (Finkel and Muller 1998: 46).

Kuran (1989, 1990, 1991*a*,*b*, 1995) posits a model of what he calls preference falsification that, strikingly without reference to Noelle-Neumann (1993),[9] closely resembles the theory of the spiral of silence. Although he refers to rational-choice explanations and collective action problems instead of psychological explanations like Noelle-Neumann, his work is implicitly full of psychological mechanisms, and the resulting theory of preference falsification is in essence the same. Individuals in Kuran's model adjust their publicly expressed opinions according to the majority opinion they perceive. Whether one puts this in a perspective of comparing the calculated costs of expressing ones view to the rewards of doing so, or whether one describes this as fear of isolation and repercussions easily degenerates in a fruitless semantic discussion. Is someone who avoids the punishment of a social group acting out of fear of punishment, for example through social isolation (Noelle-Neumann 1993), or out of calculation that the costs (punishments) are higher than the rewards (Kuran 1991*b*)? Some even talk of a "psychic reward" (Finkel, Muller and Opp 1989; Muller, Dietz and Finkel 1991) to marry the two approaches. Although perhaps interesting from a psychological perspective, in terms of the social effect of the behavior the two are equivalent.

Kuran (1991*a*) makes a distinction between private and public preferences, whereby the former are relatively stable and whereby the latter are determined by both the private preferences and the cost or risk of expressing these. Although the models of Noelle-Neumann (1993) and Kuran (1995) provide an extensive attempt to understand the mechanism by which the balance between minority and majority positions of public opinion changes,

[9]This is even the more striking considering that it is clear from footnotes in Kuran (1991*b*) that the two authors were in fact in touch with each other about opinion polls in Germany, with Noelle-Neumann in her role as director of the Allensbach Institute.

the books only sparsely discuss the formation of these attitudes in the first place. In their model the preferences are thus exogenous and fixed. In the key model of this thesis, the private preferences are what are explained by the diffusion process itself, by the social judgment theory, while it is the interaction of these preferences with the costs of protest that determine the operation of the revolutionary cascade. The idea is that the stronger someone's preference, the more difficult it will be for the individual involved to repress this preference and to avoid expressing it.

3.2.2 Preferences and cascades

> "After all, a mass uprising results from multitudes of individual choices to participate in a movement for change; there is no actor named 'the crowd' or 'the opposition'." (Kuran 1991*b*: 16)

Whereas Noelle-Neumann (1993)'s model of the spiral of silence is primarily described in the context of Western democracies and elections, Kuran (1995) applies his model directly to the revolutions in Eastern Europe around 1989, and extrapolates subsequently to other revolutions, including the revolution in Iran of 1979-1980, the Russian revolution of 1917, the French Revolution, and the coup in Germany by the national-socialists (Kuran 1991*b*: 43-44). He thus suggests a close link between the mechanism of preference falsification and the occurance of public protest, through a mechanism that one could call the cascading revolution.

The "Orange Revolution" in Ukraine is a good example of this mechanism. In the winter of 2004 a mass of protesters camped out in Kiev to protest against the attempts by the authorities to help president Kuchma's selected successor to win the elections over the more popular Yushchenko (Bunce and Wolchik 2006). Interesting in this revolution is the striking discrepancy between the reputation of the Ukrainian population of being relatively apathetic and the sudden large protest movement in the streets of Kiev. Of course, pressure from outside and campaigns help, but the discrepancy between apathy and protesting in wintry Kiev for days in a row seems too large to be explained by campaigning alone. A combination of two factors is likely to largely explain this phenomenon. The first factor would be that the little change in attitude as a result of the campaigning might have been just that little bit needed to bring people over a threshold from not protesting to protesting. In other words, their attitude was already very close to that of the protesters, but just needed that tiny little push. The second factor is probably even more important. For those people that had pro-democratic attitudes but were just not past the threshold to protest, their reluctance to protest will have become significantly lower once they saw large numbers of people on the street. Suddenly, they had somewhat less to fear from the authorities, as they would not be standing there on their own in the streets, but in a crowd, and suddenly they knew that they were not the odd exception, but that they had the support of many people in their country. Thus more people started to protest, and the more there were, the more those with a slightly higher threshold felt safe enough to go on the streets as well. This is the mechanism of the cascading revolution (Granovetter 1978; Kuran 1991*a*, 1995; Lohmann 1994).

The effect of the cascading revolution can be that a small change in preference of only a few individuals can suddenly create a cascade of protest, whereby a hitherto hidden distribution of preferences is suddenly put in a completely different light. Hence the surprise reaction to the revolutions in Eastern

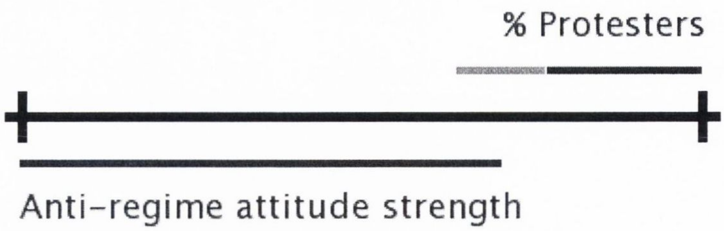

Figure 3.1: Triggering a cascade

Europe in the late eighties and early nineties (Kuran 1991*a*). Figure 3.1 helps to illustrate the mechanism. Imagine one could describe the strength of the attitude towards democracy on a one-dimensional scale of 0 to 100. Assume furthermore that the political regime under which a particular individual lives is either a democracy or an autocracy and that a certain percentage of the overall population is currently protesting against the regime. The cascading model of revolution argues that the likelihood of joining the protest, given the size of the current protest, is directly proportional to the strength of the attitude itself. In other words, if an individual has an attitude in favor of democracy of 80 and lives in a non-democratic regime, then at least 20% of the population would already have to be protesting for this individual to join the protest. The stronger this attitude, the lower this threshold - thus an attitude of 90 would mean that only 10% would already have to be on the streets. And the reverse, if the attitude is only 60, 40% of the population would have to be protesting before the individual feels concerned enough about social exclusion or future reputation costs to join in. At the other extreme, when an individual is in fact quite supportive of the authoritarian regime, with say a pro-democraticness score of only 10, this individual will still join the protest when 90% of the population is already publicly opposing the regime. In this scenario, the individual, while largely supportive of the regime, starts to take into account that under such large protests the regime is unlikely to survive and thus starts to be concerned with his or her post-revolution reputation and social status. The action of protesting thus only indirectly reflects the attitude towards the regime and is also contingent on the actions of others in the same community. One can now imagine a situation whereby many individuals are fairly strongly opposed to the regime, but nobody strongly enough to initiate the first protests. Once a small change in opinion takes place, however, a few people might start the protests, which in turns triggers those with slightly milder oppositional attitudes to join, etcetera.[10] A small change in public opinion can thus have a dramatic effect in observable actions, which could explain the discrepancy in the Ukraine between the initially observed apathy and the subsequent consistent protests in Kiev.

In this thesis we have a very local perspective on the mechanism of the cascading revolution. While attitudes cross international borders through communication and norm diffusion, protest behavior itself only 'diffuses' within countries or even subnational units, 'provinces'. An alternative interpretation could be a mechanism whereby latent protesters feel strengthened by observing their foreign comrades undertaking public action, gaining a sense of "universal camaraderie even in the face of significant domestic opposition" (Ramirez, Soysal and Shanahan 1997: 737).

3.2.3 Revolutions

Having traveled from the social-psychological literature on attitudes to electoral studies in Germany to one of the most abstract models of revolution

[10]See Kuran (1989) for the clearest exposition of this mechanism, on which this paragraph is based.

available, it is time to revert back to the more familiar democratization literature in political science.[11] As has been stated above, most theories of democratization focus either on the behavior of elites exclusively, or on structural factors that either form a breeding ground or an obstacle to democracy, or on the strength of particular classes in society or the extend to which citizens are organized independent of the state. Very little attention is paid to the extent to which general public opinion matters for democratization or for the survival of new and old democracies. "If not completely ignored, mass attitudes are either considered mere reflections of a society's structural properties or they are declared irrelevant for the elites' institutional choices" (Welzel 2006: 873). The focus in this thesis is entirely on the mass level in this complicated process. It is about individual attitudes and their relation to major political developments. In this section we will discuss somewhat more extensively the literature on mass-based political regime changes, or revolutions.

It should be made very clear that this thesis does not by any means attempt to argue that elites or structural factors do not matter for political transitions. It is undeniably the case that the actions of key players in the process can have dramatic effects on the outcome of processes of major so-

[11]The more appropriate name is of course 'transition studies', since this refers to transitions in both directions, democratization and democratic breakdown. The more common nomer of 'democratization studies' quickly leads to an unfounded optimism that most transitions that involve the breakdown of an autocratic regime will by default lead to a democratic outcome. This optimism has been most visible with regards to the study of regime change in Russia since 1991, where overly optimistic predictions have been shown to be unfounded retrospectively. A downside of the use of the term 'transition studies', however, is that at times this is interpreted as a particular section of democratization studies, namely those theories that focus on agents and elite behavior instead of structural preconditions or cultural values in the explanation of democratization (see, e.g., Grugel 2002). In this thesis we will follow suit and use the term 'democratization studies' or 'democratization literature', while acknowledging that this terminology is somewhat misleading in that it regularly refers to literature on transitions in either direction.

cial change.[12] Furthermore, it can probably be argued that for some major shifts in power, the general attitude towards the regime among the general population was utterly irrelevant. The only claim, more or less made by assumption, made in this thesis in this regard boils down to stating that for at least a substantial number of political regime changes, public opinion matters. It does not claim that this is the case for all transitions, nor does it claim that public opinion is the only important factor. It would include what is conventionally known as revolutions - popular based mass movements that overthrow the political regime - but also elite takeovers that are stimulated by the legitimizing power of public support.

Of the latter, perhaps the collapse of Suharto's regime in Indonesia in 1998 forms a good example. It was a smooth transition whereby members of the elite took over power from other members of the elite, without any violent coup. Suharto simply handed over power. The basis of this transition was the clearly visible gradual deterioration of public support for Suharto's regime, however. Whereas previously any public negative comments against the regime were more or less unheard of, now even high ranking members of society were explicitly suggesting to Suharto that he should step aside. Protests in particular by students strongly influenced the decision that the necessary legitimacy of the regime was waning (Vatikiotis 1998: 218-232), in particular thanks to the "intolerably high, and rising, level of government corruption" (Ross 2004: 236). This transition can by no means be labeled a revolution in the conventional interpretation of the term, but neither can the important role of mass public opinion, visible primarily in the intensifying student protests, be denied.

The democratization in Spain in the late seventies is generally provided

[12]See Tolstoy (1993) for a brilliant and compelling argument to the contrary, however.

as the typical example of a "pacted transition" (Burton and Higley 1987; Higley and Burton 1989; Bermeo 1997). Bermeo makes a strong argument, however, that while the transition itself took place through an elite pact, "1976 was also a year of widespread violence and unceasing mobilization (...) violence was much more pervasive in Spain than in "revolutionary" Portugal" (Bermeo 1997: 309). The actions of the members of the elite cannot be seen as entirely separate from popular pressure for reform. Instead, the elite pact that settled the transition should be seen as a measure to keep radical elements in the popular protests in check. Given that the more extreme elements had too little electoral support to upset a post-transition regime, the elites calculated that a pact would maintain a sustainable regime, while giving in to popular pressure for more democratization (Bermeo 1997). Similarly, in El Salvador and South Africa, "sustained unrest eventually persuaded elites that the costs of repression were too high and that negotiations with the insurgent counter-elite were therefore in their interest" (Wood 2000: 198). The negotiations were an elite affair, but the popular unrest cannot be ignored when attempting to understand these transitions.

Welzel (2006) makes a strong case that democracy aspirations of the general public are an important factor in processes of democratization, despite being ignored in the major paradigms of transition studies, structural, elite-choice, and cultural theories. In both structural and elite-choice theories of democratization, public opinion towards democracy can be seen as constraining factors. Dictatorial regimes with a strong public support for democracy find it more difficult to repress protests than do regimes where such support is less widespread. The area in transition studies where one would expect most evidence of a positive relationship between public opinion and democratization is that of political culture, but here most attention is paid to civic culture and social capital, rather than to explicit support for democracy or civil and political liberties themselves (Welzel 2006: 888). Welzel shows, on the basis of Freedom House scores of political and civil liberties and World Value Survey scores on attitudes towards these, that such attitudes clearly outperform both structural and cultural explanations of democracy.

In the remainder of this thesis the term revolution will be used in an unconventional manner to include all regime transitions where public opinion played a key role and it will be juxtaposed to coups, which covers all regime transitions where public opinion was irrelevant. The latter is seen as an "error term" in the model - it covers all transitions for which the theoretical model provided in this thesis provides absolutely no explanation. In Kuran's work, "[t]he term revolution is used (...) in a narrow sense to denote a mass-supported seizure of political power that aims to transform the social order. By this definition it is immaterial whether the accomplished transfer of power brings about significant social change" (Kuran 1991*b*: 13). The concept is thus further widened in this thesis to encompass all mass-supported, either through active involvement or through passive legitimizing support of elite actions, seizures of political power, with or without the aim to dramatically change the social order outside of the political regime itself. The seizure of political power is a key aspect, however, in line with the "broader and more contemporary definition of revolution: an effort to transform the political institutions and the justifications for political authority in a society, accompanied by formal or informal mass mobilization and non-institutionalized actions that undermine existing authorities. (...) this definition is strong enough to exclude coups, revolts, civil wars, and rebellions that make no effort to transform institutions or the justification for authority." (Goldstone 2001: 142). In contrast, many existing definitions of revolutions are slightly wider in scope in that a change of political regime is not necessary an effect, but narrower in type of action, in that it involved, e.g., more "violent civil disturbances" (Davies 1962: 6, note 3).

Because of the broad interpretation of the term revolution and wide range of types of transitions that would fall under the nomer of mass-supported seizures of power, some of the critique on such explanations of transitions can rather easily be brushed aside. For example, McFaul argues that "[p]acted transitions are elite affairs; mobilized masses spoil the party. Jacobins must therefore be side-lined, for if they are part of the equation, democracy is less likely to result" (McFaul 2002: 218). "To date, no stable political democracy has resulted from regimes transitions in which mass actors have gained control even momentarily over traditional ruling classes" (Karl 1990: 8). The requirement that "mass actors" gain control over members of the elite is, however, not a requirement for a transition to be classified as a revolution in the context of this theoretical framework. The transition in Indonesia is a good example of a transition where a member of the political elite - the vice president no less - seizes power partly on the basis of the substantial popular support for a transition, evidenced in the many protests at the time. It is not a mass actor, but an elite actor that takes over power and moves the regime in the direction of a more democratic one, but it is still based on popular support, on a changing public opinion towards the political regime.

The conceptualization suggested above appears to be an "easy way out" in terms of defining revolutions in such a way that many dramatic regime changes in history can be relabeled as revolutions and made relevant to this thesis. For the purposes of this thesis, this is a reasonable step to take, however. No theoretical model can explain all aspects of all revolutions, or even of some of them. Many factors, international and domestic, agent-specific and system-specific, typical and particular, and elite-based and mass-based, interact in bringing about major political shifts like the one analyzed here. In this regard, this thesis is entirely in line with the comment that "structuralism and individualism are not rival and mutually incompatible approaches to the study of revolution, as Skocpol would have it. They are essential components of a single story" (Kuran 1991*b*: 22).[13] Separate theoretical models, however, can shed lights on particular aspects of these changes that do occur with some regularity and one of these aspects is the role of popular support for the regime change. Furthermore, the key goal here is to explain the geographic and temporal clustering we observe in democratic regimes. This clustering might be caused by one particular aspect of the democratization process, or it might be caused by a number of different factors. The aim of this thesis is to demonstrate that one particular aspect, the dynamics of public opinion formation, could play a role in the generation of the geographical and temporal clustering we observe. Independent of the clustering aspect, it still remains fruitful to further study how exactly the public opinion dynamics relate to shifts in political power and under what circumstances public opinion plays an important role and under what circumstances a less important one. This will be left to further research, discussed in §6.2.2.

3.3 Conclusion

This chapter went in large steps through a very large literature on attitude formation, on attitude diffusion, on attitude revelation or falsification, on the relation between attitudes and behavior, on the cascading character of

[13]In fact, it could be argued that agent-based models, like the one applied here, are particularly suited for linking structural and individual components of complex social dynamics. The extent to which individual behavior affects the global system, which in turn affects individual behavior can be explicitly modeled and studied with such a tool.

popular protests and finally on the role of mass level attitudes in processes of democratization. Attitudes have been conceptualized as an evaluative opinion towards a particular object, which can be altered through interpersonal communication. The object in the case of the model presented here is democracy, or could also be interpreted as the action of protesting in favor of democracy. This attitude was subsequently linked to behavior through the cascading model of revolution, which argues that the likelihood of an individual to join a protest, or more generally to publicly oppose the regime, is a function of the strength of the attitude towards democracy as well as of the extend to which other actors in the same system are currently showing similar public oppositional behavior. The sheer fact that few people are protesting can thus not directly be interpreted as full support for the regime - the lack of protesting behavior itself discourages latent protesters to take action. Finally, it was argued that the fact that public opinion is assumed to matter in regime transitions is a very weak claim. Not only full popular revolutions would fall under this category of regime changes, but also transitions whereby elites make decisions on the basis of their perceived popular support. After this discussion of the many underlying building blocks of the theoretical model of this dissertation, the next chapter will turn to the detailed description of the actual model itself.

CHAPTER 4

An agent-based model

An agent-based model is a computer simulation to perform the equivalent of a thought experiment (Holland 1995: 156), where a large number of agents interact on the basis of a few relatively simple rules. Whereas game theory is usually a solid approach to understand the outcome or dynamics of games with few actors, the results of large numbers of actors that interact with each other and where the actions of one actor affect that of all other actors are generally difficult to trace analytically. Computer simulations can help to understand the dynamics of such models. In an agent-based simulation the rules of behavior are usually simple and there are few types of different actors. While the rules are simple, the resulting patterns in the system as a whole can be highly complex and often surprising given the rules of interaction, hence the term *emergence* (Holland 1998; Johnson 2001). Like game theory, agent-based modeling is thus a purely theoretical exercise. The occasional attempt to compare computational models with experiments is misleading in this regard (Shadwick 2007). Agent-based simulations generally need lengthy computations and thus have become popular only after the widespread availability of computing power. Although early examples exist, most notably Schelling (1978), most applications are of more recent date. Examples of such models in political science are a model of democratic survival and geographic clustering (Cederman and Gleditsch 2004), models of cooperation (Axelrod 1997*a*), a model of secessionism in multi-cultural states (Lustick, Miodownik and Eidelson 2004), a model endogenizing the international state system (Cederman 1997), and a model of policy and party competition (Laver 2005).

In relation to more common formal models of human behavior, agent-based models do not only provide a methodological alternative for models that are intractable due to the nonlinear dynamics, but also make a different assumption about the kind of decision making that underlies human behavior. When analyzing a formal model, the usual assumption is that the actors involved are making calculations similar to those the researcher makes in drawing conclusions from the model. Not only the researcher, also the actor is forward looking. This line of reasoning can then be extrapolated to agent-based models, where just as the outcomes of the model are to some extent unpredictable for the researcher, so are the outcomes for the actors involved. Actors in a social system are, thanks to the nonlinear dynamics and complex patterns of social interactions, unable to predict the future and have to base their actions and decisions on more approximate heuristics. Simple reactions to a changing environment can generate adaptive behavior suitable for such an unpredictable reality, similar to the simple rules of behavior in an agent-based simulation. Thus, "intractability for real agents involves a substantive shift in the most plausible behavioral assumption about agents' decision-making inside the complex system - from deep strategic look-forward to adaptive learning" (Laver and Sergenti 2007: 1).

There is a strong tendency among agent-based modelers and other researchers working in the wider area of complex systems modeling to claim that their models are more "realistic" than other, linear models. The idea is that because it is easier to argue that the world is non-linear in nature than that it is linear - why would it be linear? - and that thus a non-linear model must be more likely to resemble the way the real world works. At times even a somewhat condescending tone can be heard: "The variables in realistic models interact in strongly nonlinear ways that give rise to the phenomena described here. Linear models are used regularly not because they are more accurate but because they are easier to handle mathematically. (Economists and physicists sometimes adopt the same research policy as the proverbial drunkard. Asked why he was looking for his keys under the street lamp when he lost them up the block, the drunkard replied that the light was better there.)" (Paulos 1995: 25-26). This reasoning appears to be based on the fallacy to assume that because the world is complex, only complex models can help in understanding this world. Generally, however, simpler models are more constructive in understanding the world than are complex models. With every added feature to a model, further assumptions are made about how the world works, with each creating additional possibilities to be wrong. Overtly complex models can easily lead to something akin to a computer game, whereby the output can be very interesting, but the relation to the real world more and more obfuscated. Simple models are preferred, and analytical solutions are more useful, but at times simple models lead to complex dynamics that can only be studied using computational models (DeAngelis and Rose 1992: 81-83), and this is the foundation on which computer simulations can help rather than distract from scientifically interesting questions.

As discussed extensively in the previous chapter, the model presented here is a bottom-up approach to explain the international diffusion of democracy. Regime transitions are explained from the perspective of the individual citizen, from the presence or lack of popular support among the general population for an existing political regime. Agent-based models are typical for the study of the relation between micro-level patterns of behavior and macro-level patterns such as the clustering of political regimes. Many theories concerning large social processes such as political revolutions are based on assumptions about individual behavior - such as the relative deprivation theory - but the overall argument remains located at a macro, social level. The exact mechanism through which the individual attitudes relate to the behavior of groups is rarely explicated. A model like the one presented here is excellent for making just this connection.

Because in the cascading model of revolution the actions of one actor depend on those of all other actors, the micro-level behavioral patterns lead to complex non-linear dynamics in the overall system. As mentioned above, agent-based models are particularly suited to analyze such non-linear dynamics. Through computer simulations we can see how the minimal assumptions about individual level behavior translates to system-wide patterns when many agents act with each other in the same environment. More typical empirical methods, such as econometric models, have serious difficulty capturing such dynamics, while more typical theoretical methods, such as game-theory, cannot deal with the complex dynamics of having large numbers of actors.

In general, these simulations are more useful in establishing under what conditions particular patterns *can* occur, rather than whether they *do* occur. The model is based on assumptions and it is always questionable whether these assumptions are correct. Empirical analysis can help to validate these claims, but never fully establish their correctness. What the computer simulation can do is to establish that under certain assumptions and certain parameter settings, particular patterns of system-wide outcomes are possible. In the case of the model presented here we are thus interested in under what conditions we do see the spatial, temporal, and spatio-temporal clustering patterns of democratization given the few assumptions we made about the diffusion of attitudes and the translation of these attitudes into protesting behavior and indirectly into regime changes. To study these patterns we will have to run many simulations across a large space of possible parameter configurations. The next chapter does exactly that.

In this chapter the different elements of the agent-based model will be discussed. To assist the reader in forming an intuition with regards to the building blocks of the simulation, two small agent-based models that describe only subparts of the overall model will be presented and analyzed. At the end of the chapter a schematic overview of the entire simulation is presented in Figure 4.15.

4.1 Creating a map of countries

Let us first turn to the creation and placement of the provinces. The provinces form the cells of what is commonly known as cellular automata of size W by H. Cellular automata are a grid of adjacent square cells which keep changing state using simple rules, on the basis of information from the previous state of the cell and the state of cells in what is called the Von Neumann neighborhood, the four cells directly adjacent.[1] The most famous example of a cellular automata is the Game of Life, which is a small set of very simple

[1] Figure 4.5 provides a visualization of the cellular automata underlying the model.

rules which leads to cyclic patterns, *perpetua mobilia*, and patterns far more complex than expected from the initial rules (Gardner 1983; Johnson 2001). Although cellular automata form the basic foundation and inspiration for this model, in line with a model of democratic diffusion by Cederman and Gleditsch (2004), the model does deviate on several fronts from regular cellular automata models. One aspect that is unusual for cellular automata is that the map of this model wraps around its borders. The cells at the edges of the map are directly adjacent to those on the opposite edge - similar to creating cellular automata on the surface of a torus, a toroidal lattice. This is common in computer games that are based on cells and that try to simulate the fact that the world is round.

Once the provinces have been created country borders are added to the map. Drawing borders grouping together certain provinces is of course a clear deviation from any common cellular automata. The country borders are created by an algorithm where countries 'conquer' neighboring provinces which become part of the country of the conquering province, unless this leads to a fragmented country that the province is originally from. The algorithm runs through a number of iterations, determined by the size of the map ($W \times H$) and a configurable multiplier (M), resulting in $W \times H \times M$ iterations. Each iteration a random pair of two neighboring provinces is selected, which we will refer to as P_1 and P_2. If the two provinces are not part of the same country ($C_{P_1} \neq C_{P_2}$), the first country (C_{P_1}) will conquer the second province (P_2), unless this leads to a disconnected country of which P_2 was originally part. This algorithm results in a somewhat realistic looking map (see Figure 4.5), with varying forms and sizes of countries, and is derived from the model of Cederman and Gleditsch (2004). Their model models wars between countries, whereas in this model these wars are only applied to the setup stage to create the borders. Each country C is subsequently assigned a random level of isolation, $\varphi_0 \sim N_{[0,100]}(\phi, \sigma_\phi)$.[2] With a probability π, the country is set to be a democracy ($\Omega = 1$), otherwise it is set to be an autocracy ($\Omega = 0$). Randomly one of the provinces of the country is assigned as the capital.

It should be pointed out that this implies a blatantly unrealistic assumption of a static state system. Once the country borders in the simulation have been determined, as part of the setup of the model run, they will not change. The real world data, which we are trying to model, shows of course significant changes over time, both in terms of the number of states in the system and in terms of the locations of borders. To endogenize the formation of states makes sense in a model that studies democratic diffusion as a side-effect in the patterns of international conflict, as in the model of Cederman and Gleditsch (2004). In a model that is based on the diffusion of norms among individual citizens, however, such an endogenization would require too many additional mechanisms and parameters, which would require additional potentially incorrect assumptions about the world. The analytic insights from such a model would be too distant from any empirical behavior. To study the effect of a changing state system on the patterns found in this thesis will be left for later research.

4.1.1 Countries as a social network

The analysis of a process of diffusion across a set of geographical units that are connected by territorial borders, the so-called *object view* or *lattice data* approach to spatial patterns (Anselin 2002: 255), closely resembles the anal-

[2]Throughout this paper, $N_{[a,b]}(c, d)$ is a draw from a normal distribution, with mean c, standard deviation d and truncated to the interval $[a, b]$.

ysis of social networks, and indeed networks in general. Abstracting slightly from the map of countries, one can easily see how this can be represented as a set of vertices (the countries), which are connected by ties (the country borders).[3] Figure 4.1 does exactly this. Once we acknowledge that territorial units in this type of analysis indeed resemble a network, we can also compare the randomly generated map to both the map of the real world and to other randomly generated networks, as well as the latter two with each other, in terms of common network characteristics. Comparing network statistics between these three networks allows us both to validate our random algorithm in terms of its resemblance to the real world network, and to investigate the extent to which typical abstract models of diffusion analyzed on artificial networks (see, e.g., Cowan and Jonard 2004), might or might not apply to patterns of diffusion between countries. Similar to how types of random networks have been compared to typical social networks (see, e.g., Watts 1999), we can compare artificial networks to the structure of the international network of polities.

For example, in the study of social networks, the likelihood that a person will have a tie with another person is not based on a random selection of the world population, but on circumstances as mutual acquaintances, work relations, etc. Thus the chances that two people connect to each other are higher when they live in close geographic proximity or when they have mutual ties. Because of the effect of mutual ties, individuals with more ties are more likely to get new ones than are those with fewer ties, creating a power law distribution of the number of ties per individual (Barabási and Albert 1999; Kang 2007). The result is a network that is locally strongly clustered - friends

[3] In this paper the common terminology for networks will be used where the units or points are referred to as vertices, which are connected through edges or ties to other vertices.

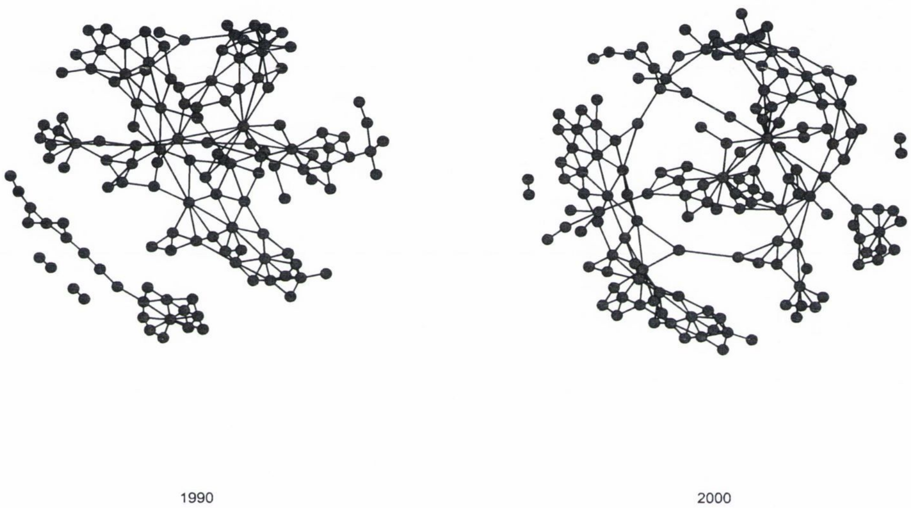

Figure 4.1: The world as a network in 1990 and 2000. Vertices and edges represent countries and land borders, respectively. The country with the black dot is the United States, the disconnected countries in 2000 are the islands around Micronesia.

are likely to be friends of each other - while being globally more separated with only weak ties between different clusters. These weak ties - weak in the sense that they are not 'reinforced' by different short trajectories between the same individuals - are crucial in the study of social diffusion - or, in fact, other forms of contagion - because they are necessary for the transmission of information from one closely connected cluster to another (Granovetter 1973; Watts 1999). The extent to which the international states system resembles such a network would shed light on the applicability of these theories on international patterns of diffusion.[4]

The effect of this type of network is that there is a combination of structure, due to the way the network forms and due to geographical factors, while there is also substantial randomness, some unexpected connections between individuals. Such a combination of structure and randomness can, under certain circumstances, lead to the phenomenon of a small-world network, a network that is locally strongly clustered, but where the distance between two random individuals in terms of the number of ties needed to cross to reach one from the other is never very large (Watts and Strogatz 1998; Watts 1999). In the real world this is rumored to be only six steps, six degrees of freedom that separate any individual from any randomly selected other individual, on average (Watts 2003).

Below we will make a comparison between the random network generated by the 'conquer' algorithm (henceforth *conquer-network*), more simple random networks (*Bernoulli-network*), and the real world network of countries

[4]It should be emphasized that this section is about the network of countries and its effect on diffusion. In this thesis, two different types of networks are overlapping each other - the network of countries and the social network of individual citizens. The latter will be abstracted from in the main model (see §4.3.1) and only in this paragraph will we talk about the international state system as a (social) network, but the two should not be confused.

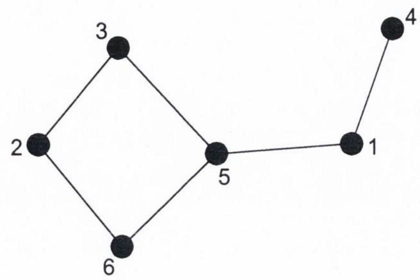

Figure 4.2: Simple network of six vertices

(*world-network*). The Bernoulli-networks are, as the name suggests, based on Bernoulli trials for each edge, where the probability of an edge between any two vertices is independent of any of the other edges.[5] For the world-network, the contiguity data is based on the Correlates of War project, using the land borders data in the Direct Contiguity dataset (version 3.0) for countries worldwide between 1816 and 2000 (Stinnett et al. 2002). It should be pointed out that the comparisons made below are all based on characteristics of static networks. As mentioned previously, the map of the world in the main simulation is a static one which does not evolve over time, and for the comparisons below snapshots will be used of the world-network, rather than the dynamic system.

The network characteristics presented below are based on statistics that

[5]These networks are generated using the *rgraph* command in the **sna** package for R (Butts 2007). This method of generating the random network is particularly distinct from the scale-free random networks of Barabási and Albert (1999). The idea that connected vertices have a higher probability of acquiring more edges is certainly plausible for social or many engineered networks, but makes little sense in a network of the international state system.

describe individual vertices in a network, which are subsequently aggregated to give a description of the network as a whole. Perhaps the most interesting to compare different networks, especially in the context of norm diffusion, would be measures of centrality. Centrality refers to the position of an individual vertex in the network as a whole, which, depending on the particular type of centrality, can have strong influences on the importance of the vertex in communication, epidemiological processes, or power projection,[6] among other things. The first and most basic measure of centrality is based on the *degree* of a vertex. The degree refers to the number of vertices the vertex is connected to through a single edge. Figure 4.2 provides a simple small network that will illustrate the concepts used here. In this network, vertex 5 has the highest degree of 3, while vertex 4 has the lowest degree of 1. Although various different definitions of degree centrality exist, the most commonly applied and most straightforward is the one developed by Freeman (1979). In his analysis, the degree centrality is simply the degree itself, possibly divided by $n-1$ to normalize over network size. Freeman describes how degree can be seen as a measure of centrality:

> "With respect to communication, a point with relatively high degree is somehow 'in the thick of things'. (...) As the process of communication goes on in a social network, a person who is in a position that permits direct contact with many others should begin to see himself and be seen by those others as a major channel of information. In some sense he is a focal point of communication, at least with respect to the others with whom he is in contact, and he is likely to develop a sense of being in the mainstream

[6]For example, Hage and Harary (1995) provide an interesting example whereby the choice of the location of the capital by chiefs on the Marshall Islands can be analyzed from the perspective of network centrality.

of information flow in the network" (Freeman 1979: 219-220).

In addition to degree centrality, we will be looking at the concepts of *betweenness centrality* and *closeness centrality.*[7] The *geodesic* of two points on a network refers to the shortest distance between the two vertices. For example, the geodesic of points 1 and 6 in Figure 4.2 is 2, that between vertices 3 and 5 is 1 and that between 2 and 1 is 3. The betweenness centrality then refers to the proportion of geodesics that cross a particular vertex.

$$C_B(p_k) = \sum_{i}^{n}\sum_{j}^{n}\frac{g_{ij}(p_k)}{g_{ij}} \qquad i < j, \tag{4.1}$$

whereby C_B refers to the betweenness centrality, g_{ij} is the number of geodesics between vertices i and j, and $g_{ij}(p_k)$ the number of those geodesics between i and j that include vertex p_k. Since the scale here is dependent on the size of the network (Butts 2006), this can be normalized as follows:

$$C'_B(p_k) = \frac{2}{n^2 - 3n + 2}\sum_{i}^{n}\sum_{j}^{n}\frac{g_{ij}(p_k)}{g_{ij}} \qquad i < j, \tag{4.2}$$

where n is the number of vertices in the network (Freeman 1979: 223-224). In a social network, the person located on the geodesics between two other members of the network is considered to be in a strategic position in terms of communication. This person "can influence the group by withholding or distorting information in transmission" (Freeman 1979: 221), or such a node "has a capacity to facilitate or limit interaction between the nodes it

[7]There are many other measures of centrality. One interesting concept discussed by Bonacich (2007) is eigenvector centrality, which is a centrality measure somewhat comparable to Google's PageRank (Page et al. 1998; Page and Brin 1998), whereby the centrality of one vertice is positively correlated to the centrality of connected vertices. In other words, vertices with a high degree have a higher centrality when connected to other vertices with a high degree, compared to vertices with a similar degree, but that are connected to vertices with lower degree levels.

links" (Marsden 2002: 410). If we accept the idea that direct contiguity between countries in a diffusion network matters, then a situation where the interaction between two particular countries always has to go through a particular third country, strongly affects the influence this third country can have on the diffusion process between the former two.

A vertex in a network is located in a central position in closeness terms when the distances to other vertices are relatively short. Communication originating from this point have the lowest cost or highest efficiency in terms of reaching all other nodes (Freeman 1979: 225). Freeman (1979: 225-226) suggests a measure of closeness centrality, by simply taking the inverse of the sum of the lengths of all geodesics from a particular point. This measure of closeness, however, cannot be performed on a disconnected network. As can be seen in Figure 4.1, the real world is a disconnected network when only land borders are taken into account. An alternative measure is used here as proposed by Tallberg (2004: 210), which makes the slight modification of taking the sum of the inverses of the distances instead of the inverse of the sum. Standardized to reach a scale from zero to unity this leaves us with:

$$C_C(p_k) = \frac{1}{n-1} \sum_{i=1}^{n} \frac{1}{d(p_i, p_k)}, \tag{4.3}$$

whereby $n-1$ is the longest possible geodesic and $\frac{1}{\infty} = 0$.

While centrality refers to the position of a particular vertex in the network, we are interested in the characteristics of the network as a whole. One could take two approaches to go from the micro to the macro level in this case. One could look at the average level of centralization, but Freeman prefers another concept of network centrality, whereby "the centrality of an entire network should index the tendency of a single point to be more central than all other points in the network. Measures of a graph centrality of this type are based on differences between the centrality of the most central point and that of all others. Thus, they are indexes of the centralization of the network." (Freeman 1979: 227). If consider $C_X(p_i)$ to be any of the three measures of centrality of point i, and $C_X(p^*)$ to be the largest value of $C_X(p_i)$ across the network, then $\sum_{i=1}^{n}[C_X(p^*) - C_X(p_i)]$ is a reasonable measure of the extent to which the vertex with the highest centrality score stands out from the network as a whole. To correct for the size of the network, we correct for the maximum value this measure can take:

$$NC_X = \frac{\sum_{i=1}^{n}[C_X(p^*) - C_X(p_i)]}{max \sum_{i=1}^{n}[C_X(p^*) - C_X(p_i)]}, \tag{4.4}$$

where NC_x is the network centrality for measure X and $max \sum_{i=1}^{n}[C_X(p^*) - C_X(p_i)]$ the maximum sum of differences possible on a graph of size n (Freeman 1979: 227-228).[8] Table 4.1 shows the centrality and centralization measures for the example network in Figure 4.2.

Figure 4.3 compares the three types of networks on the basis of average degree and degree centralization. In terms of average degree across the network, it is striking that the variance in degree for a given number of countries is far larger for the Bernoulli-network than for the other two types of networks. Although the variance is slightly larger in the world-network for the years where there are a small number of countries, the conquer-network and the world-network follow each other very closely in this graph. In terms of

[8]See Freeman (1979) for a full derivation of the maximum value of the sum of the differences for the measures of centrality based on degree and betweenness. For the closeness centrality, the following derivation is used: the highest level of closeness is obtained in a network where one vertex connects to all other vertices and there are no direct edges between the other vertices (Freeman 1979: 229). The highest level of closeness in this case is $\frac{1}{n-1}(n-1)\frac{1}{1} = 1$. All the other vertices have closeness $\frac{1}{n-1}(1 + (n-2)\frac{1}{2}) = \frac{n}{2(n-1)}$. The sum of differences in this case is $(n-1)(1 - \frac{n}{2(n-1)}) = \frac{1}{2}n - 1$.

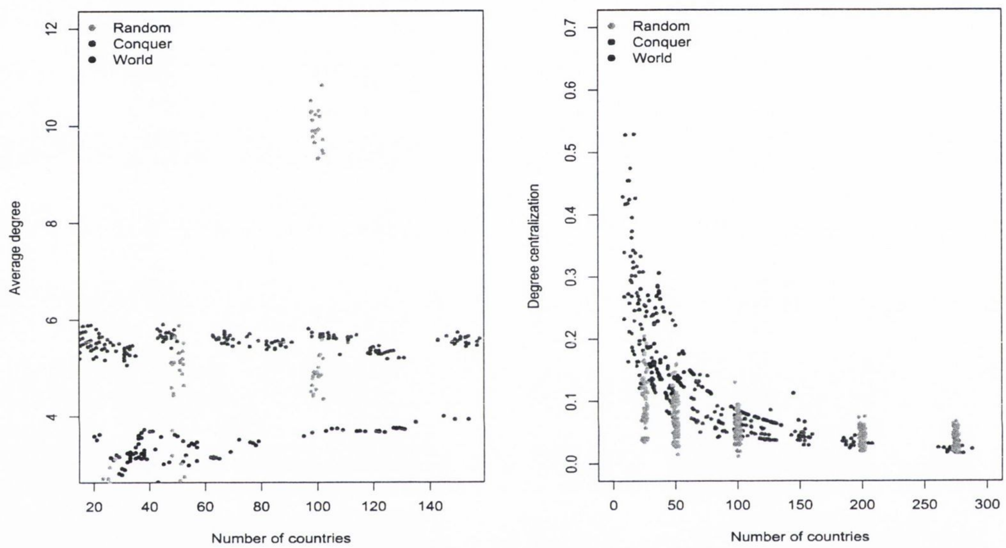

Figure 4.3: Average degree and degree centralization by the number of countries. The number of countries for the artificial networks vary by design. For the world-network, the number of countries varies over the years, in an ever increasing fashion.

Vertex	Degree	Betweenness	Closeness
1	0.40	0.40	0.67
2	0.40	0.05	0.62
3	0.40	0.15	0.67
4	0.20	0.00	0.48
5	0.60	0.65	0.80
6	0.40	0.15	0.67
Centralization	0.30	0.50	0.45

Table 4.1: Example of centrality and centralization measures. The vertices refer to those in Figure 4.2.

degree centralization, the pattern of the conquer-network is much closer to the world-network than the Bernoulli-network, although the world-network does show a consistently higher level of degree centralization.

Although the centralization measures based on the degree suggest that the conquer algorithms comes much closer to the world-network than does the Bernoulli-network, the centralization measures displayed in Figure 4.4 show less clear results. The level of centralization is generally higher in the world-network than in both artificial ones. In terms of closeness centralization, at least the relation between centralization and network size shows a similar pattern to the world-network. With betweenness centralization the level of centralization only changes in terms of variance, not in terms of expected value, while that for the world-network shows a remarkable erratic pattern. For the smaller networks, the conquer-network and the world-network are very close, while the Bernoulli-network shows a much higher variance. Overall we can conclude that the conquer-network gets slightly closer to the world-network than does the Bernoulli-network in terms of centralization by size.

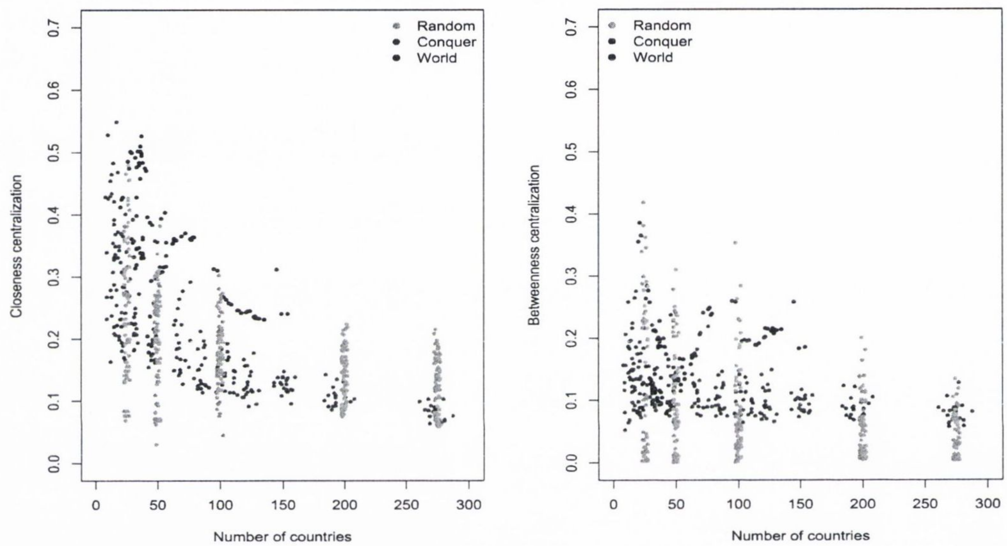

Figure 4.4: Closeness and betweenness centralization by the number of countries. The number of countries for the artificial networks vary by design. For the world-network, the number of countries varies over the years, in an ever increasing fashion.

4.2 Isolation and censorship

Different countries have a different susceptibility for outside information, ranging from strict control over information as in China, where the government closes Internet cafes for their 'dangerous influences', to countries with extensive cross-border traffic of individuals and many international sources of news and information, like most liberal democracies. Kopstein and Reilly (2000) use the concept of openness as part of their explanation of regime transformation. They regress the level of political and economic freedom on the level of openness of the regime. This approach is problematic, however, as the level of openness is very closely related to the type of regime in the first place. Generally, democratic or free regimes are more open to outside influence than are autocratic or restricted regimes, hence their measure of openness captures the effect of having a certain regime in the time period previous to the current rather than the effect of international influence. In other words, when a country is a democracy, it is likely to be open, and likely to be a democracy in the next time period, but this does not mean that this is a democracy due to its openness and international influence. Moreover, one would not expect a positive correlation per se between openness and democracy within a framework of diffusion, but rather a stronger effect of neighboring regimes in the case of a open regime, whether those neighbors are democratic or not.

Autocratic countries are likely to be concerned, to varying degrees, with the restriction of this information flow to their citizens. North Korea and China are prime examples, where access to the Internet is highly restricted and filtered for political purposes, as was the Soviet restriction on a large number of publications. These policies isolate citizens from foreign influences, including or primarily those that promote democratization. In this model, this enforced isolation of the citizens of non-democratic regimes limits the effects of broadcasting attempts by neighboring countries and lowers the chances of cross-border communication between individuals of the country and foreign individuals. Democratic regimes are assumed not to limit any international communication. One could argue that such limitations on free information would contradict the concept of democracy sufficiently not to consider such a country a democracy. In terms of implementation, the effects of broadcasting and isolation are what distinguish democracies from authoritarian regimes.

The level of isolation for each country is updated. It is reasonable to assume, perhaps even true by definition, that democracies do not limit the communication of their citizens with foreigners. For this reason, each country that is a democracy in this model resets the level of isolation to zero. For all other countries, the level of isolation is modeled as a straightforward random walk:

$$\varphi_{t+1} = \begin{cases} 0 & \text{if} \quad \Omega = 1 \\ \varphi_t + U | U \in \{-1, 0, 1\} & \text{otherwise,} \end{cases} \tag{4.5}$$

whereby φ_{t+1}, the level of isolation, is truncated to $[0, 100]$. With a random walk it is implicitly assumed that levels of suppression generally change over time, and that without other predictive variables, the best prediction of the level of isolation at time t is likely to be the level of isolation at time $t-1$.

4.3 Communication and persuasion

The second concept that is crucial to the simulation is that of an individual's attitude towards democracy. The idea is that any person has a particular attitude towards the concept of democracy, on a scale from strong support for democracy to strong opposition to democracy. In reality, it is unlikely that such a scale exists within someone's political outlook. Rather, the attitude towards one's own current political regime and that towards the concept of democracy in general is likely to be a complex combination of a multitude of different attitudes, expectations, experiences, and beliefs. The intricacies of such psychological and ideological preferences are assumed to be of little relevance to the overall pattern of democratic diffusion, and a relatively simple scale should therefore suffice. The attitude towards democracy scale measures the actual attitude towards the regime of an individual, rather than the demonstrated preferences. There is thus a strong distinction between attitude and resulting behavior; a distinction that is empirically difficult to discern (Ajzen and Fishbein 1977) but in a model straightforward to implement.

For each province a random number of citizens is set ($N_{citizens}$), with an initial attitude towards democracy (α_i). Subsequently, the threshold values of the social judgment model are assigned (t_i and u_i). By default, a citizen is not protesting, $\psi_i = 0$.

$$\begin{aligned} N_{citizens} &\sim N_{[1,\infty)}(C, \sigma_C) \\ \alpha_i &\sim N_{[0,\lambda-1]}(A, \sigma_A) \\ t_i &\sim N_{[0,\infty)}(T, \sigma_T) \\ u_i &\sim N_{[0,t_i]}(U, \sigma_U) \end{aligned}$$

In the above, C and σ_C are the mean and standard deviation for the initial values of the population sizes of each province; α_i is the initial attitude towards democracy for the individual citizen, set using a normal distribution with mean A and standard deviation σ_A; t_i is the latitude of rejection of the individual, or the outgroup threshold for communication, which is set randomly with mean T and standard deviation σ_T; u_i is the latitude of acceptance, or the ingroup threshold for communication, set with mean U and standard deviation σ_U.

Through communication, the third concept underlying the model, these individuals change their attitudes. By talking to others about democracy and about their ideological outlook on the world, one can gradually change one's own opinion towards democracy. This communication takes most likely place between citizens of the same country and to a lesser extent between randomly selected citizens of neighboring countries. The underlying assumption is thus that geographical distance matters for the frequency of interpersonal contact. The direction of this change, according to the social judgment theory, depends on the similarity between the two individuals at the outset. Individuals that have opinions very similar to one another are likely to refine their attitudes through the interaction and to move closer towards each other in terms of their attitudes and beliefs, while individuals with very different attitudes will diverge even more.

$N_{citizens}/10$ times a random citizen (S) is selected to initiate communication. The probability for each of the four provinces P in the Von Neumann neighborhood that a citizen will be targeted from this province is:

$$Pr(P) = \begin{cases} \tau/4 & \text{if} \quad C_P = C_S \\ \frac{\tau \max(\varphi_{C_S}, \varphi_{C_R})}{400} & \text{otherwise.} \end{cases} \tag{4.6}$$

The maximum level of isolation between the two countries is taken, as it is assumed that what really matters for communication to occur is whether the more restricted of the two countries can be reached.[9] If a neighbor-

[9] τ is divided by four because there are four neighbors in the Von Neumann neighbor-

ing province is selected, a citizen (R) will be randomly selected from this province, otherwise this will be done from the province of S. Once a sending (S) and a receiving (R) citizen have been selected, given that their attitudes towards democracy differ, the attitude of R is updated in line with the social judgment model of communication:

$$\alpha_R = \begin{cases} \alpha_R + sign(\alpha_S - \alpha_R) \times \delta & \text{if} \quad |\alpha_S - \alpha_R| < u_R \\ \alpha_R - sign(\alpha_S - \alpha_R) \times \delta & \text{if} \quad |\alpha_S - \alpha_R| > t_R \\ \alpha_R & \text{otherwise,} \end{cases} \tag{4.7}$$

whereby α_R is truncated to $[0, \lambda - 1]$. The sign function operates as follows:

$$sign(x) = \begin{cases} +1 & \text{if} \quad x \geq 0 \\ -1 & \text{if} \quad x < 0 \end{cases} \tag{4.8}$$

The attitude toward democracy of the receiver, α_R, is thus increased by the configurable size of the communication effect, δ, when the attitude towards democracy of the sender, α_S, is larger than that of the receiver, but within the latitude of acceptance of the receiver, u_R. It is decreased by δ when α_S is larger than α_R, but outside the latitude of rejection, t_R, thus generating a counter reaction. The reverse also holds: α_R is decreased when α_S is smaller than α_R and within the latitude of acceptance and increased when this is the case with α_S being outside the receiver's lattitude of rejection.

The assumption that proximity matters for the strength of social ties is intuitive and common in the literature on diffusion (see, e.g., Gleditsch 2002: 3-6). The most clear statement along those lines is the so-called first law of geography: "Everything is related to everything else, but near things are

hood. The division by 100 is because φ is scaled from 0 to 100.

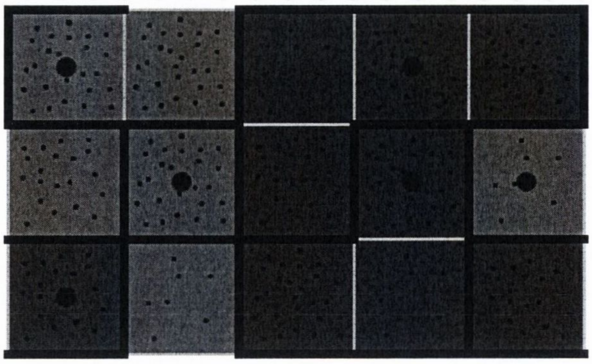

Figure 4.5: Example setup of the model ($W = 5$, $H = 3$). Grey lines denote the province borders; black lines the country borders. Large dots denote the capitals; small dots are individual citizens (in the actual simulation, they do not have a precise location other than being part of a province). The map wraps around both borders. The color represents the regime type (light is democratic, dark is autocratic).

more related than distant things" (Tobler 1970: 236). Mok, Wellman and Basu (2007) provide an extensive overview of the literature on the assumption that proximity matters for interpersonal ties, as well as an empirical analysis of the effect of distance on social networks in a Toronto neighborhood. Faust et al. (1999) relate the social ties between Thai villages in terms of their shared use of schools, temples, labor, and tractors, to the geographical location of these villages. Taking both the Euclidean distance and the characteristics of the geography into account, they find preliminary evidence that proximity matters for social ties. They also find that villages that are closer to the district border have a higher chance of having labor or school ties with villages outside the district, which provides slight additional support for the way the democracy promotion mechanism has been implemented in this model (see §4.5).

4.3.1 A simulation of attitude diffusion

To establish some intuition concerning an important element of the main simulation model, this section describes a model of communication where a set of individuals contacts other individuals randomly and adjusts the attitudes according to the social judgment theory. In this mini simulation presented here we will replicate the analysis of Jager and Amblard (2004) and attempt to improve our intuition with regards to this model by varying various parameters.

The configurable parameters of the model in equation 4.7 are δ, the initial values of α, and the individual threshold values u and t. In this simplified version of the model there are no countries, provinces, or regimes. We are talking only about a set of agents communicating randomly with each other. Appendix C contains the implementation code of this simple model.

Jager and Amblard (2004)'s main claim is that they find uniformity, bipolarization, and pluriformity depending on the distribution of the parameters u and t. Jager and Amblard use identical values of u and t for all agents in the model and then vary these values across different simulations. To replicate their results, the values presented in Table 4.2 are used.[10] Figures 4.6 through 4.9 show that the results from Table 4.2 are replicated in this analysis.[11] The figures are produced as snapshots rather than the one used by Jager and Amblard, which show the dynamics over time. The snapshots are used because they in fact make the dynamics more clear in this case. Figure D.1, as well as Table D.1, show the level of convergence of this replication model, measured as the amount of change in α per 1000 iterations ($\Delta\alpha$) dur-

[10]The range of values for α in the Jager and Amblard (2004) model is $[-1, 1]$, while the one in these simulations is $[0, 100]$. The values of u and t are adjusted accordingly.

[11]Unless otherwise stated, for the analyses below, α_0 is randomly, uniformly distributed over the entire range $[0, 100]$ and $\delta = 1$. For a simulation exactly equivalent to that by Jager and Amblard (2004), a setting of $\delta = 5$ would be needed.

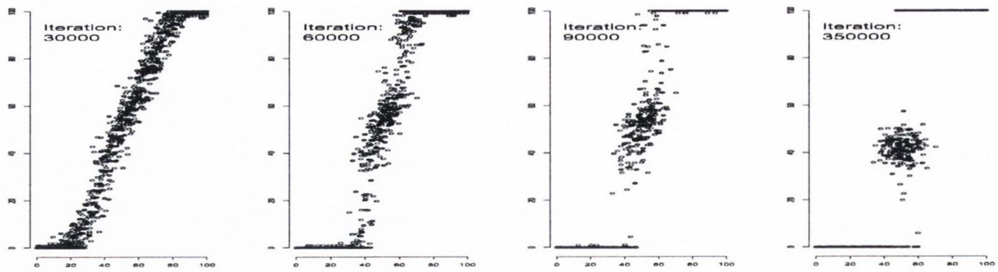

Figure 4.6: Replication of Jager and Amblard (2004). $u = 20$, $t = 30$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

ing the last fifth of the simulation (see Appendix A). These statistics show clearly that the simulations have been run long enough to reach a stable state. In some simulations there is still variation, but at a fairly constant level.[12]

u	t	Result
20	30	bipolarization
60	80	uniformity
30	60	pluriformity (3 clusters)
10	80	pluriformity (5 clusters)

Table 4.2: Replication parameters for Jager and Amblard (2004).

The simulation in Figure 4.6 shows largely a pattern of bipolarization, as found under similar conditions by Jager and Amblard. Except for the agents with an initial attitude at the center of the scale, there are always more agents in the latitude of rejection than in the latitude of acceptance, thus

[12]Note that these convergence statistics only show the amount of change in α. The fact that the change is constant does not necessarily imply that some stable state has been reached. The level of α might still be consistently moving. The combination with the simulation output in Figures 4.6 through 4.9 clearly suggests that stability has been reached, however.

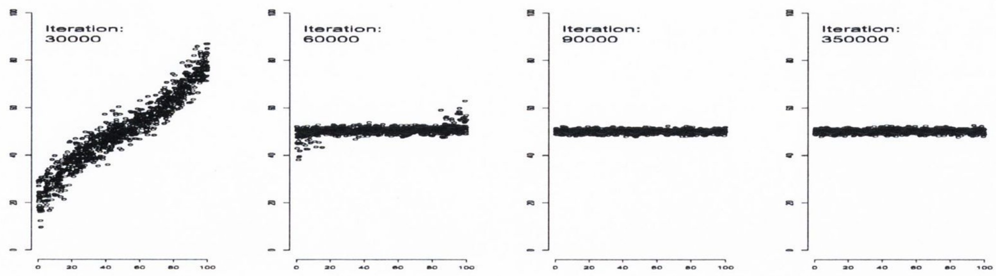

Figure 4.7: Replication of Jager and Amblard (2004). $u = 60$, $t = 80$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

opinions tend to diverge to the extremes. The center opinion holders more or less persist, but in an unstable fashion - if they coincidently move somewhat closer to either of the extremes they fall out of this center region and end up at the extreme. The bipolarization is thus due to the low threshold for the latitude of rejection.

Figure 4.7 shows a simulation where the latitude of acceptance is relatively large, and that of rejection relatively small. For any agent in the simulation only very few other agents are in this rejection zone and for most of the agents, those with an initial position between 20 and 80, there are no such agents initially. Thus, for all agents the tendency to converge is much stronger than the tendency to diverge. A clear central cluster emerges, in line with the findings of Jager and Amblard.

In Figure 4.8 the three clusters of Jager and Amblard are clearly visible. The latitudes of acceptance and rejection are similar in size and each approximately a third of the range of possible attitude values. The result is that for attitudes $40 < \alpha < 60$ there are no other agents in the latitude of rejection, thus the only force pulling those agents is a converging one. Hence

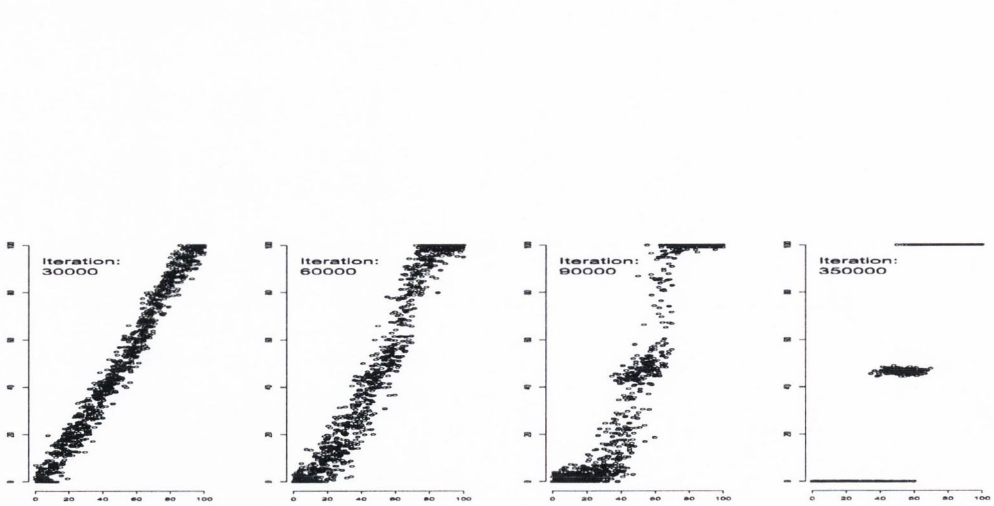

Figure 4.8: Replication of Jager and Amblard (2004). $u = 30$, $t = 60$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

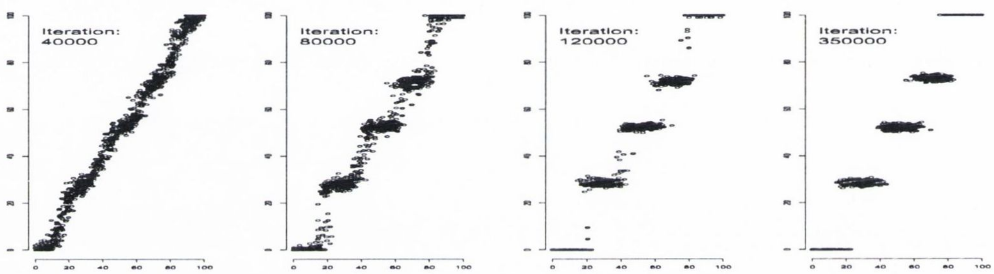

Figure 4.9: Replication of Jager and Amblard (2004). $u = 10$, $t = 80$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

the clustering around the mean of the attitudes of the agents in this range, at 50. The other agents are attracted to the agents at the same side of the scale, but rejected by those on the other side, and converge to one of the extreme positions, resulting in three distinct groups of agents.

Figure 4.9 describes a situation with two similarly sized, but smaller latitudes of acceptance and rejection. In this scenario, the agents with attitudes $20 < \alpha < 80$ experience no repulsing forces and converge to groups exactly two times the latitude of acceptance away from each other. Within these bands, the simulation resembles that of Figure 4.7, while beyond the band, the agents have no effect on other agents. The remaining agents again converge to each other, but are also forced away by those at the other end of the scale, resulting in extreme positions.

The main thing that stands out from these results is their straightforward interpretation due to the fact that the initial attitudes are uniformly distributed and that the values for u and t are constant across agents in each simulation. The combination means that one can describe exactly what pressures affect the agents, which makes the results straightforward to interpret. The idea that concerning a particular subject the attitudes are uniformly distributed over the entire range of possible attitudes before the diffusion process starts is not necessarily a very intuitive assumption. That the threshold values of the social judgment theory would be constant across agents certainly is an unrealistic assumption and appears to be at odds with the theory itself. Since these assumptions appear to drive the simulation results to a significant extent, it will be necessary to validate the simulation outcomes with those assumptions relaxed. The first assumption because in our democratic diffusion model it appears unrealistic that at the start of the nineteenth century there would be as many democrats as there would be autocrats and the second assumption because it is particularly unrealistic and appears to have a strong effect on the simulation outcome.

Normally distributed initial attitudes do not lead to a drastically different result. That is to say, the resulting distribution of attitudes is rather different from those presented above, but this change is a very direct effect of the density of the original distribution. If most agents had initial positions towards the lower end of the scale, then in a simulation which results in a larger number of groups (Figure 4.13), the groups towards the lower end of

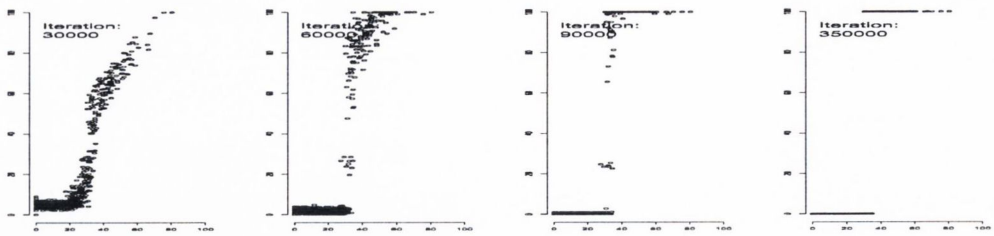

Figure 4.10: Normally distributed initial attitudes in Jager and Amblard (2004) with $\alpha_0 \sim N_{[0,100]}(10, 20)$. $u = 20$, $t = 30$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

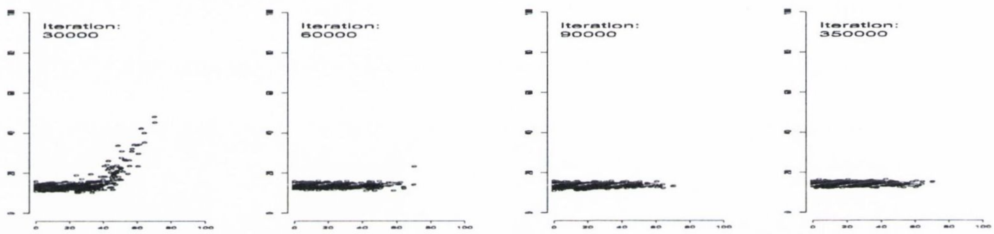

Figure 4.11: Normally distributed initial attitudes in Jager and Amblard (2004) with $\alpha_0 \sim N_{[0,100]}(10, 20)$. $u = 60$, $t = 80$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

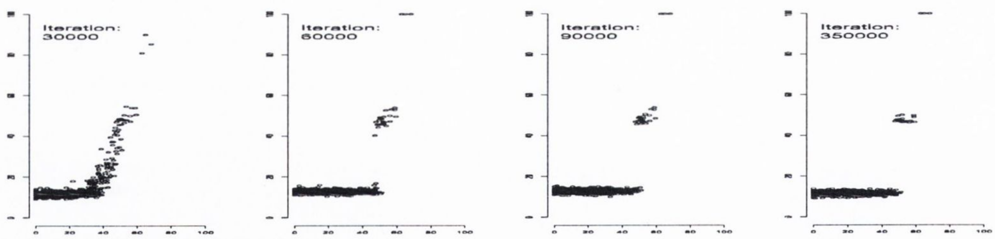

Figure 4.12: Normally distributed initial attitudes in Jager and Amblard (2004) with $\alpha_0 \sim N_{[0,100]}(10, 20)$. $u = 30$, $t = 60$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

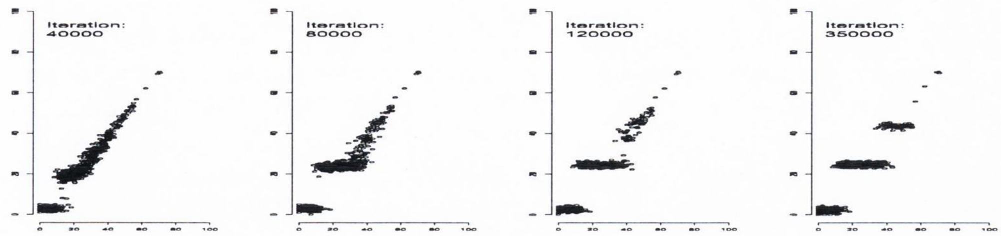

Figure 4.13: Normally distributed initial attitudes in Jager and Amblard (2004) with $\alpha_0 \sim N_{[0,100]}(10, 20)$. $u = 10$, $t = 80$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

the scale are much more densely populated than those at the other end of the scale. In other simulations, where the result is one homogeneous opinion (Figure 4.11), it affects the location of this position, which will be at the mean value of the original distribution.

Of course, this effect of the location of the original density and its spread can lead to a situation where particular opinions largely disappear from the outcome. In Figure 4.12, where we replicate the simulation of Jager and Amblard which leads to a "tripolarization", we can see that one of the groups is practically non-existent because it is too far away from the original mean density, while the middle group is very small. Had the variation in original opinions been smaller, even fewer groups would appear, from which we can deduce that would the original attitudes be constant across agents at the start of the simulation, then whatever the values for u and t, the end result would be similarly homogeneous. You need some agents to have different opinions from others for a diffusion to take place, unless another mechanism that affects these opinions operates, e.g., democracy promotion as in our main model.

Normally distributed u and t, although far more plausible than fixed u

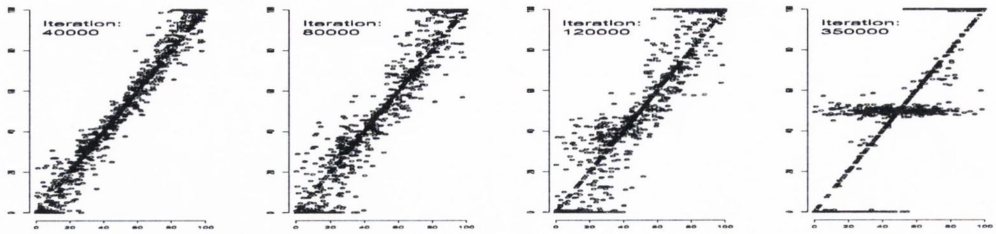

Figure 4.14: Normally distributed thresholds in Jager and Amblard (2004) with $u \sim N_{[0,100]}(10, 20)$ and $t \sim N_{[u,100]}(80, 20)$. The initial attitudes are on the x-axis and the attitudes at the given iteration on the y-axis.

and t as in the simulations of Jager and Amblard, appears to have only limited effect on the simulation results. For this reason, only the results for the simulation with very small latitudes of acceptance and rejection are presented here, as they deviate most. As one can see in Figure 4.14, in this simulation the distributions of latitudes creates a situation where many agents are as much attracted as repulsed by other agents, which leads to a situation where many of them hardly leave their original position. This results in the 45 degree linear pattern visible in the graph that plots current against initial values. Thus no sign of the five groups as found by Jager and Amblard and instead three main clusters of attitudes and a wide presence of attitudes outside these focal points. In a sense, this is the strongest evidence thus far of the possibility of a pluriformity of attitudes in a model of attitude diffusion.

One further generalization is important to consider. In the current simple model, all agents can communicate with any other agent. In terms of social network terminology, this would be a fully connected network, where all vertices are connected to all other vertices, and vertices are randomly activated to initiate communication. The obvious generalization would thus be to study alternative network specifications, for example social networks that are random in the sense that from the fully connected network, ties are removed in a random fashion. The expectation would be that this has no substantive effects on the results of the simulation and is an unnecessary complication. One should verify this assumption, however, before continuing to the more complex model presented in this chapter. Jager and Amblard (2004: 299-300) fortunately already perform this analysis and conclude that it does not affect the substantive results of their analysis. After replicating their results and furthering our intuition with regards to this model, the results with more realistic social networks will not be explicitly replicated, and we accept the findings of Jager and Amblard.

In the main model of this chapter, both the initial attitudes and the latitudes of acceptance and rejection of the different agents, or citizens, are normally distributed. The above simple agent-based model demonstrates that these modifications to the model of Jager and Amblard (2004) should only to a limited extent alter their findings. One would expect, depending on the settings of parameters U, σ_U, T, and σ_T, to generate either uniformity, bipolarization, or pluriformity in attitudes. To generate the highest level of pluriformity, normally distributed latitudes with a low value for U and a high one for T would be the suggested parameter settings. The result of moving the mean of the initial distribution of attitudes, A, on the outcome of these simulations does not have a major impact on the number of clusters formed, but should be kept in mind when the concept of the cascading revolution is taken into account. A very skewed distribution of attitudes towards democracy can either dramatically increase the probability of protest or turn it into an almost entirely impossible event.

4.4 Taking to the streets

The sixth concept of the model is what has been labeled protesting. Similar to broadcasting, protesting should here be seen as an abstraction of a broad spectrum of forms of political action. It includes all those publicly visible manifestations of individual attitudes towards democracy, or rather, the current political regime. Protests might literally mean protesting on the street, like in the Ukraine or the demonstrations in the DDR before the fall of the Berlin Wall, but it might also include dissenter writings or other forms of protesting art, mobilization for political action, like Solidarity in Poland, speech-acts or jokes (Johnston 2001), or votes for an opposition party in limitedly competitive elections. The protests have to be public, however, to qualify for this protest category, as the mechanism of the spiral of silence or preference falsification requires the visibility of these protests. Observing fellow individuals having the courage to take to the streets, literally or figuratively speaking, might lower the threshold for opponents of the regime to join the protests. The combination of the attitude towards democracy and the protesting status of an individual implements the spiral of silence into the agent-based model.

It should be pointed out that an alternative modeling approach to such publicly visible protest actions would be possible in terms of the model of cascading revolutions. The assumption underlying this modeling decision in this thesis is that it is the dynamics between public and private preferences (Kuran 1990) that generate the cascading effect of revolutions. An alternative specification, however, would pose that it is the dynamics between the perceived and the private preferences. Instead of using the publicly professed preferences as a direct indicator of the perceived support, one could assume a stochastic, uncertain estimation of this support, determined by the true support and some amount of uncertainty or error. In such a model, each individual citizen makes its own judgment of the amount of support for the opposition movement and subsequently decides whether or not it is advisable to join. Koster et al. provide a computational model based on exactly that assumption. They find that the more uncertainty around the size of support, the more likely opposition movements will be able to grow substantially (Koster et al. 2008: 70). For brevity's sake, this further extension will be ignored in the remainder of this thesis.

After the order in which citizens are being processed has been randomized, each citizen determines whether or not to start or stop protesting. In line with the cascading model of revolution as described above, a citizen will join the protest if the attitude against the current regime is strong enough relative to the proportion of protesters in the citizen's province to dare to risk the costs of protesting.

$$\psi_i = \begin{cases} 1 & \text{if} \quad \Omega_{C_i} = 0 \quad \& \quad \Upsilon \geq \frac{\alpha_i}{\lambda} \quad \text{or} \\ & \text{if} \quad \Omega_{C_i} = 1 \quad \& \quad \Upsilon \geq (1 - \frac{\alpha_i}{\lambda}) \\ 0 & \text{otherwise,} \end{cases} \tag{4.9}$$

where

$$\Upsilon = \frac{\sum_{j \in P_i} \psi_j}{N_{citizens,P_i}}. \tag{4.10}$$

Here ψ_i reflects the protesting status of individual i: $\psi_i = 1$ if i is protesting and 0 otherwise. Υ is the proportion of citizens in the province of citizen i that is protesting. α_i is the attitude towards democracy of individual i, in this equation normalized by dividing by the maximum value of this scale, λ. If the country is an autocracy, $\Omega_{C_i} = 0$, the proportion of protesters, Υ needs to be at least as large as the strength of the attitude towards democracy,

α_i/λ. If the country is a democracy, $\Omega_{C_i} = 1$, the proportion needs to be at least as large as the strength of the attitude against democracy, $(1 - \alpha_i/\lambda)$.

The manner by which individuals decide whether or not to protest is a good example of strategic adaptive behavior that agent-based models are particularly good at highlighting (Laver and Sergenti 2007: 1). Just as opinion pollers have difficulty gauching the extent to which anti-regime attitudes are present in a population (Noelle-Neumann 1993; Kuran 1995), so are individuals inside that society unable to know who would support the government and who would not. It is the proxy of this information communicated through public protest that provides this information and it is this proxy that the individual responds to. Instead of determining whether a majority in a society supports the regime or not, and follow this majority in a pure bandwagon fashion, the individual decides whether or not to protest only on the basis of his or her knowledge about the existing protest, as well as the personal attitude towards the regime.

4.5 Democracy promotion

Regimes will not helplessly watch how citizens change their attitudes towards them. Instead, they are likely to attempt to influence those attitudes, both locally and internationally. Especially democratic countries tend to make a serious effort trying to stimulate democratization abroad. Sometimes by using pressure towards foreign political leaders, but often also by stimulating grass-roots organizations in non-democratic countries or by providing alternative news sources to those provided by autocratic governments. A good example would be *Radio Free Europe*, which presented regional news from the Western perspective across Eastern Europe. In the model this element has been labeled broadcasting, for lack of a better term, which encompasses all forms of attempts by democratic governments to stimulate positive attitudes towards democracy in neighboring countries. Radio broadcasting is a good example, but this also includes supporting local organizations, distributing newspapers or pamphlets, or any other form of 'educating' individuals abroad by democratic governments. The presence of Serbian advisors in the Ukrainian Orange Revolution is another good example.

This mechanism of broadcasting relates closely to the increased references to 'grassroots' support when establishing democracy promoting activities abroad. The idea is that democratic governments can affect the attitudes and the likelihood of public anti-regime behavior among common citizens in foreign countries, independent of their government structures. The broadcasting effect here is very local in nature, however, whereby democratic countries attempt to influence opinions in neighboring countries more so than they do in countries farther away.

In terms of the implementation of this mechanism in the model, one randomly selected democratic capital will broadcast its democratic values to citizens in neighboring provinces. In this case not the Von Neumann neighborhood is taken into account, but all nine provinces that are either in the Von Neumann neighborhood or diagonally adjacent, including the capital itself. For each of the nine provinces, the probability of receiving the broadcast is one when the province is part of the same country, or one minus the maximum level of isolation of the two countries involved. For a province that receives the broadcast, all citizens update their attitude towards democracy by the size of B.

4.6 Revolutions and coups

Finally, regimes can change, democracies can turn authoritarian or vice versa. Such transitions can be largely due to actions by the political elite (Burton and Higley 1987) and/or due to a public demonstration of a serious lack of support among the general population. Elite transitions independent of actual or perceived legitimacy among the population are not part of this particular model and are considered exogenous to it. More in general, coups in this model encompass all those regime changes that are not explained by the level of protest in a country. The chances for such a regime change that is not explained by public protest - protest in the more general sense described above - is assumed to be higher directly after a regime change took place. This models the concept of regime consolidation (Linz and Stepan 1996). The second form of regime change is right at the core of the model. Regimes make a transition when all individuals in the capital protest. Due to the cascading nature of this protest mechanism, many countries where a substantial number of inhabitants are protesting have a good chance of falling into a state where all citizens are protesting the regime. It is thus the clustering of these revolutions that the model is trying to illuminate.

Each country determines whether or not a revolution will take place:

$$\Omega_{t+1} = \begin{cases} 1 - \Omega_t & \text{if} \quad \Upsilon_{capital} = 1 \quad \& \quad D \geq s \quad \text{or} \\ & \text{with probability} \max(K, \beta e^{\gamma s}) \\ \Omega_t & \text{otherwise,} \end{cases} \tag{4.11}$$

where s is the time since the last revolution or coup. D is a configurable parameter reflecting the number of days after a revolution or coup during which no revolutions should take place. β and γ are parameters determining the exponential function of the probability of a random coup, over time (s). This probability is capped at the configurable level K.

There are two ways in which a country can make a transition from or towards democracy: a country makes a transition when all citizens in the capital province are protesting[13] or randomly, with a probability which decays with the age of the regime. The latter can be seen as external shocks to the model, the many forms of revolutions in the world that are not caused by the diffusion of attitudes or even by public opinion. When Gorbachev let the Soviet Union slip and the regime collapsed, this can hardly be seen as an effect of democratic diffusion, but the subsequent collapses of many regimes in Eastern Europe to some extent can. The external shock of Gorbachev thus generated a diffusion effect as it is modeled here. Another way to look at this would be to see K as the built-in error in the model, much like any econometric model will include an error term. The fact that for endogenous revolutions only the capital is taken into account can be defended by a quick glance at most coups and revolutions in the world. Protests are generally more threatening when they take place in the capital, and rarely can a country where all but one region are opposed to the regime sustain its political system. Taking into account all provinces leads to unrealistic assumptions, while taking the capital into account seems in line with general perceptions of revolutions. Finally, for a number of iterations after a revolution, set by D, revolutions are not possible. Regimes are considered fragile in the first iterations after a coup or revolution, and coups or exogenous regime changes are more likely directly after a regime change, decaying over time. This decay

[13] A requirement of 100% protesters might seem too strong, but given the mechanisms of the cascading model of revolution this is theoretically the most appropriate and also in practice does not hold back many countries where there is a large proportion of protesters smaller than 100%. An alternative would be to use a 50% threshold instead (Koster et al. 2008; Kuran 1989: 54).

is parameterized with β and γ.

4.7 Conclusion

Combining all the various elements to the model together leads to one large computer simulation analysis. Appendix G provides all details concerning the underlying computer code used for these simulations. Each simulation run, first the system is set up by generating provinces as a cellular automata, country borders using the 'conquer' algorithm, and citizens based on the initial parameters. Then each simulation run goes through thousands of iterations. Each iteration the following steps will take place, in this order: the level of isolation for each non-democracy will change following a random walk; a random set of individual citizens will communicate with randomly selected fellow citizens; each citizen will determine whether or not to join the anti-regime protest; one randomly selected democratic capital will broadcast a pro-democratic message; and for each country it will be determined whether there are sufficient protesters for a revolution, or whether a random, exogenous coup will take place. Figure 4.15 highlights those steps graphically. Of course, the simulations produce large amount of data providing information about the dynamics of attitudes, protesting behavior, and political regimes.[14] This data can be analyzed using statistical analysis, either descriptive or inferential. The next chapter will discuss in more detail the various parameter settings under which the simulations were run and, based on these statistical analyses, discuss the conclusions we can draw from these simulations.

[14]In fact, the sheer size of data output - more than 60 Gigabytes - created problems of data analysis rare in political science.

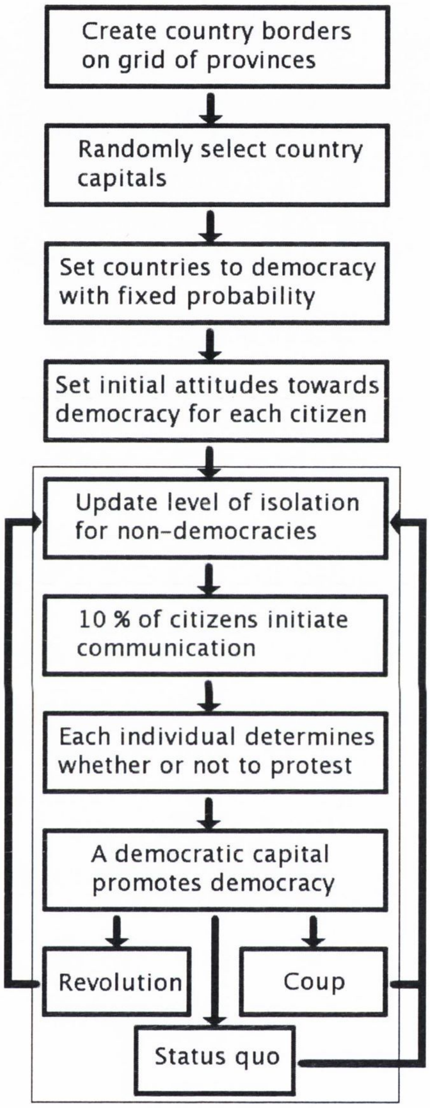

Figure 4.15: Schematic overview of the agent-based model.

CHAPTER 5

Simulation results

In the previous two chapters we gradually developed an attitudinal diffusion model of the global clustering of democracy. First the social judgment theory was presented, which describes how individuals alter their opinions through communication, how this process of cohesion and adhesion between individual norms can lead to a diversification of norms in a society. Not only did we provide a theoretical presentation of this theory in Chapter 3, but also a simulation-based demonstration of the mechanisms as implemented by Jager and Amblard (2004), and analogous to the implementation in the main model, in §4.3.1. Furthermore, we discussed the cascading model of revolutions, which demonstrates how a small change in some configurations of norms can lead to massive changes in public, or visible opinion. What was hidden before suddenly turns out to be hugely powerful and able to throw over governments. Once we discussed these foundations of the model, we presented the many aspects of the simulation model in detail in the previous chapter. Finally, Figure 4.15 provides a schematic overview of the agent-based model.

Once the many details as described have been implemented in a computer program,[1] we can run a large battery of simulations. The subsequent question is of course what the many simulation runs tell us. What does the model show about the diffusion of democracy and what are the implications for our theoretical understanding of the diffusion of democracy, if we assume this model to be relevant? We will focus our discussion on two main outcomes of the simulation runs. The first are the overall equilibrium outcomes of the model, in terms of the eventual overall level of democracy in the world. Under what conditions does a diffusion model like the one presented here lead to worldwide democracy or autocracy, under what conditions does democracy have a good chance of taking off and under what conditions does it not, and what does this suggest about the relevance of this model in understanding empirical trends? As mentioned previously, Modelski and Perry (1991, 2002) extrapolate on the basis of their empirical analysis that by 2113 the world well be fully democratic - under what conditions would we draw similar conclusions on the basis of this theoretical model?

The second element we will focus on is the clustering of political regimes, both spatially and temporally, as we discussed in Chapter 1. Under what conditions do we see substantial clustering and when do we not? And what does this say about the match between the model and the empirical observations we discussed in Chapter 1? Under what conditions do we see the spatial groups of democracies and autocracies and under what conditions the temporal waves of democratization seen in empirical data? As we discussed, three different types of clustering can be distinguished: spatial clustering whereby countries nearby other democracies are more likely to be democratic themselves; temporal clustering, whereby countries are more likely to

[1]See for the code Appendix G.

make a transition to democracy in times when many such transitions are taking place in the world; and spatio-temporal, whereby countries are more likely to make a transition to democracy when close-by countries make similar transitions. In the analysis of the simulation output, we will attempt to identify each of these three patterns.

5.1 Equilibrium outcomes

From their analysis of the diffusion of democracy as a technological innovation, Modelski and Perry (1991, 2002) conclude that the beginnings of an S-curve of democratization is visible, suggesting that eventually all countries will become democratic. In their analysis, it is a matter of early and late adopters, but eventually conversion to democracy will be inevitable. Although predicting the future is a hazardous activity for social scientists, it is interesting to see under what conditions, given the model of democratic diffusion presented here, the diffusion of democracy will lead to a fully democratic, or a fully autocratic world. To provide some insights to this question, a large number of simulations have been run using the parameter settings as listed in Table 5.1.[2] As one can see, most parameters have been set at a fixed setting for all simulation runs, with the exception of the theoretically most interesting parameters, which are related to the mechanism of the diffusion of democracy itself. These include the effect size of the broadcasting mechanism (B), which captures the activity of democracy promotion by democratic governments; the effect size of regular communication between individual citizens (δ), which captures the diffusion of norms among individuals; and the chance of such individual communication taking place across

[2]These settings for the map size and the multiplier result in a number of countries ranging from approximately 75 to 100.

international borders (τ). Note that when any of these parameters is set to zero, an entire aspect of the model is disabled. When $B = 0$, democracies do not promote democracy abroad; when $\delta = 0$, citizens do not communicate; and when $\tau = 0$, citizens only communicate within subnational regions. The one additional parameter that we vary is the random chance of coups (K), which can be interpreted as the 'error term' of the model, or alternatively, as a mechanism of exogenous shocks to the system.

To study the convergence to equilibrium of either democracy or non-democracy, we will look at the average proportion of citizens living in countries that are democratic in the last ten percent of iterations in a simulation. Table 5.2 provides the results of our parameter sweep. The resulting pattern is fairly clear and shows a clear interaction effect between democracy promotion and the diffusion of norms through individual communication. Strikingly, the extent to which the world, in the long run, democratizes is negatively related to the amount of communication among individuals. When individuals share norms, and the effect of democratic promotion is no stronger than the effect of this individual communication ($B \leqslant \delta$), the world eventually turns almost entirely autocratic.[3] On the other hand, when there is no individual communication at all ($\delta = 0$), and the mechanism of democracy promotion is in play ($B > 0$), the world eventually turns entirely democratic. The forces that could counter this trend are entirely disabled and there is no effect of the presence of autocratic norms in the population on the spread of democracy. Another striking result is that the presence or absence of cross-border communication between individuals has absolutely no bearing on the diffusion of democracy. Whether $\tau = 0$ or $\tau = \frac{1}{2}$, the results and standard errors are identical. Communication within national borders thus has an

[3]A result showing entirely democratic or autocratic results are unlikely, since the parameter for random coups is positive ($K > 0$).

Parameter	Description	Setting
W	Field width	20
H	Field height	20
M	Border multiplier	6
ϕ	Isolation mean	50
σ_ϕ	Isolation standard deviation	10
π	Initial proportion of democratic countries	$\frac{1}{10}$
C	Number of citizens, mean	50
σ_C	Number of citizens, standard deviation	10
λ	Number of levels on attitude scale	100
τ	Chance of cross-border communication	$\{0, \frac{1}{2}\}$
A	Initial attitude mean	10
σ_A	Initial attitude standard deviation	20
U	In-group threshold, mean	10
σ_U	In-group threshold, standard deviation	0
T	Out-group threshold, mean	80
σ_T	Out-group threshold, standard deviation	0
B	Size of the effect of broadcasts	$\{0, 1, 5\}$
δ	Size of the effect of communication	$\{0, 1\}$
K	Random chance of coups	$\{0, \frac{1}{10000}, \frac{1}{5000}\}$
D	Regime delay	50
β	Starting point of decaying regime fragility	$\frac{1}{20}$
γ	Strength of regime fragility decay	$-\frac{3}{20}$
	Number of iterations	100,000
	Thinning for global statistics	50
	Thinning for country statistics	50

Table 5.1: Model parameters and value settings. For the parameter sweep simulations, each combination of settings was used 100 times, resulting in 3600 simulation runs. For all simulations where $B = 5$ and $\delta = 1$, 500,000 instead of 100,000 iterations were used because of slow convergence.

important effect on the diffusion of democracy, negating weak democracy promotion, but cross-border communication has no bearing on this process. See Table 5.9 for a confirmation of these results in a Monte Carlo framework, where parameters are set randomly within predefined ranges, as opposed to fixed in advance.

Cross-border chance (τ)	Broadcast effect (B)		
Communication effect (δ)	0	1	5
$\tau = 0, \delta = 0$	.01	.98	.98
	(.00)	(.02)	(.02)
$\tau = 0, \delta = 1$	.01	.01	.73
	(.01)	(.01)	(.06)
$\tau = \frac{1}{2}, \delta = 0$	.01	.98	.99
	(.00)	(.02)	(.01)
$\tau = \frac{1}{2}, \delta = 1$	.01	.02	.74
	(.01)	(.02)	(.06)

Table 5.2: Average percentage of democratic states in the last 10% of iterations. Standard errors in parenthesis. Based on 3600 simulations.

It should be emphasised that what these different configurations represent are different "possible worlds" in terms of the international diffusion of democracy, given our model. The parameters do not distinguish particular countries from each other. The real world is likely to be more heterogeneous than the simulation world. Some countries have very effective democracy promotion campaigns ($B > 0$), while other countries' campaigns are entirely ineffective ($B = 0$). Some cultures will be very sensitive to foreign input ($\delta > 0$), while in other cultures there is a much more isolationist tendency ($\delta = 0$), even without government prevention of international communication. For the sake of brevity, and interpretability, the simulations only represent different possible worlds, in each of which the configuration of states is highly homogeneous. There are differences in regime type, in level of government generated isolationism for autocracies, in population and country sizes, but all democracy promotion campaigns are equally effective (B is either 0, 1, or 5 for all countries), international communication is equally likely to take place, conditional on the level of isolationism (τ is either $\frac{1}{2}$ or 0 for all individuals), and communication is equally effective for all individuals (δ is either 0 or 1 for all citizens). The differences presented in Table 5.2 thus show variation across different possible worlds, not across countries or individuals. Table 5.3 provides an overview of the various different possible worlds in the parameter sweeps, which can function as a guide in reading the various tables in this section.

	$B = 0$	$B = 1$	$B = 5$
$\delta = 0$	Attitudes towards democracy are fixed. No communication between citizens affects their attitudes, nor do democracies make any attempt to promote democratic attitudes.	Although the communication between citizens does not affect their attitudes, democracies actively promote democracy in the provinces close to the capital, both nationally and across borders. These campaigns are moderately effective.	Individuals do not transmit attitudes, but democratic governments provide highly effective campaigns to promote democracy.
$\delta = 1$	Individuals change their attitudes through communication either with fellow countrymen or also across borders ($\tau > 0$). Governments do not influence attitudes.	Individuals change their attitudes through communication, nationally and internationally ($\tau > 0$) and democratic governments actively promote democracy. The democracy promotion is relatively ineffective compared to the inter-individual diffusion of attitudes.	Individuals change their attitudes through communication and in addition, democracies provide highly effective campaigns to promote democracy in the region of the capital.

Table 5.3: Six types of "possible worlds" in the parameter sweep.

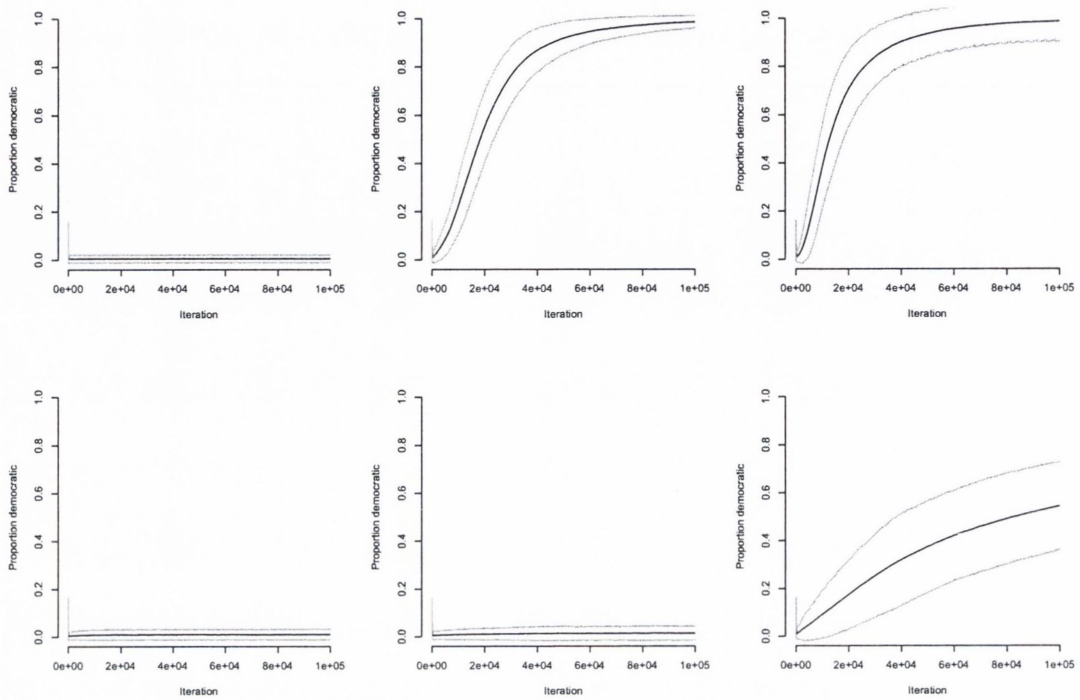

Figure 5.1: Patterns of long term equilibria. The data is the same as in Table 5.2, for $\tau = 0$ and $\tau = \frac{1}{2}$ combined, presented over time. The grey lines denote the standard errors. The top row concerns the cases where $\delta = 0$, the bottom $\delta = 1$. The first column $B = 0$, the second $B = 1$, and the last $B = 5$.

Figure 5.1 shows the same simulation runs. The small standard errors[4] towards the end of each simulation are a strong indication that the simulations converged to a particular value,[5] with the exception of the last graph, when $\delta = 1$ and $B = 5$. Here the variance remains large in the long run, even given

[4]Note that these are 'empirical standard errors', i.e. they denote the standard deviation of the various simulation runs.

[5]An alternative, statistical approach to this would be to look at whether the mean of the last ten percent of the iterations differs substantially from the mean halfway the simulation. In this case, all but the last configuration of parameters lead to convergence in the proportion of citizens living in democracies during the first 100,000 iterations, within a margin of error of 5% of the number of countries.

that for these settings the number of iterations used are five times as large. It should be noted that these figures show plots smoothened by the fact that averages across simulation runs have been used. The individual simulation runs are substantially more erratic than these averages suggest, in particular in those areas where the standard errors plotted here are large. In both scenarios where the world evolves to an almost entirely democratic state, one can clearly see the S-curve as projected by Modelski and Perry (1991, 2002). This is not overly surprising, since a model that includes democratic promotion, does not include autocratic promotion, and does not include other individual effects of communication ($\delta = 0$), is by far closest to the conceptualization of democratic diffusion as a process of diffusion of innovation. The mechanism of the cascading revolution plays a role in the extent to which we can predict for an individual country the timing of the revolution, but it does not alter our understanding of the global pattern of the diffusion of democracy.

5.2 Clustering

In Chapter 1 we made a clear distinction between three different types of clustering of political regimes: spatial, temporal, and spatio-temporal (see Table 1.1). The first refers to the fact that when one takes any random cross-section of countries in the world at any point in time, one can see that countries that are surrounded by democracies have a higher chance of being democratic themselves. This is a purely static observation and could be related both to the higher likelihood of particular regime transitions in certain international environments, or it could be related to common external factors affecting neighboring countries, or to higher chances of survival of democratic regimes when surrounded by other democracies. Various alternative explanations were discussed in Chapter 2. The key observation here is the clear, statistically significant geographic clustering of political regimes, without any reference to the timing of transition or the causes or underlying mechanism of this spatial clustering.

The second refers to the often mentioned waves of democracy, whereby one can see that transitions to democracy and away from democracy tend to occur in international waves. Particular periods in time show many transitions in one direction, while other periods show many transitions in the opposite direction. The transitions are thus clustered in time, rather than randomly spread over the years. Here we ignore the geographical pattern of this clustering. A change in international norms with regards to democracy can occur due to regime changes in one part of the world and subsequently affect developments in entirely different parts of the world. Although one could imagine some spurious relationships, for example when a slowdown in the world economy affects the stability of various regimes in very different areas of the world, most explanations of such temporal clustering would have to be related fairly directly to a form of diffusion. Regime changes elsewhere change the international climate of democratic norms, or their success provide examples for activists or political elites to change their regimes, etc. Again, various explanations are discussed in Chapter 2. The key here is the clustering in time regardless of geographical location. Any cross-section of the data at a fixed point in time will not reveal any such waves.

Finally, one can see countries in larger geographic regions make transitions to democracy virtually simultaneously, thus creating both a temporal wave of democratization and a spatial cluster of new democracies. The two mechanisms thus coexist when temporal waves are spatially concentrated. This is what we labeled spatio-temporal clustering. The two types of clustering are combined by the fact that not just political regimes are geographically clustered, but also transitions from one regime type to another. Clear examples are the transitions visible over the years in Latin America (see Markoff 1996) or the collapse of the Soviet Union, when many regimes in Eastern Europe made a transition to democracy, in a short period of time (Kuran 1995).

In Chapter 2 we discussed extensively how various different theoretical contributions can help explain each of these three forms of clustering. In this section, we will study under what conditions the model of diffusion presented here could. In other words, are there particular configurations of parameter settings under which we indeed observe these three patterns of regime clustering? We will use the same simulations as in the previous paragraph, with the settings presented in Table 5.1.

5.2.1 Spatial clustering

The point of departure for this thesis is the repeatedly observed spatial autocorrelation in measures of democracy. As many empirical studies suggest, an important factor that cannot be ignored in explaining the level or presence of democracy is the extent to which democracy is present in neighboring countries. The more democratic the region or the bordering nations, the more likely a country is to make a transition to democracy or to stay democratic (Starr 1991; Ward et al. 1996; O'Loughlin et al. 1998; Ward and Gleditsch 1998; Gleditsch and Ward 1997, 2000, 2006; Brinks and Coppedge 2006; Gleditsch 2002; Elkink 2003; Doorenspleet 2001, 2004; Wejnert 2005; Fordham and Asal 2007). This observation is discussed more extensively in Chapter 1. The model developed in this thesis is an attempt to provide a possible theoretical explanation of this phenomenon - thus far, the amount of theory to support the empirical claims has been fairly limited. Although the fact that proximity matters in terms of international influence sounds certainly plausible and appealing, it does not in any sense highlight the causal relationships between these two phenomena - neighboring democratic regimes and local transitions or local survival of democracy.

Although other explanations can be imagined - see, for example, Cederman and Gleditsch (2004) for an entirely different approach - in this thesis we try to explain the influence of neighboring countries through both the communication between individual citizens of countries, both domestically and internationally, and the practice of democracy promotion abroad by democratic governments. Using the same simulations as those presented in the previous paragraph, we can attempt to interpret the extent to which this model does, as a thought experiment, indeed lead to a spatial clustering of democracies as observed in empirical research.

As discussed in Chapter 1, the standard measure for the level of spatial clustering is Moran's I (see equation 1.2), which has an expected value of $-1/(n-1)$ (Griffith and Arlinghaus 1995: 27). To measure the level of spatial clustering in our simulations, it thus makes sense to use this statistic. The level of spatial clustering of course varies over time, during the course of a simulation run. For that reason, we will look at the average deviation of Moran's I from its expected value, over the entire duration of the simulation run. The expected value depends only on n, the number of countries, which is constant for the duration of a simulation run. Taking the deviation from this number instead of the 'raw' value of Moran's I makes the results easier to interpret. If the deviation is significantly larger than 0, significant spatial clustering is present. Table 5.4 presents the results of this analysis.[6]

[6] A similar approach to convergence can be taken as discussed in Footnote 5. In 100,000 iterations, the average level of clustering converged for all cases where $B = 0$. The lowest

Cross-border chance (τ) Communication effect (δ)	Broadcast effect (B) 0	1	5
$\tau = 0, \delta = 0$	.008	.060	.074
	(.001)	(.031)	(.034)
$\tau = 0, \delta = 1$	.006	.018	.187
	(.008)	(.028)	(.057)
$\tau = \frac{1}{2}, \delta = 0$	.008	.059	.074
	(.001)	(.034)	(.034)
$\tau = \frac{1}{2}, \delta = 1$	.006	.026	.184
	(.007)	(.038)	(.054)

Table 5.4: Average deviation from expected Moran's I. Standard errors in parenthesis. Based on 3600 simulations.

If one investigates whether the results in this table are statistically significant, using the typical threshold value of $\alpha = .05$,[7] one finds statistically significant results for all cases where $\delta = 0$, as well as for the $\delta = 1$ and $B = 5$. In the case of $\delta = 0$ and $B = 0$, the level of spatial clustering is very low and substantively insignificant. Why this level of clustering is nevertheless statistically significant is somewhat unclear. The same cases where the simulation leads to, eventually, a completely democratic world are thus the cases where we observe significant levels of spatial clustering. The most stark level of spatial clustering is visible for the case where $\delta = 1$ and $B = 5$, thus where both are positive and the effect of broadcasting is stronger than the effect of inter-personal communication. The former generates the international clustering patterns, while the latter is too weak to bring the average attitude in a country back to where it was before the broadcasting effect, while at the same time being strong enough to reinforce local attitudes. Again, it is striking that the effect of τ is negligible. Whether or not norms diffuse between

score of convergence is 80%, found when $\delta = 0$ and $B = 1$.

[7]Note that for the calculation of the standard errors in this table, the observed standard deviation of the simulation results is presented, which is unrelated to the type of calculation of the standard errors as in equation 1.5.

individuals across (sub)national borders has, given this model, no effect on the extent to which democracies cluster geographically. The key mechanisms are the broadcasting by democratic regimes to promote democracy abroad and the inter-personal communication between citizens of the same country to stabilize or reinforce attitudes within the country.

5.2.2 Temporal clustering

With temporal clustering we refer to the world-wide waves of democratization as highlighted by Huntington (1991). Certain periods in the history of the last two hundred years showed many democratizing countries, while in other periods either a lot of reversals to non-democratic regimes took place or simply fewer regimes made a transition. Some doubts can be cast on the main finding by Huntington (see, e.g., Doorenspleet 2000; Paxton 2002), but nonetheless we can see temporally denser and sparser periods of democratization. The interesting question in this model is thus under what circumstances we can find similar waves of democratization and autocratization in our model of diffusion.

The patterns over time presented in Figure 5.1 suggest an absence of any such waves of democratization, but one should not forget that these are based on averages across simulations. Wave-like patterns in individual simulations will, if not coinciding in their timing, be averaged out against each other. The smooth patterns in these figures do not imply the absence of a temporal clustering of democracy.

To measure the presence of waves, we will simply look at the first order autocorrelation in the world-wide number of democratizations. Each iteration of the simulation we count the number of transitions to democracy and the number of transitions away from democracy, adding both coups and popular revolutions together. We then look at the extent to which the number of transitions at time t correlates to that at time $t-1$. In one iteration in this model not an awful lot happens. To look at the autocorrelation between t and $t-1$ seems unreasonable, considering the small movements per iteration in each simulation run. For that reason, the iterations have been divided in blocks of 100 time periods and thus the autocorrelation between blocks of size 100 will be studied.

	Broadcast effect (B)		
Communication effect (δ)	0	1	5
0	0.14	0.12	0.13
	(0.03)	(0.03)	(0.03)
1	0.14	0.14	0.13
	(0.03)	(0.03)	(0.02)

Table 5.5: Average first order autocorrelation of number of transitions to democracy, based on time blocks of length 100. Standard errors in parenthesis. Based on 3600 simulations.

	Broadcast effect (B)		
Communication effect (δ)	0	1	5
0	0.13	0.13	0.13
	(0.03)	(0.03)	(0.03)
1	0.14	0.13	0.12
	(0.03)	(0.03)	(0.02)

Table 5.6: Average first order autocorrelation of number of transitions to autocracy, based on time blocks of length 100. Standard errors in parenthesis. Based on 3600 simulations.

Table 5.5 shows the results of this analysis for transitions to democracy; Table 5.6 for transitions away from democracy. As is clear from this table, there is very little autocorrelation visible, albeit statistically significantly different from zero, and thus very little evidence of waves of democracy. The model as presented in this paper provides little in terms of an explanation of the global waves of democratization and reversal of democracy as described by Huntington. Although the model provides some explanation of how the diffusion of individual attitudes towards democracy can affect the international geographical clustering of processes of democratization, the same individual diffusion of attitudes does not appear to explain the global waves, at least not in the model as analyzed here.

Although not presented in these tables, the parameter sweep also varies the random chance of experiencing a coup, K. No effect is visible when taking this variation into account. This is a rather striking result. The random chance of coups were implemented to add the equivalent of a statistical error term to the model. An additional effect was also envisaged, however. Although the coups could not themselves explain the spatial or temporal clustering patterns in the data, since they are modeled as effects purely independent of events in either neighbouring countries or in the past of the same country, they would still be expected to have an effect on the form temporal clustering takes. One would expect that a coup that is random for as far as the model is concerned could still instigate a regional or global wave of democratization. In other words, they could take the effect of external shocks to the system, which initiate the temporal clustering without having an effect on the process of diffusion itself. No such pattern is visible in the simulation results, however. The absence ($K = 0$) or presence ($K = \frac{1}{10000}$ or $K = \frac{1}{5000}$) of this type of exogenous shock does not affect the absence of presence of temporal waves of democracy.

5.2.3 Spatio-temporal clustering

Whereas the statistics presented in §5.2.1 provide information on the extent to which democracies cluster geographically in our virtual simulation world, this section will discuss the extent to which *democratization*, as opposed to *democracy*, clusters geographically. In other words, to what extent can we speak of local waves of democratization in the simulation output - independent of the fact that we do observe spatial clustering of democracies in a static sense and that we do not observe temporal clustering of democracies at a global level? To analyze these spatio-temporal patterns we will take an approach similar to the one in §1.2.3. By using a Markov chain model we can study the probability of transitions to democracy and the probability of survival for existing democracies on the basis of global and local levels of democracy, as well as local transitions to and away from democracy. Because of the large volume of simulation output - 3600 simulation runs with each over 100 countries and 100000 iterations - a regular ordinary least squares model was estimated instead of a logistic regression. Furthermore, no control variables or random effects were added, as they are less relevant in this idealised world of simulations. For example, no economic or historical factors would have to be taken into account.[8] The model estimated for each simulation run is thus:

$$
\begin{aligned}
y_{it} = &\beta_0 + (\beta_1 G y_{i,t-1} + \beta_2 \tilde{W} y_{i,t-1} + \beta_3 \tilde{W} \Delta y_{i,t-1}) y_{i,t-1} \\
&+ (\beta_4 G y_{i,t-1} + \beta_5 \tilde{W} y_{i,t-1} + \beta_6 \tilde{W} \Delta y_{i,t-1})(1 - y_{i,t-1}) + \varepsilon_{it},
\end{aligned}
$$

where y_{it} measures the presence of democracy for country i at time t; G is a matrix with zeros on the diagonal and $\frac{1}{N_t - 1} \sum_{j, j \neq i} C_{jt}$ in all other cells, where

[8]The code for this part of the analysis can be found in Appendix G.4.

N_t is the total number of countries at time t; $\Delta y_t = y_t - y_{t-1}$ is the first difference of the level of democracy; and the β's are regression coefficients. For democracies at time $t-1$, β_1 will indicate the effect of global waves of democratization, of temporal clustering; β_2 the effect of spatial clustering; and β_3 the effect of local waves of democratization, thus spatio-temporal clustering. For autocracies, β_4, β_5 and β_6 will capture the equivalent effects. Because of the size of the data set, ordinary least squares regression was executed despite the fact that the dependent variable is binary. The resulting beta coefficients are reported in Table 5.7. The standard errors reported in this table are based on the distribution across simulations, not on the estimated standard errors in the individual regressions.

		Gy_{t-1}		$\tilde{W}y_{t-1}$		$\tilde{W}\Delta y_{t-1}$	
B	δ	Aut.	Dem.	Aut.	Dem.	Aut.	Dem.
0	0	-0.46	5.30	-0.00	0.03	0.01	-0.18
		(0.32)	(6.83)	(0.01)	(0.04)	(0.01)	(0.24)
0	1	-0.47	5.75	-0.00	0.03	0.01	-0.19
		(0.32)	(7.60)	(0.01)	(0.04)	(0.01)	(0.27)
1	0	-0.37	3.04	0.03	0.01	-0.00	-0.09
		(0.17)	(3.08)	(0.05)	(0.04)	(0.02)	(0.08)
1	1	-0.47	5.83	-0.00	0.03	0.01	-0.19
		(0.32)	(7.83)	(0.01)	(0.04)	(0.01)	(0.26)
5	0	-0.43	2.77	0.04	0.02	-0.01	-0.09
		(0.24)	(3.10)	(0.05)	(0.02)	(0.02)	(0.09)
5	1	-0.48	5.17	0.01	0.02	0.01	-0.15
		(0.32)	(6.86)	(0.03)	(0.04)	(0.01)	(0.20)

Table 5.7: Regression coefficients estimating spatial, temporal, and spatio-temporal clustering effects in the simulations. Gy_{t-1} captures temporal clustering; $\tilde{W}y_{t-1}$ captures spatial clustering; and $\tilde{W}\Delta y_{t-1}$ captures spatio-temporal clustering. Standard errors in parenthesis. Based on 3600 simulations.

As it turns out, Table 5.7 shows almost no significant coefficients. Given the parameter settings for the broadcast effect (B) and the communication effect (δ), all regression coefficients show both negative and positive results across the various simulations. The only exception is the coefficient for the temporal clustering of transitions from autocracies to democracies, when $B = 1$ and $\delta = 0$. We will have to be reluctant to attribute this to anything other than chance. Correlations that we found earlier in this section are thus not confirmed by this regression, perhaps due to the small amount of changes in the simulations in any given time period. These regressions are based on the full 100000 iterations, rather than on some aggregation similar to the analysis of the temporal clustering.

5.3 Monte Carlo simulations

The common method for analyzing the output of agent-based models is a systematic sweep across particular parameter settings. Taking this approach, one selects a number of reasonable settings for each parameter and then runs a potentially large number of simulations for each particular setting to distinguish the stochastic from the systematic part of the model. The advantage is full control over the parameters of interest and thus a clearer picture of the effect of changing one parameter, while keeping all others constant. To sufficiently study interesting combinations of parameter settings, however, this approach can be computationally expensive and important combinations of parameters can easily be missed. The background of agent-based modeling is, after all, in the theory of complex systems, whereby the unexpected, non-linear interactions between individual agents in the model is the key concern of the paradigm. A parameter sweep necessarily leaves a virtually infinite[9] number of points in the parameter space untouched, possibly missing unex-

[9]Real numbers in a computer are only finite approximations of such, hence computer simulations have by definition a finite parameter space.

pected interactions at these particular points in the parameter space.

An alternative approach to selecting parameter settings is a Monte Carlo approach, whereby parameters are randomly set for each simulation run, within predefined ranges (see also Plümper and Martin 2006; Laver and Sergenti 2007). The advantage is that there is a much smaller chance of missing important subspaces of the parameter space, where interesting interactions take place; the disadvantage is that it is much harder to interpret the outcomes. Changes in the simulation output cannot be directly attributed to particular changes in the parameter settings, since all parameters change and any given parameter setting has too few simulation runs to distinguish the stochastic from the systematic component of the simulation. Thus, a Monte Carlo scheme of selecting parameter settings requires a more statistical approach to studying simulation output. In this section, we will provide such an analysis as a validation of the outcomes that have been described in the previous sections, which made use of systematic parameter sweeps. The parameter space over which random settings are selected is provided in Table 5.8.

Because the parameter space is very large, we will need statistical analyses to summarize the results. The immediate question then is what statistical analyses to use and what parts of the parameter space to focus on. Since this section is primarily an attempt to validate the results from the parameter sweep in the previous section, we will use the parameter sweep as a guide in seeking out particular regions of the parameter space. The other variation in parameters will be seen as nuisance parameters, where we are interested in the robustness of the parameter sweep results given this variation, rather than in the variation itself. As the statistical method we will use straightforward linear regressions in which we allow for all possible interactions between variables in the model. To make the results comparable to those provided in the previous section, we will use the regression results to predict the equivalent values reported above. In other words, if we assume the linear models to be correct, what would have been the result of a parameter sweep, controlling for the variation in the other variables? The predictions are based on samples from the posterior distribution, so that the uncertainty measured by the standard errors in the linear model are carried through into the calculation of the predictions. The resulting standard errors are also reported.

Table 5.9 reports the Monte Carlo equivalent of Table 5.2. This table has been produced by running a linear regression of the equilibrium value of the proportion of democratic states (weighted by population size) on the broadcast effect B, the probability of cross-border communication τ, the size of the communication effect δ, and the probability of coups K.[10] Based on this regression result 1000 regression coefficients were sampled from the posterior distribution and used to predict the proportion of overall citizens living in democracies given the reported values for B and δ, holding the other variables at their mean.

Because these are predicted values, and there is much more variation in all other parameters in the model, it is no surprise that the result is less stark than in Table 5.2. It also underlines the argument that parameter sweeps lead to more clearly interpretable results, while the many changes in parameters through random sampling generates less clear patterns that are at times hard to interpret. Nonetheless, it is encouraging that the results of the Monte Carlo analysis here show a pattern somewhat similar to that of the parameter sweep, with the possible exception of $B = \delta = 1$, where the proportion of citizens living in a democracy is clearly higher than the

[10]The regression results themselves can be found in Table E.1, in Appendix E. Table 5.9 is based on Model 3.

Parameter	Description	Range
W	Field width	$[20, 40]$
H	Field height	$[20, 40]$
M	Border multiplier	$[0, 20]$
ϕ	Isolation mean	$[0, 100]$
σ_ϕ	Isolation standard deviation	$[1, 30]$
π	Initial proportion of democratic countries	$[0, \frac{1}{5}]$
C	Number of citizens, mean	$[40, 100]$
σ_C	Number of citizens, standard deviation	$[1, 40]$
λ	Number of levels on attitude scale	100
τ	Chance of cross-border communication	$[0, 1]$
A	Initial attitude mean	$[0, 100]$
σ_A	Initial attitude standard deviation	$[1, 50]$
U	In-group threshold, mean	$[0, 80]$
σ_U	In-group threshold, standard deviation	$[1, 20]$
T	Out-group threshold, mean	$U + [0, 80]$
σ_T	Out-group threshold, standard deviation	$[1, 20]$
B	Size of the effect of broadcasts	$[0, 5]$
δ	Size of the effect of communication	$\{0, 1, 2\}$
K	Random chance of coups	$[0, \frac{1}{5000}]$
D	Regime delay	$[50, 100]$
β	Starting point of decaying regime fragility	$[0, \frac{1}{5}]$
γ	Strength of regime fragility decay	$[-\frac{1}{3}, 0]$
	Number of iterations	100,000
	Thinning for global statistics	50
	Thinning for country statistics	50

Table 5.8: Model parameters and value ranges. 2000 simulations were run with parameters drawn from uniform distributions within the ranges presented above and, for all parameters that so require, rounded to the nearest integer.

parameter sweep suggests. Although the parameter sweep suggested that the effect of broadcasting clearly had to be stronger than the effect of inter-personal communication ($B > \delta$), the Monte Carlo parameter settings do not support this conclusion. The inter-personal communication clearly slows down the international growth of democracy by somewhat neutralizing the

	Broadcast effect (B)		
Communication effect (δ)	0	1	5
0	.60	.69	.86
	(.02)	(.01)	(.02)
1	.30	.39	.55
	(.02)	(.01)	(.01)

Table 5.9: Predicted percentage of democratic states in the last 10% of iterations. Standard errors in parenthesis. Based on 2000 simulation runs and 1000 samples from the posterior of a linear regression.

broadcasting effect, but also in the case where the broadcasting effect is present but weak does the world gradually move towards a more democratic configuration. It should be kept in mind, however, that these results are based on linear predictions of the level of democracy, rather than observed levels. If we assume that the relation between B, δ and the equilibrium democracy value is not linear, it will still be picked up accurately by the parameter sweep, but incorrectly by this approach to interpreting the Monte Carlo results.

	Broadcast effect (B)		
Communication effect (δ)	0	1	5
0	.005	.018	.079
	(.005)	(.003)	(.004)
1	.004	.017	.078
	(.005)	(.003)	(.003)

Table 5.10: Average deviation from expected Moran's I. Standard errors in parenthesis. Based on 2000 simulation runs and 1000 samples from the posterior of a linear regression.

Table 5.10 reports the Monte Carlo equivalent of Table 5.4, using regressions identical to those used for Table 5.9.[11] Again, the results are somewhat

[11]The regression results themselves can be found in Table E.2, in Appendix E. Table 5.10 is based on Model 3.

akin to those from the parameter sweep, but not identical. One exception is again the situation where $B = \delta = 1$, which is in this case not distinguishable from the other cases where $B > 0$. Again, this could be due to a non-linear relationship between B, δ and the simulation outcomes. A more striking result, however, is the case of $B = 5$, $\delta = 1$. Whereas in the parameter sweep this scenario lead to a much higher level of spatial clustering, in this Monte Carlo analysis the situation is indistinguishable from, again, all cases where $B > 0$. The Monte Carlo analysis would suggest that if there is an active attempt by democratic countries to influence democratic attitudes in neighboring countries, there will be spatial clustering visible, and if there is not, there will not be such clustering visible. The level of inter-personal communication does not appear to have any effect on the spatial clustering of democracies.

5.4 Conclusion

The original aim of the model was to study under what conditions a model of the diffusion of attitudes towards democracy, related to a cascading model of revolution, can explain the different kinds of clustering patterns we observe over the past two centuries. It turns out that under some conditions we observe some of these clustering patterns, but not all three types are visible in the simulation runs. The most striking is the absence of any form of temporal waves. The four waves of democracy visible in long term democracy data is not visible in the simulations here, irrespective of the configuration of parameters.

The spatial clustering of democratic regimes, the observation of which was the original trigger for this research, is clearly visible under some parameter conditions. The most clear result in this regard is that international democracy promotion by democratic regimes is a more important factor than individual norm diffusion. The two factors clearly interact: only when the effect of democracy promotion is clearly stronger than the effect of individual norm diffusion does democracy cluster geographically. The individual norm diffusion has a somewhat unexpected effect on the outcome of the model, probably related to the process described earlier (§3.1.3), whereby an unlimited diffusion of norms drags opinion to the average level, reducing any chances for revolution. The social judgment model should mediate this mechanism, but might be doing so insufficiently. The effect is that the individual communication can reduce the effects of democracy promotion, by 'dragging' the opinion back to the average and thus reducing the level of geographic clustering.[12]

Irrespective of levels of clustering, the simulation results suggest that in any scenario where the level of democracy promotion is stronger than the effect of individual communication, the world will eventually turn democratic. As long as the communication between individual citizens cannot neutralize external influences on popular opinions, democratic states have a chance to stimulate democracy abroad and push the world a little closer to a worldwide Kantian democratic peace. Such predictions are always hazardous, of course, and should be taken with a grain of salt. The next chapter will discuss in much more detail what we can conclude from these simulation results and more importantly, what kind of questions these results raise for future research in the diffusion of democracy.

[12]See also the negative coefficient for the communication effect in Table E.2.

CHAPTER 6

Discussion

In the previous chapter the results of the various simulations were presented that give a picture of how the key factors under study in this dissertation relate to each other. We have seen how the active promotion of democracy by democratic regimes across country borders and more informal communication between citizens of countries relate to the temporal and spatial clustering of political regime types. We have observed the interaction of these mechanisms of democracy promotion with the presence of regime changes exogenous to this model. Although these simulation results suggest some tentative hypotheses concerning the relations and interactions between these variables, there is still plenty of room for further study and the results also beg an empirical validation of the results. Based on the imaginary world of the computer simulations, these results can indeed not be seen as much stronger than that: fruitful, but tentative hypotheses for further research.

Given these results, the question arises as to what the implications are for both the academic study of regime transitions and democratization in particular and for the more practical world of policy making with regards to regime transitions. Although most transitions are highly complex events with many actors and influences culminating in a large change of the way a society works, some more concrete policies both by democracies and by authoritarian regimes do appear to have a discernible effect on regime transitions. An example would be the active involvement in the transition to democracy of contemporary Iraq by the United States. In this case not only did a democratic regime have a clear and influential effect on the transition of a regime, it also went combined with a rhetoric suggesting that the presence of a democratic regime in the Middle East might lead to a spatial diffusing effect, improving the chances for democracy in the region. To what extent can the simulation results presented above shed light on the chances for such a diffusion effect emanating from Iraq?

In this chapter attention will turn to the more theoretical implications of the interpretation of democratic diffusion suggested by these simulations. How would existing theoretical approaches have to be adjusted to incorporate the findings presented here? To what extent do these findings suggest new pathways for future democratization research? How can one go about finding empirical validation of these results? And more in general, what kinds of conclusions can we draw from simulations and agent-based models like these? This chapter thus comments briefly on the role of agent-based modeling in the social sciences and its relation to the positivist approach to political science.

The next chapter will conclude by considering the more practical implications. What would this theoretical perspective suggest for practical policy by both democratic and non-democratic regimes? What does it suggest about activities like *Radio Free Europe* to promote democracy across borders? Or the effect of extensions of trade relations with non-democracies by democracies on the chances of regime changes in the partner country? What does it suggest about the diffusing effects of forcibly changed political regimes?

When considering the theoretical implications of the model presented above, it should be emphasized again that this research is not an attempt to replace any of the existing theories of democratization, international factors in democratization, or the diffusion of democracy. Instead, an additional perspective is provided that might explain parts of the otherwise hugely complex phenomenon of transitions to democracy. Looking solely at ideological factors at the level of the individual citizen, as this model does, could never be sufficient in explaining such tremendous processes of change as with the collapse of an autocratic regime and the establishment of new institutions and new rules of the game. Elite attitudes and behavior can not be ignored, nor can economic and military factors, or the developments in the social structure of a country. A straightforward theoretical implication that this model is right and therefore all other models are wrong is thus inappropriate and also not sought.

One could conclude, however, that the diffusion of attitudes through individual citizens could be part of the explanation of what causes the geographical clustering patterns in democratic regimes across the world that we observe. For such an explanation we do not necessarily have to revert to theories of elite emulation and learning, or of survival strategies in a militarily hostile world, but instead we can relate it to conventional models of mass-based democratization and revolution. The fact that international developments matter, which is evident given the international patterns that are visible in the end result, does not imply that the explanation of democracy in those countries will have to be an elite based theory.

In the following sections we will first turn to some very general comments on agent-based modeling, with particular focus on issues in the empirical validation of agent-based models. We will then turn more specifically to the various venues of future research suggested either directly or indirectly by the research in this thesis, starting with some comments on the empirical validation of the model presented here itself.

6.1 Some comments on agent-based modeling

We will start this discussion of the simulation results with some more general comments on the use and interpretation of agent-based models in the social sciences. The use of agent-based models in the social sciences is a relatively new development, which is quickly gathering pace. In political science they have an increasing presence at the major international conferences as well as in the major academic journals. They gradually move from the specialized niche journals like the *Journal of Artificial Societies and Social Simulation* to the more general and widely read ones such as the *American Political Science Review*. As mentioned earlier, various relatively well-known recent applications of this research strategy now exist in fields such as democratization (Cederman and Gleditsch 2004), international relations (Cederman 1997; Lustick, Miodownik and Eidelson 2004), cooperation (Axelrod 1997*a*), and party competition (Laver 2005). Because of its novelty and increasing presence, it is worthwhile to carefully contemplate what exactly the role of this type of modeling is or can be in the empirical positivist enterprise in political science, and what kind of conclusions can or cannot be drawn from such simulation methods.

6.1.1 Social science and complexity

The basic premise of the agent-based modeling strategy is founded on the concept of complexity. This concept, closely linked to chaos theory, relates to how simple patterns of behavior for many small elements in a system can lead to complicated and unpredictable patterns in the system as a whole. The classical example is the weather system developed by Lorenz. By using a very small set of equations regarding the interactions of different weather elements he built and analyzed a model that turned out to be highly sensitive to initial conditions. Whereas the common wisdom in the physical sciences used to be that small changes in initial conditions in a system should lead to small changes in overall system behavior, his simulations showed that this was not at all necessarily the case. This discovery has lead to a large literature in chaos theory and the related complex systems approach and, primarily in the social sciences, agent-based modeling.[1]

Laver and Sergenti (2007: 2-3) argue that both highly parsimonious formal models and complex computational models can be seen as tools for discovery, to establish better intuition into human behavior. They attribute the substantive intuitions arising from formal models to "informal esthetics and gut feeling", however, while computational models would be more realistic and therefore generating more substantively interesting intuitions and the possibility of "unexpected discoveries and intuitions". The latter is indeed an important aspect of computational modeling, usually under the label of *emergence* (Johnson 2001), but the former is a rather strong and weakly founded claim. There is no reason to believe that more complex implies more realistic and that therefor conclusions drawn from studying such models are

[1]See Holland (1995, 1998) and Johnson (2001) for accessible, popular scientific introductions to the subject.

necessarily more substantively interesting. Since the real world is almost infinitely complex, parsimony is an important tool to highlight particular mechanisms that are likely to explain aspects of human behavior, without attempting to explain entire realities. The more bells and whistles are added to a theoretical model, the more likely it is to deviate from the way different mechanisms relate to each other in reality. With each additional feature usually a substantial number of new assumptions are being made about human behavior, increasing the risk of being unrealistic. The risk of crossing the border between a substantively interesting model and a computer game increases with each added feature.

It could be argued, of course, that some dynamic, agent-based models are based on removing, rather than adding, assumptions to a more tractable formal model. Sometimes the tractability in a formal or game-theoretical model is acquired by adding assumptions about knowledge, certainty, memory, etc., whereby removing those assumptions lead to dynamics that can only be traced through simulations. In this case, agent-based models can be argued to be likely to be more realistic. When agent-based models are created by adding features, like different types of agents or a more explicit modeling of interactions, the more complex model can appear more realistic because it ignores fewer interactions, but it is perhaps likely to be in fact less realistic, as the added mechanisms are all based on additional assumptions about the real world that might not be true. I would argue that more complex models are less likely to be realistic, but models that remove dynamics by adding assumptions about behavior, are less likely to be realistic than models that do not. A straightforward correlation of computational models being more realistic and formal models being less realistic, as claimed by Laver and Sergenti (2007), does not seem obvious.

This critique on Laver and Sergenti's evaluation of the purpose of parsimony, however, does not imply a critique on the use of computational as opposed to formal models for theorizing. The fact remains that some mechanisms, however simple in design, can lead to complex and mathematically intractable dynamics. In fact, the very idea that the interaction of very simple patterns of behavior can lead to very complex patterns in a system as a whole is what underlies the field of complex systems. Parsimony thus does not relate directly to tractability. Agent-based or computational models can be used to study such parsimonious, interactive, and complex systems. The model presented here is an example of a model where the building blocks are relatively parsimonious, yet the interaction between the various elements too complex to analyze with conventional game-theoretic tools.

6.1.2 Empirical validity of agent-based models

While agent-based models have become gradually more common in the social sciences, little attention has so far been paid to the empirical validation of these models. As any kind of theoretical modeling, a model is only useful from a scientific perspective if it can be corroborated with or at least falsified by empirical research. Many agent-based models, however, suffer from a serious lack of falsifiability (Popper 1962). Agent-based approaches are generally used to model nonlinear patterns that are impossible to predict and difficult to measure, but we will argue that this does not imply that empirical validation is entirely impossible.

When talking about the validity of a theoretical model one can make a distinction between two types of validity, internal and external or empirical validity. The internal validity relates to the internal consistency of the model, to the extent to which the underlying assumptions of the model indeed lead to the outcome as predicted by the theorist. Simply due to the fact that an agent-based model as the one described above is implemented in a computer simulation, there is little doubt as to the internal validity for the model. The mechanism of the model has been explicitly implemented in the computer code and computers do not make mistakes, i.e. these mechanisms really do lead to the outcome as observed in the simulations. The one major exception here is of course mistakes in the code itself and the more subtle effects of inevitable rounding - after all, the whole theory of chaos started with the rounding of a few numbers in Edward Lorenz' simulations in 1960. The extent to which such errors affect the simulation is difficult to gather, but one solution might be to implement each simulation always in different programming languages. The simulation above was originally written in Java, using the RePast library (North, Collier and Vos 2006) because of its extensive graphical capabilities, and then rewritten in C++ for faster simulations, with similar[2] results.

In terms of internal validity, agent-based modeling is a relatively easy type of modeling, since a large part of the internal consistency is guaranteed by the method. For external validity, however, this approach to modeling becomes rather complex. Agent-based modeling is invented in part to deal with the complex nonlinear patterns that arise from the interactions between large numbers of individuals. These nonlinearities lead to patterns that are unpredictable and difficult to capture with straightforward models that can be statistically tested. Nonlinear systems tend to be highly sensitive to initial conditions, thus small changes in parameters or circumstances can lead to dramatic changes in outcome, rendering the use of standard statistical

[2]Similar rather than identical, since there are influential random elements in the model - one cannot establish that the simulations are in fact identical on the basis of the simulation outcomes. High levels of similarity provide the necessary trust in the simulation implementation, however.

techniques virtually impossible.

It would be a mistake, however, to simply conclude that such empirical testing is impossible and consider the analysis of the simulation results as the end stage of the research cycle. The unpredictability of the outcome of nonlinear systems does not imply that there are no observable implications to the model that can be empirically tested. In this respect the same critique as delivered by Green and Shapiro (1994) towards rational choice models can be applied to recent developments in agent-based modeling as well: the fact that models are internally consistent is not a sufficient validation and empirical validation is a crucial next step. Few examples of empirical tests of agent-based models can so far be found.

Even a simple model as that of Schelling (1978), describing how the preference of individuals to have at least some neighbors of their ethnic group already leads to an almost complete social segregation of ethnic groups, could well be corroborated with empirical data. Survey data could demonstrate the prevalence of such attitudes towards neighbors and ethnic groups, while regional statistical data could demonstrate the levels of segregation in communities. The link between the two can be demonstrated by an agent-based model as the one developed by Schelling, while the empirical validity of the model can be verified by statistical data both at the micro and at the macro level. The agent-based model can thus establish the plausibility of the supposed interconnection between different elements of the dynamics, while the empirical tests can validate these segments individually, both at micro and macro levels of social systems. This combination of agent-based modeling and partial empirical validation could thus contribute importantly to the attempt to link the gap between macro and micro level theories.

One strategy for establishing the empirical validity of an agent-based models would be to look at what Moss calls 'statistical signatures' (Moss 2001; Laver 2005). The idea behind these statistical signatures is that although we cannot empirically validate directly the theoretical model we have, we can look at the distributions and other statistical properties of particular variables of interest in the simulation and see if these match with real world distributions. For example, in the example of Laver (2005) we can see whether parties do indeed have changing numbers of supporters in opinion polls as suggested by his dynamic model of party competition. Such a mapping is a relatively weak corroboration of the model, however. Many mechanisms could lead to a similar level of variation. Furthermore, the amount of variation is often due to very arbitrary aspects of the simulation, such as the scale on which particular things are measured or the number of agents in the simulation. Alter these and the resulting levels of variation change. The statistical signatures approach is thus both very limited and somewhat unreliable.

Agent-based models in the social sciences are particularly useful for establishing the link between individual, micro-level behavior to overall, macro-level behavioral patterns. How does the decision-making of individuals entering a cinema affect the order in which seats in a cinema are filled? How can counting the number of passers-by in one street leading to a square help in estimating the overall number of people on the square itself? How can exit-routes in a town be designed in such a way that individual choices and behavior in case of a major calamity will lead to the most efficient emptying of the town? Most agent-based models are concerned with this type of micro-macro link, a link in general most problematic in social science analysis. Most other frameworks approach major social phenomena either as purely macro-level concerns, e.g. by explaining them through class relations or the interaction of groups in the elite; or by purely looking at individual-level explanation, e.g. in social psychology; or by reducing the individual interactions to two, three, or four-player games, neglecting the complications due to interactions between large numbers of actors. Agent-based models come in exactly at that point, where the concern is with the interaction between large numbers of agents and their macro-level effects.

This positioning of agent-based modeling in the theory-building context also has implications for the empirical validation of such models. The actual mechanism that links the micro-level behavior with the macro-level patterns is generally difficult if not impossible to capture with empirical data analysis. Scenarios could be imagined where the relations suggested by the theoretical model can indeed be captured through statistical analysis using data measured at both individual and aggregate levels, but in these scenarios one can likely conclude that the agent-based model can be replaced with a more tractable one. Developing the agent-based model might have helped arriving at this stage in a theoretical development, but it does imply that the agent-based model is in the end unnecessary and thus not optimal. In cases where the agent-based modeling strategy is indeed warranted, capturing the overall mechanism with a statistical model is likely to be impossible.

Given that we cannot capture the entire theoretical model in a statistical one does not imply that we should give up empirical validation altogether, however. In general, causal inference requires theory (Morgan and Winship 2007), and the agent-based model fulfills exactly that role. What empirical analysis can do in terms of validating the theoretical model, however, is to find corroborating evidence with regards to both the micro-patterns assumed and the macro-patterns observed in the simulations, separately. The two sides of the story, micro and macro, lead to probably quite different empirical strategies for validation. The overall strategy would be one of triangulation, i.e. find as much corroborating evidence as possible that supports the main mechanisms and assumptions of the model, either separately or in terms of how they relate to each other.

Although further research is definitely required in this area, in particular in terms of providing an actual example of the idea in practice, one would suspect that agent-based models can be best validated by validating subparts of the model separately, in particular the micro level mechanisms separately from the macro level patterns. If we take Schelling (1978) as example again, we can check the individual attitudes towards other ethnic groups and we can look at the macro level patterns of the housing market. The link between the two can be provided by theory or, indeed, Schelling's agent-based model. The link between the two is non-linear and difficult to capture with an empirically testable model, but the two mechanisms independently are not. The model can provide the linkage, while the data can provide the corroboration for the assumptions made about individual behavior.

6.2 Future research

The theoretical and methodological framework of this thesis is a combination of a number of different fields that normally develop independently of each other. The first is the literature on democratization, which is the prime literature to which this thesis contributes. This literature attempts to explain why some countries democratize and others do not, when this is most likely to happen, and under what circumstances these new democracies are most likely to survive on the long term. The second literature is that related to the way attitudes spread between individuals in a community, either at a very local level or indeed the international diffusion of attitudes. This somewhat smaller literature is concerned with how and when attitudes diffuse, what kind of circumstances enable or catalyze this process, and what factors constrain it. A key puzzle here is why, if attitudes indeed diffuse between individuals, the end result is not one "average attitude" held by every individual in a community, but a persistent polarization of attitudes (Axelrod 1997*b*; Jager and Amblard 2004). The third field is the methodological approach taken in this thesis, based on the relatively new but quickly growing agent-based modeling strategy. Given the still somewhat unconventional nature of this approach, it makes sense not to see this as a mere tool, but to reflect explicitly on its usefulness and future. The fourth area that will be briefly reflected upon in this section is the area of spatial econometrics, in particular the estimation of patterns of diffusion whereby the observed dependent variable is dichotomous in nature. Finally, one could argue that the diffusion literature is a separate literature that deserves comments in terms of future research, but the way diffusion is treated in this thesis, it is entirely incorporated in the first two literatures. A large part of the diffusion literature is concerned with learning and emulation, for example in terms of state policies (Volden 2007), but exactly to this part this thesis does not substantially contribute.

In this section we will comment on the various venues for future research highlighted by the research in this thesis. We will first turn to the next step in this research project itself, namely the empirical validation of the model presented here. Subsequently we turn to the democratization and diffusion literatures and the agent-based modeling and spatial econometric literatures to discuss some ideas for future research in each of these areas.

6.2.1 Empirical validation

As discussed above, the strategy suggested here for the empirical validation of agent-based models is to separate the different elements in the model that might be related to each other in a non-linear way, but that are more 'linear' of themselves. In most cases this will be the micro versus the macro or system-wide parts of the model. In terms of the macro-level patterns, two main types of observations appear most suitable for empirical exploration: time series of key variables in the model or the main dependent variable and correlations between core variables of the model. For the model presented in this thesis, the presentation of the time series result has been the starting point rather than the validation of the model. The empirical time series of both the level and the clustering of democracy are presented in figures 1.3 and 1.2. Matching the simulation output with those plots is thus the first step towards an empirical validation of the model. Since the model is designed to match these graphs, however, this cannot be seen as much more than a first step. Additional implications of the model at macro level could be derived that would further corroborate the findings that were not the starting point of the research project. To do this type of correlational research for this model would have to be an important first step in evaluating the model - the empirical validation process thus starts with some further analysis of the simulation output itself. This would include correlations and patterns that might not be substantively interesting - the purpose of finding them is to corroborate them with empirical evidence so as to support the overall model.

Although in general the distinction between micro and macro level patterns of behavior makes sense in the interpretation and validation of agent-based models, in this particular case it would be fruitful to distinguish one more level, a meso level. The model contains in effect three different levels at which behavior can be observed: the individual (micro) level, the international (macro) level, but also the country (meso) level. Some aspects of the model would have to be validated at this particular level, for example when it concerns the mechanisms of the cascading model of revolution, or particular variables that the model suggests are related to the timing of transitions. In terms of the cascading model it is particularly difficult to get any empirical validation: after all, the model itself predicts that to a large extent what is happening is invisible and impossible to measure. Citizens keep their true opinions private and tell public lies (Noelle-Neumann 1993; Kuran 1995). Hence the conclusion that we should not be surprised that we did not see the collapse of the communist system in most of Eastern Europe coming (Kuran 1991*a*).

For the remaining issue, the micro-level idea that attitudes towards democracy diffuse across borders, one could use international survey data measuring such attitudes. For example, use could be made of the World Values Study (Inglehart 2000).[3] An attempt could be made to find an underlying attitude towards democracy dimension along the lines of what Treier and Jackman (2008) do to find the underlying democracy score in the Polity IV data and subsequently investigate whether this underlying attitude diffuses geographically. If the model holds, we would expect the average levels of these attitudes in countries to show geographical clustering patterns and we would expect that individuals who are more likely to be exposed to foreign influences to be more affected by these foreign attitudes. For example, more wealthy individuals and migrant workers should be more likely to have attitudes similar to near foreigners.

[3]The attitude towards democracy could for example be measured using an index composed of the scores on the statements: "In democracy, the economic system runs badly", "Democracies are indecisive and have too much squabbling", and "Democracies aren't good at maintaining order".

Since in this case the macro patterns were the starting point of the simulation, stronger evidence would be found at the micro level. The patterns at the macro level were simulated using assumptions and theories concerning individual level behavior - validating these would provide strong evidence for the simulation results. §6.2.3 will discuss in more detail the possibilities for further research with regards to the social judgment theory, which forms a key foundational part of the theoretical model. This theory has been the key vehicle to get around the concern that attitudes do not, in fact, all converge to the same mean attitude, but that a diversity of opinions persists in a population. Further research on the validity of this theory, as well as the presumed theoretical implications in terms of diversity, would implicitly also help corroborate, or falsify, one aspect of the theoretical basis of this model of democratic diffusion. In terms of the social judgment theory, also, empirical testing is not very straightforward. One would have to look into relatively expensive experimental research designs rather than available survey data.

Perhaps the most promising approach would be when prior to the empirical data analysis a clear specification exists as to what one should observe to disqualify the theoretical model. What patterns, both at micro and at macro level, should one *not* observe if one assumes the model to be correct? A number of suggestions come to mind:

- The model assumes a reasonably strong relationship between public opinion and regime stability. Although it is not necessarily suggested that most regime transitions are popular revolutions, it is still assumed that public opinion matters, that it is difficult for leaders to stay in power when being largely unpopular. A single country where a dictator manages to stay in power despite lack of popular support would not falsify the model, but a clear lack of correlation between popular support and regime stability would cast doubt on the main findings.

- Although harder to establish empirically, if we would find that individuals adjust their opinions more strongly on the basis of inter-personal communication than on the basis of democratic propaganda, it is unlikely that the model presented here can explain the global rise and clustering of democracy that we observe. The simulation results show clearly that the democracy promotion effect has to be stronger than the inter-personal communication effect ($B > \delta$).

- In terms of the overall trend, the way democracy promotion has been implemented in the model here does imply a steady increase of the number of democracies over time, or at least a more or less stable level given particular parameter configurations. If the current upwards trend of the number of democracies would end in the future and a decline would be visible, this would cast some doubt on the model as well.

6.2.2 Democratization and autocratization

In relation to the literature on political transitions, the theoretical model of this thesis makes the assumptions that (1) public opinion matters crucially for the survival of a political regime, that (2) public attitudes towards political regimes diffuse through communication with both fellow citizens and individuals abroad, and that, in addition, (3) international propaganda affects these attitudes. It also suggests, or rather, assumes, that this international propaganda is asymmetric, that democracies spread their ideas more than do autocracies. Whether framed as an empirical validation of the model itself, as in §6.2.1, or as a matter of future research in the area of transition studies, these assumptions lead to interesting further questions.

The idea that public opinion would at least to some extent matter for transitions to and from democracy would not be controversial in the transition literature. Political actors that can claim popular support are generally in a stronger position than actors that cannot. Furthermore, in many contemporary cases of democratization, elections played a crucial role. 'Electoral autocracies' or 'illiberal democracies' are becoming more and more common, where a dictatorial leader attempts to strengthen his position by demonstrating his power status through holding elections and, usually, committing extensive fraud. In these cases, the political regime runs a calculated risk of the possibility of losing elections in exchange for the ability to demonstrate power - a risk that sometimes fails. Recent democratizations therefore often come about through losing the election. In these cases popular support is thus a crucial factor, as soon as it becomes too weak to compensate through electoral fraud. Serbia, Ukraine and Kyrgyzstan are the most common examples of this scenario.

Given that popular support, or public opinion, matters, however, is a very limited claim. It was sufficient to proceed with the modeling activity presented in this thesis, but rather unsatisfactory for a broader research agenda in transition studies. This thesis thus strongly supports and underlines the claim by Welzel (2006) that the role of public opinion in studies of democratization is under-studied and would be a fruitful area for further research. Although the emphasis on elite-level explanations is reasonable based on the idea that in the end it is the actions of the elite that determine the outcome of the process, the mass-level part cannot be ignored entirely (see also Finkel, Muller and Opp 1989; Muller, Dietz and Finkel 1991; Opp and Gern 1993; Bermeo 1997; Finkel and Muller 1998; Paxton 2002; Welzel 2006).

Several questions could be raised for further investigation. Under what structural conditions is mass opinion more important in transitions? What are the constraints on the relationship between public opinion and regime transitions? How does the transition process itself affect public opinion? What is the role of feedback in this process? What types of transitions are more affected by public opinion? What types of actors or actions are most affected by public opinion? All these questions are under-studied and relevant and thus open for further research.

What the above discussion of the role of public opinion in political transitions brushes over is what this opinion really is. What attitudes are we talking about? §3.1.1 provides a brief discussion of what kind of attitude is meant by a 'pro-democratic attitude', but it also emphasizes the multi-faceted character of this attitude. Citizens cannot be assumed to have a one-dimensional, single attitude towards democracy in general. For the sake of brevity, or parsimony, it makes sense to assume such a one-dimensional attitude in the development of an agent-based model. To what extent such an attitude exist, or at least an underlying dimension along these lines in the attitude structure of an individual, is a matter for empirical research. Further research in this area, perhaps using a latent variable approach along the lines of Treier and Jackman (2008) on data measuring democratic attitudes such as those reported by Welzel (2006) could be very helpful in the study of the role of public opinion in political transitions.

The 'broadcasting' concept implemented in this model would be an area of primary interest in the study of the diffusion of political regimes, one that has thus far been largely neglected. Although a reasonably extensive literature exists on the relation between political and developmental aid and democracy promotion, less literature is available on more direct forms of democracy promotion. The 'broadcasting' concept in the model presented here is also more directly related to attempts to influence individual opinions of citizens of non-democratic countries in the neighborhood of democracies. The aid literature, in contrast, focuses primarily on the kinds of sticks and carrots that democracies can apply for the political elites of these countries, through the application of conditional aid. The literature on more grassroots oriented attempts to affect democratic attitudes is much more limited and an interesting area for future research.

Finally, when further studying the relationship between the diffusion of individual norms and the empirical clustering of democratic regimes, more attention should be paid to other factors that affect this relationship. For example, how does the diffusion of attitudes towards democracy interact with the more general clusters of cultures? Cultures could be interpreted as a large set of different attitudes and ideas that might diffuse independently (see, e.g., Axelrod 1997*a*) - how would this concept of attitudinal diffusion then affect the model presented here? And what is the interaction of this kind of diffusion with geographical factors that affect the accessibility of particular countries or particular borders between countries (Starr 1991)? And in addition to geographic proximity, what other ties between nations would affect the diffusion of pro-democratic attitudes? One could think of trade relations, common religions or language, former colonial ties, or perhaps even plain similarity between countries. Further research into the interaction between such factors and the spread of norms towards democracy, or indeed the clustering of regimes more in general, is certainly warranted.

6.2.3 Attitudinal diffusion

The previous section discusses the possibilities for further research in terms of the role of public opinion in democratization studies. As mentioned, the model is based on the assumption that attitudes diffuse through interpersonal communication, both across and within international borders. This section will pay further attention to this part and see what the model suggests for future research.

One obvious aspect of interpersonal attitudinal diffusion is the relation between such diffusion and methods of communication. The rise of the Internet in recent years has had a huge impact on international communication across borders. The diffusion of norms and values through communication is thus likely to have a gradually less local character. The way the rise of the Internet affects models of attitudinal diffusion is a matter of future empirical research. It is tempting to make sweeping statements about its impact on the basis of common sense, but extensive empirical research appears to be missing. Lack of methods of communication, for example through government restrictions, can also affect the diffusion of norms towards the political regime. For example, Opp and Gern (1993: 662) suggest that the restrains on tools of communication enforced by the East German state severely affected the extent to which individuals could communicate and mobilize for protest. The role of social networks is likely to be more important in societies where other forms of communication or media access are limited. For example, in the Leipzig demonstration in 1989 in the GDR, personal ties appeared to be a particularly important mobilizing factor, probably because of the restraints on, for example, the telephone, enforced by the regime (Opp and Gern 1993: 673-674).

As discussed extensively in §3.1.3, one of the issues of interest in the literature on attitudinal diffusion is the fact that both local clustering and global polarization are visible patterns in norms and values. The social judgment theory as employed by Jager and Amblard (2004) as well as the norm diffusion model of Axelrod (1997*a*) are attempts to deal with this issue. In both cases, particular channels of communication are blocked - in both cases individuals do not communicate with other individuals when they are too different from themselves exactly on the attitude or norm that is diffusing. Because of this restriction on communication, global polarization can persist despite the strong homogenizing force of the diffusion process. Both models are somewhat unsatisfactory, however, as they make very strong assumptions to be able to reach their findings. Future research, along the lines of §4.3.1, is definitely required to study further what the limitations are of these models - what assumptions can be relaxed, and how far, before the results of the simulations break down? For example, one would expect communication between individuals who are very different to be unlikely but not impossible. When one implements this, however, the results of both models will collapse entirely. What would be particularly interesting to study is whether a combination of the two models - a model along the lines of the social judgment theory, but with localized rather than global paths of communication, more akin to the Axelrod model - can bring us a step further.

6.2.4 Agent-based modeling strategies

In terms of the methodology used in this thesis, one area that certainly deserves more attention and specific research is the empirical validation of these models. Extensive attention has been paid to this topic already in the previous paragraphs (§6.1.2 and §6.2.1) and the arguments presented above can be straightforwardly extended to a question of future research. What are the various tools available to corroborate results from agent-based models? What are the levels to look at? How do we decide what aspects of a model are not empirically verifiable due to their non-linearity and complexity and what aspects are? How does one determine what observations would falsify a model? How does empirical validation of an agent-based model work when the assumption is made that the model is a parsimonious representation of only part of an explanation? All these questions are interesting for future research. This section, however, will concern itself with other areas of future research on agent-based modeling that have received less attention in previous paragraphs.

The first question that arises and that is already discussed fairly extensively in the literature on agent-based models or complex systems approaches more in general is under what circumstances such approaches are most useful and how. For example, debates exist as to whether agent-based models are primarily of interest purely as a thought experiment or to generate new intuitions about social mechanisms, or whether such models are also some form of proof, if not empirically then at least of the internal consistency of a theory. Furthermore, should agent-based models be applied very widely because of the likely inherent non-linearity of many social patterns, or should they be applied very sparsely only in circumstances where the search for parsimony leads to an intractable model. I would certainly be inclined to discourage applications where there are many assumptions made about the real world, which is likely to lead to a less and less realistic model (see, e.g., Epstein and Axtell 1996). The logic here is perhaps somewhat counterintuitive: to understand the complex world better we need simple models. Complex models are less likely to correspond to the real, complex world than are simple models.

This leads to the next big issue in agent-based models, which has hitherto drawn very little attention. If we hold that agent-based models are a type of game-theoretic models, with the one distinction that they are mathematically intractable, usually due to the larger number of agents, then how do we know whether the model is indeed intractable? To my knowledge, no agent-based model provides any mathematical evidence that it is indeed intractable other than through simulations. A relatively simple model such as that by Jager and Amblard (2004) would certainly raise doubts whether simulations are really necessary. They help the intuition of the researcher at first instance, but once the results are clear, it is not obvious that we cannot provide the mathematical proof of those results, without actual simulations. The locations of the various norm groups in the polarized outcome of the simulations appear to be very constant and, in retrospect, predictable. This is not to suggest that the researchers were wrong in doing these simulations or that they were useless, but it does suggest that perhaps, in retrospect, the model could be mathematically tractable, and if it is, that is how it should be presented.

The models by Jager and Amblard (2004) and Axelrod (1997*a*), to stick to the same examples, also suggest another possible critique of agent-based modeling. In statistical analysis, there is always an assumption made that the model only explains part of the data. The rest is captured into an error term, which captures all factors related to the dependent variable but not explicitly part of the model. This renders the conclusions from these analyses more robust: the fact that other factors play a role is acknowledged and the conclusions drawn are despite these factors.[4] Agent-based models, however,

[4]Of course, there are many further complications with regards to endogeneity, interdependence, etcetera, that this argument brushes over, but the principle in general still holds.

are generally purely deterministic. The model covers the entire process, no other factors are acknowledged. Perhaps most modelers would not claim that the real world is indeed so deterministic and that their model is the only mechanism operating, but in terms of the design of their model, they do make this assumption. In some cases, including these two models, this assumption can drive the results. Both of these models do not contain anything akin to an error term. In Axelrod (1997*a*)'s model, agents *never* communicate with other agents that are entirely different in cultural norms. If they would do so with just a very low probability, the model would again lead to one homogeneous result. No cultural groups would remain. Similarly, in Jager and Amblard (2004)'s model, agents *never* adjust their attitudes towards the other when communicating with another agent outside their latitude of rejection. If an error term was added here, the results would most likely collapse as well. This reliance of the model results on the lack of exogenous factors implies that it is less likely to hold in real life as a model of attitudinal diffusion. The models are not very robust. More attention should thus be paid to how 'error terms' can be added to agent-based models or at least how the robustness of the model can be checked for such exogenous factors.

When adding these 'error terms' it also creates an additional concern to be wary of. These exogenous factors will generally be implemented as random shocks to the system. The analysis for the results will then have to be concerned with distinguishing the effects of these random shocks from the more systematic patterns of the simulation. In an earlier iteration of the model as presented here, clear waves of democracy were visible. It was gratifying to see both the spatial and the temporal patterns of clustering so clearly visible in the simulation output, only until the discovery that these 'waves' could be entirely explained by the exogenous shocks, the 'coups' in the model. In other words, I was looking at random walks of democracy! Distinguishing a random walk from a patterned one can be very difficult, however, especially when there are strong stochastic factors incorporated in the model. This issue is probably of much larger concern than just the area of agent-based modeling, but it would be advisable for agent-based modelers to take the literature on the detection of random walks into account.

Finally, and perhaps related, there is the issue of how to interpret and present the results of agent-based models. In this thesis the main conclusions are based on a parameter sweep with simple summary statistics for each parameterization, combined with more inferential statistical models to get at the more complicated patterns and to deal with the large parameter space in the Monte Carlo approach to setting the parameters. For the research presented here it turned out that the sweep approach with straightforward descriptive statistics was more informative than the statistical modeling approach and random parameter settings. Arguments concerning the non-linearity of such models and the possibility of missing these non-linearities due to skipping large parts of the parameter space are still of concern, however. Further guidelines on how to study the output of simulations would be useful for the further expansion of the application of agent-based models in the social sciences.

6.2.5 Spatial econometrics

In the area of statistics or econometrics, this thesis also suggests another area where spatial dependencies are of substantial interest. The presence of such spatial autocorrelation - the interdependence between spatially proximate units - has implications for the empirical estimation of models where these spatially correlated data form the dependent variable of the model. It leads to inconsistency and bias in the estimates of the coefficients for the independent variables (Anselin 1988: 59) when ignoring the spatial autocorrelation in the errors of the model. Furthermore, in most cases the spatial effect will itself be of interest, and should thus not be ignored either by treating it as a nuisance in the error term (Beck, Gleditsch and Beardsley 2006; Beck 2007; Franzese and Hays 2007*a*) or by estimating a fixed or random effects model that can, in particular circumstances, lead to consistent albeit perhaps inefficient estimates of the effects of exogenous variables (Case 1991; Swinton 2002).

When the complications of spatial autocorrelation are avoided by correcting in the error term or by estimating a fixed or random effects model, one loses efficiency as well as the ability to learn about the structure of the spatial effect. One might for example be interested in comparing different types of connections between units, in the strength of the spatial effect, or in the interaction between characteristics of a particular unit and the spatial autocorrelation. Furthermore, the distinction between spatial correlation in the outcome variable or in the error can often be substantively of interest. Finally, in some cases the use of fixed effects makes it difficult to deal with independent variables of interest that do not vary within units (Case 1991: 959-960).[5]

Spatial autocorrelation is to a large extent comparable to serial autocorrelation, on which a much more extensive literature is available. There are three key differences however, that complicate matters considerably. Firstly, serial correlation is directional: y_{t-1} might affect y_t, but the reverse is not possible given common assumptions regarding causality. In spatial autocorrelation, the effect is two-directional (or non-directional): under most assumptions, if

[5]See, however, Plümper and Troeger (2007) for an alternative solution to dealing with variables that do not vary within the fixed effects units.

y_x affects y_s, the reverse is also true, or at least indistinguishable. Secondly, due to the geographic nature of spatial autocorrelation, the effect is two- rather than one-dimensional. When a more abstract, network-based view of space is taken into account, one could even speak of a multi-dimensional framework. Thirdly, spatial observations are rarely observed on as regular a grid as the time intervals in most time series observations (Pinkse and Slade 1998: 127).

In many of these studies where a spatial dependence structure can be assumed, the outcome variable is binary in nature. A state implements a particular policy or it does not, it has particular institutions or it does not, etcetera. The study in this thesis is a typical example of such a variable, where we measured countries as either being democratic or not. It is generally accepted that in such a context the application of a regular ordinary least squares regression is not appropriate and probit and logit models are more common. In the case of spatial autocorrelation, regular probit estimators are consistent but inefficient, since for an efficient estimator it is necessary to condition on all values of y, not just y_t (Poirier and Ruud 1988). In most realistic spatial econometric contexts, however, the spatial interdependence also leads to heteroscedasticity, which in turn leads to not only inefficient, but also inconsistent estimators (McMillen 1992, 1995). The literature on the combination of a binary dependent variable combined with spatial interdependence is sparse, however, and thus far largely disconnected from the political science and policy studies literatures, with the exception of Franzese and Hays (2007*b*).

Franzese and Hays (2007*a*) discuss the application of spatial econometric models in political science and provide suggestions for the interpretation and presentation of the coefficients of such models. While the effect of spatial autocorrelation is similar in some respects to time series autocorrelation, in particular the explicit presence of feedback creates complications specific to spatial models. In a subsequent paper they extend their work to the binary context (Franzese and Hays 2007*b*), but further research in this field of econometrics is certainly warranted. Only limited information is available on the performance of various different estimators and the circumstances under which those estimators perform well or badly. The field of spatial econometrics - and the related statistical analysis on social networks - is relatively underdeveloped and further research is certainly necessary. In this thesis we got around the problem, in line with Gleditsch and Ward (2006), by assuming that the spatial dependence disappears when controlling for the temporally lagged spatial variables. In other words, we did not related $y_{x,t}$ with $y_{s,t}$, but rather $y_{x,t}$ with $y_{x,t-1}$. This solves most of the issues in terms of estimation, but it would be better if econometric models can indeed deal with spatial interdependence proper. For this, further research is indeed warranted.

6.3 Conclusion

In this last chapter of the thesis we discussed extensively what questions this research raises both in terms of substantive and in terms of methodological issues. It is clear that the model presented here leads to perhaps more questions than that it answers. Considering the cumulative nature of science a simulation experiment that leads to new venues and critical comments regarding the applied method is itself a constructive contribution to the social sciences. It serves no purposes to repeat all points mentioned above, but it is clear that especially in methodological terms there is still a lot to be explored within this new tool in social science research, agent-based modeling. The research also re-emphasizes questions of importance to the study of processes of democratization, especially the somewhat ignored factor of public opinion. The model presented here contributes to the literature on norm diffusion and especially the question under what condition such diffusion leads to a globally polarized but locally homogeneous distribution of norms and also this area is still open for further contributions.

CHAPTER 7

Conclusion

Whereas in the previous chapter we focused on the technical or theoretical implications of the model and the avenues for future research, we will conclude the dissertation in this chapter by concentrating on the more practical conclusions we can draw from the simulations. What do the results of this modeling exercise suggest for the future of democratization in the world and what would be the practical policy implications of these results? The idea of the diffusion of democracy, or the idea that countries affect their neighbors in terms of their political regimes, has policy implications in particular in terms of foreign policy. The relative importance of a particular transition to democracy increases if this transition, in addition to its intrinsic value, increases the probability of future transitions in the region. The theoretical model also sheds some light on the effectiveness and role of more direct policy tools, such as campaigns to promote popular democratic attitudes abroad. Keeping the many caveats presented in the previous chapter regarding the validity of the results in mind, we will first briefly reflect on the implications of these results for the future of the international spread of democracy and subsequently to the more immediate policy implications.

7.1 The future

Attempting to predict the future is always a hazardous activity for a social scientist and any such prediction should always be taken with a large grain of salt. This section will therefore be kept very short. Some observations on the basis of these simulations are interesting to make, however. The most clear prediction in terms of the future is that in any scenario in the model where spatial clustering is visible - and we do see this spatial clustering in the real world - the world turns eventually fully democratic. The speed of this process is determined by the relative strengths of democracy promotion versus interpersonal communication in attitude formation, but the end result is constant: almost full democracy. The overall outcome of the model is not at all very dissimilar from that presented by Modelski and Perry (1991, 2002) - we see the same typical S-curve of diffusion.

It should be noted, however, that the model is very much based on a local diffusion of attitudes. Perhaps the explanatory power of the model will reduce over time as increased methods of virtual communication render locality gradually less important (see, e.g., Rosenau 1988: 359). The growing importance of the Internet results in a reduced relevance of distance, at least in terms of opinion formation. Many forms of communication take place over long distances. Increased globalization, however, also leads to a stronger tendency to emphasize the local identity, including probably values, hence including attitudes towards democracy. For example, perhaps the increased levels of globalization lead to a reduction of pro-democratic attitudes in the Muslim world, as their lack of democracy is part of their identity, of their otherness from Western liberal democracies. The fear for losing one's own identity in a globalized world encourages people to emphasize local norms and communicate them more strongly to their peers. The Internet can easily have an effect in both directions and which one is more prominent is a concern for empirical research.

7.2 Policy implications

In attempting to relate policy implications to a model that attempts to provide an explanation for the temporal and spatial diffusion of democracy, one can look at two different types of arguments. The first is the extent to which we can count on diffusion to help the promotion of democracy. To what extent can and should we focus on the stimulation of democracy in a particular country in one region instead of the region as a whole. Changing the regime in one country, or at least adjusting policy to be as stimulating for democratization as possible, is presumably a lot easier and cheaper than attempting to affect these processes in a large number of countries at once. Thus, if one could influence democratization in one country in a largely non-democratic region and subsequently observe the diffusion of this new democratic presence across the region, it could be advisable to indeed focus such efforts on a small number of countries. Explaining the clustering of democracies as an effect of the chances for survival of new democracies (Cederman and Gleditsch 2004) would not suggest such a possibility - do the simulations presented above suggest an alternative interpretation? The second interpretation is that if the model suggests an explanation of temporal and spatial diffusion that is similar to the patterns we clearly observe in empirical data, then this might suggest that these mechanisms as modeled are indeed present. The macro-level match between model and empirics would suggest a similar match between micro-level mechanics of the model and the empirical world. Such a model would then suggest specific interpretations of processes of democratization and have particular policy implications for democracy promotion independent of the process of democratic diffusion itself. The latter interpretation is arguably somewhat more contingent on a future empirical validation of the model than the former.

In this section we will look at these two types of implications for two potential areas of policy where an effect on processes of democratization can be expected. The two areas relate to the two main mechanisms in the model for the diffusion of democracy: informal communication between citizens across international borders and the active promotion of democracy by democratic regimes by way of affecting public opinion towards democracy in neighboring countries. The first would imply policies by democratic regimes that would increase the probability of communication between citizens of authoritarian countries with the democratic counterparts. Opening the borders and increasing the possibilities for cross-border communication would be the prime focus of democracy promoting policies. The second would imply less of a focus on regime openness and more on the direct influence of public opinion in the non-democratic regime, for example by establishing media that can reach these citizens or by sending agents to mobilize democratic elements in these societies.

7.2.1 Channels of communication

Policies that are aimed at stimulating communication across international borders between citizens could take many possible forms. One clear example would be to stimulate international trade between democracies and non-democracies. Given the likely mutual economic benefit of international trade, one would imagine it easier to convince a non-democratic political elite to slightly open up their borders for international trade than it would be to convince them of the need of a transition to democracy. Such an increase of trade would likely involve an increase in the extent to which citizens of the non-democratic and the democratic country interact, increasing the potential for cross-border individual communication. Another example of such a policy would be to stimulate tourism from ones own (democratic) country to the non-democratic country targeted for democracy promotion. The increase of tourism could lead to more communication between citizens and to an exchange of ideas, including political ones. Yet another interesting example would be to allow large international events like the Olympic Games to take place in non-democratic regimes one hopes to affect, which would have similar effects to the increase of tourism. These policies are clearly very different from an active promotion of democracy and in effect constitute a cooperative rather than adversarial attitude towards the non-democratic regime, albeit with the long term agenda of bringing about a change of regime.

To what extent increased cross-border communication is likely to contribute to transitions to democracy would affect the calculation in terms of the benefits of such policies. If increased cross-border communication does indeed lead to increased chances of democratization, such policies could be defensible towards pro-democratic groups, while if such an effect cannot be expected, these policies are likely to be considered undemocratic. If increased trade leads to increased communication and in turn to increased probabilities of democratization, this would be a very different story from such policies helping the economic position of the affected regime. If these policies do not improve democracy, they are likely to have an opposite effect in that they stimulate the economies of non-democratic regimes.

One clear finding emanating from the simulation results is that the spatial clustering we observe in the real world only appears in these simulations where the democracy promotion effect is stronger than that of communication between individual citizens. This suggest that the diffusion of ideas through interpersonal communication, be it through tourism or trade relations or family relations, in fact counteracts the process of the clustering of regimes. Although this appears counterintuitive at first sight, it is understandable when one thinks through what the effect could be. At this stage it is useful to recall, again, the models of Axelrod (1997*a*) and Jager and Amblard (2004)[1] and the logical observation these models try to circumvent. When one assumes that attitudes simply diffuses by attitudes becoming more similar whenever two individuals communicate, the straightforward prediction for a society as a whole is that everybody will end up with an opinion close to the average. With attitudes near the average, there are no extreme opinions present and nobody is strongly enough opposed to the current regime to instigate a revolution. A flat distribution of opinions means no revolution if the model of the cascading revolution holds true to any extent. The straightforward diffusion of attitudes leads to a reduction in the chances for revolutionary behavior and to a reduction of the tendency of countries bordering on democracies to make transitions to democracy themselves. This result would strongly argue against policies that increase the interaction between democratic and non-democratic citizens. These policies are unlikely to lead to an increase in the probability of democratization, while they are likely to reduce the negative effects of the presence of the non-democratic regime on its citizens. It will have an anti-democratic rather than a democratic effect.

[1]See §3.1.3 and §4.3.1, respectively.

It also suggest that one cannot rely on this mechanism to help democracy diffuse in an area after one country has made a transition to democracy. To force one country in a largely non-democratic region to install democracy, for example by military occupation, might affect the attitudes of the citizens in the country in question, but this does not mean that, through communication, it will positively affect the opinions of citizens in surrounding countries. Based on just the mechanism of the diffusion of attitudes between individuals, one cannot rely on diffusion as a catalyst of international democratization.

7.2.2 Democracy promotion

The broadcasting effect implemented in the simulations, to the contrary, clearly contributes to the diffusion and clustering of democracy. In this case, the effect is much more directed and the tendency of opinions is in only one direction.[2] The term broadcasting in this dissertation refers to a broad category of activities that a democratic regime can perform to actively promote democratic abroad. It encompasses all activities that relate to attempts to change opinions of individual citizens abroad towards democracy. A typical practical example of such an activity would be the support of *Radio Free Europe*, which has likely played an important role in influencing the perceptions of democracy and of their own political regimes by citizens of the then Soviet Union. It would include financially supporting pro-democratic media outlets in non-democratic (neighboring) countries, broadcasting news across borders over radio waves or the Internet, sending political mobilizers abroad to influence opinion, or any other form of influencing public opinion abroad one can think of. The simulations suggest that this has a much

[2]It should be kept in mind that the model as described in this dissertation implemented only democracy promotion and no anti-democratic ideological propaganda by political regimes. The model is clearly biased in its design, in this regard.

stronger potential effect on generating the spatial clustering we observe than the communications between individual citizens.

In terms of the extent to which a democracy promoting country can focus its efforts on a limited number of countries in a largely non-democratic geographical area, one would conclude that simply installing a new democracy is not sufficient. The next step required would be to encourage this newly established democratic regime to actively promote democracy in its immediate surroundings. Simply being present is not sufficient, but a local active effort to promote democracy can have a positive effect. For example, for Iraq to have a democratizing effect on the Middle East as a whole, it would have to actively promote democracy in neighboring countries. After installing the new regime, this would suggest that the United State should focus on encouraging the Iraqi regime to establish such policies, for example by the establishment of Arabic pro-democratic broadcasting media on its territory. Just through tourism and trade relations between Iraq and its neighbors, we should not expect the democratic diffusion to take place.

To the extent that the micro-level mechanisms of the model are indeed correct representations of the empirical reality, we could conclude that the most important policy of democracies to increase the regional level of democracy would be to actively influence attitudes of neighboring citizens. Projects like *Radio Free Europe*, which directly attempt to provide an alternative to the autocratic news media usually present in non-democracies, are key policy instruments to promote democracy. The international presence of democracy could substantially be improved by developing innovative ways of affecting such opinions. In effect, it suggests more focus on democratic propaganda, despite the negative connotations of such a concept.

Bibliography

Acemoglu, Daron and James A. Robinson. 2001. "A theory of political transitions." *American Economic Review* 91(4): 938–963.

Acharya, Amitav. 2004. "How ideas spread: whose norms matter? Norm localization and institutional change in Asian regionalism." *International Organization* 58: 239–275.

Ajzen, I. and M. Fishbein. 1977. "Attitude-behavior relations: a theoretical analysis and review of empirical research." *Psychological Bulletin* 84(5): 888–918.

Almond, Gabriel A. 1989. "The international-national connection." *British Journal of Political Science* 19(2): 237–259.

Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney and D. Sorensen. 1999. *LAPACK Users' Guide.* 3rd ed. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Anselin, Luc. 1988. *Spatial econometrics: methods and models.* Dordrecht: Kluwer Academic Publishers.

Anselin, Luc. 2002. "Under the hood. Issues in the specification and interpretation of spatial regression models." *Agricultural Economics* 27: 247–267.

Arat, Zehra F. 1991. *Democracy and human rights in developing countries.* Boulder: Lynne Rienner Publishers.

Axelrod, Robert. 1986. "An evolutionary approach to norms." *American Political Science Review* 80(4): 1095–1111.

Axelrod, Robert. 1997*a*. *The complexity of cooperation: agent-based models of competition and collaboration.* Princeton, NJ: Princeton University Press.

Axelrod, Robert. 1997*b*. "The dissemination of culture. A model with local convergence and global polarization." *Journal of Conflict Resolution* 41(2): 203–226.

Ayres, Jeffrey M. 1999. "From the streets to the Internet: the cyber-diffusion of contention." *Annals of the American Academy of Political and Social Science* 566: 131–143.

Barabási, Albert-László and Réka Albert. 1999. "Emergence of scaling in random networks." *Science* 286: 509–512.

Beck, Nathaniel. 2007. "From statistical nuisances to serious modeling: changing how we think about the analysis of time-series-cross-section data." *Political Analysis* 15: 97–100.

Beck, Nathaniel, Kristian Skrede Gleditsch and Kyle Beardsley. 2006. "Space is more than geography: using spatial econometrics in the study of political economy." *International Studies Quarterly* 50(1): 27–44.

Beetham, David, ed. 1994. *Defining and measuring democracy.* London: Sage Publications.

Bell, James E. and Lynn A. Staeheli. 2001. "Discourses of diffusion and democratization." *Political Geography* 20: 175–195.

Bergesen, A. 1992. "Regime change in the semi-periphery: democratization in Latin America and the socialist bloc." *Sociological Perspectives* 35: 405–413.

Bermeo, Nancy. 1997. "Myths of moderation: confrontation and conflict during democratic transitions." *Comparative Politics* 29(3): 305–322.

Bollen, Kenneth A. 1991. Political democracy: conceptual and measurement traps. In *On measuring democracy. Its consequences and concomitants*, ed. Alex Inkeles. New Brunswick, NJ: Transaction Publishers.

Bollen, Kenneth A. 2001. "Cross national indicators of library democracy, 1950-1990." Inter-university Consortium for Political and Social Research.

Bollen, Kenneth A., Robert W. Jackman and Hyojoung Kim. 1996. "Suffrage, registration, and turnout: a comparative analysis." Unpublished paper, Department of Sociology, University of North Carolina at Chapel Hill.

Bonacich, Phillip. 2007. "Some unique properties of eigenvector centrality." *Social Networks* 29(4): 555–564.

Box, G.E.P. and Mervin E. Muller. 1958. "A note on the generation of random normal deviates." *The Annals of Mathematical Statistics* 29(2): 610–611.

Braun, Dietmar and Fabrizio Gilardi. 2006. "Taking 'Galton's Problem' seriously. Towards a theory of policy diffusion." *Journal of Theoretical Politics* 18(3): 298–322.

Brinks, Daniel and Michael Coppedge. 2006. "Diffusion is no illusion. Neighbor emulation in the third wave of democracy." *Comparative Political Studies* 39(4): 463–489.

Browne, Edward Granville. 1910. *The Persian Revolution of 1905-1909.* Cambridge: Cambridge University Press.

Bunce, Valerie J. and Sharon L. Wolchik. 2006. "International diffusion and postcommunist electoral revolutions." *Communist and Post-Communist Studies* 39: 283–304.

Burkhart, R.E. and M.S. Lewis-Beck. 1994. "Comparative democracy: the economic development thesis." *American Political Science Review* 88(4): 903–910.

Burton, Michael G. and John Higley. 1987. "Elite settlements." *American Sociological Review* 52: 295–307.

Butts, Carter T. 2006. "Exact bounds for degree centralization." *Social Networks* 28: 283–296.

Butts, Carter T. 2007. "sna: Tools for Social Network Analysis.". *http://erzuli.ss.uci.edu/R.stuff/*

Case, Anne C. 1991. "Spatial patterns in household demand." *Econometrica* 59(4): 953–965.

Cederman, Lars-Erik. 1997. *Emergent actors in world politics.* Princeton: Princeton University Press.

Cederman, Lars-Erik. 2002. "Endogenizing geopolitical boundaries with agent-based modeling." *PNAS* 99(3): 7296–7303.

Cederman, Lars-Erik and Kristian Skrede Gleditsch. 2004. "Conquest and regime change: an evolutionary model of the spread of democracy and peace." *International Studies Quarterly* 48(3): 603–629.

Centola, Damon, Juan Carlos González-Avella, Victor M. Eguíluz and Maxi San Miguel. 2007. "Homophily, cultural drift, and the co-evolution of cultural groups." *Journal of Conflict Resolution* 51: 905–929.

Cliff, Andrew D. and J. Keith Ord. 1973. *Spatial autocorrelation.* London: Pion.

Collier, David and Robert Adcock. 2001. "Measurement validity: a shared standard for qualitative and quantitative research." *American Political Science Review* 95(3): 529–546.

Coppedge, Michael and Wolfgang H. Reinicke. 1991. Measuring polyarchy. In *On measuring democracy. Its consequences and concomitants*, ed. Alex Inkeles. New Brunswick, NJ: Transaction Publishers.

Cowan, Robin and Nicolas Jonard. 2004. "Network structure and the diffusion of knowledge." *Journal of Economic Dynamics and Control* 28(2): 1557–1575.

Crescenzi, Mark J.C. and Andrew J. Enterline. 1999. "Ripples from the waves? A systemic, time-series analysis of democracy, democratization, and interstate war." *Journal of Peace Research* 36(1): 75–94.

Cutright, Philip. 1963. "National political development: measurement and analysis." *American Sociological Review* 45: 253–264.

Dahl, Robert A. 1971. *Polyarchy. Participation and opposition.* New Haven: Yale University Press.

Davies, James C. 1962. "Toward a theory of revolution." *American Sociological Review* 27(1): 5–19.

DeAngelis, D.L. and K.A. Rose. 1992. Which individual-based approach is most appropriate for a given problem? In *Individual-based models and approaches in ecology. Populations, communities and ecosystems*, ed. Donald L. DeAngelis and Louis J. Gross. New York: Chapman & Hall pp. 67–87.

Doorenspleet, Renske. 2000. "Reassessing the three waves of democratization." *World Politics* 52(3): 384–406.

Doorenspleet, Renske. 2001. *The fourth wave of democratization: identification and explanation.* Leiden: Leiden University.

Doorenspleet, Renske. 2004. "The structural context of recent transitions to democracy." *European Journal of Political Research* 43: 309–335.

Edwards, K. and E.E. Smith. 1996. "A disconfirmation bias in the evaluation of arguments." *Journal of Personality and Social Psychology* 71(1): 5–24.

Elkink, Jos. 2003. "Democratic and economic diffusion." Paper presented at the 31st Joint Sessions of Workshops of the European Consortium of Political Research, Edinburgh (March 28-April 3, 2003).

Epstein, Joshua and Robert Axtell. 1996. *Growing artificial societies: social science from the bottom up.* Washington: Brookings Institution Press.

Ethier, Diane. 2003. "Is democracy promotion effective? Comparing conditionality and incentives." *Democratization* 10(1): 99–120.

Faust, Katherine, Barbara Entwisle, Ronald R. Rindfuss, Stephen J. Walsh and Yothin Sawangdee. 1999. "Spatial arrangement of social and economic networks among villages in Nang Rond District, Thailand." *Social Networks* 21: 311–337.

Finkel, Steven E. and Edward N. Muller. 1998. "Rational choice and the dynamics of collective political action: evaluating alternative models with panel data." *American Political Science Review* 92(1): 37–49.

Finkel, Steven E., Edward N. Muller and Karl-Dieter Opp. 1989. "Personal influence, collective rationality, and mass political action." *American Political Science Review* 83(3): 885–903.

Fordham, Benjamin O. and Victor Asal. 2007. "Billiard balls or snowflakes? Major power prestige and the international diffusion of institutions and practices." *International Studies Quarterly* 51: 31–52.

Franzese, Robert J. and Jude C. Hays. 2007*a*. "Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data." *Political Analysis* 15(2): 140–164.

Franzese, Robert J. and Jude C. Hays. 2007*b*. "The spatial probit model of interdependent binary outcomes: estimation, interpretation, and presentation." Presented at the 24th Annual Summer Meeting of the Society of Political Methodology.

Freeman, Linton C. 1979. "Centrality in social networks. I. Conceptual clarification." *Social Networks* 1: 215–239.

Fukuyama, Francis. 1989. "The end of history?" *The National Interest* .

Gardner, Martin. 1983. *Wheels, Life, and Other Mathematical Amusements.* W.H. Freeman.

Gibbon, Peter. 1993. The World Bank and the new politics of aid. In *Political conditionality*, ed. Georg Sørensen. London: Frank Cass pp. 35–62.

Gleditsch, Kristian S. 2002. *All international politics is local: the diffusion of conflict, integration, and democratization.* University of Michigan Press.

Gleditsch, Kristian S. and Michael D. Ward. 1997. "Double take: a reexamination of democracy and autocracy in modern polities." *Journal of Conflict Resolution* 41: 361–83.

Gleditsch, Kristian S. and Michael D. Ward. 2000. "War and peace in time and space: the role of democratization." *International Studies Quarterly* 36.

Gleditsch, Kristian S. and Michael D. Ward. 2006. "Diffusion and the international context of democratization." *International Organization* 60(4): 911–933.

Goldstone, Jack A. 2001. "Toward a fourth generation of revolutionary theory." *Annual Review of Political Science* 4: 139–187.

Granovetter, Mark. 1973. "The strength of weak ties." *American Journal of Sociology* 78: 1360–1380.

Granovetter, Mark. 1978. "Threshold models of collective behavior." *American Journal of Sociology* 83: 1420–1443.

Green, Donald and Ian Shapiro. 1994. *Pathologies of rational choice theory : a critique of applications in political science.* New Haven: Yale University Press.

Griffith, Daniel A. and Sandra L. Arlinghaus. 1995. *Practical handbook of spatial statistics.* Boca Raton, FL: CRC Press.

Grugel, Jean. 2002. *Democratization: a critical introduction.* Palgrave.

Hage, Per and Frank Harary. 1995. "Eccentricity and centrality in networks." *Social Networks* 17: 57–63.

Hak, Bin Chua. 1993. On spillovers and convergence PhD thesis Department of Economics, Harvard University Cambridge, Massachusetts: .

Herbst, Susan. 1993. *Numbered voices. How opinion polling has shaped American politics*. University of Chicago Press.

Higley, John and Michael G. Burton. 1989. "The elite variable in democratic transitions and breakdowns." *American Sociological Review* 54: 17–32.

Holland, John H. 1995. *Hidden order: how adaptation builds complexity*. Reading, MA: Addison-Wesley.

Holland, John H. 1998. *Emergence: from chaos to order*. Reading, MA: Addison Wesley.

Huntington, Samuel. 1991. *The third wave: democratization in the late twentieth century*. Norman, Oklahoma: University of Oklahoma Press.

Inglehart, Ronald. 2000. "World Values Surveys and European Values Surveys, 1981-1984, 1990-1993, and 1995-1997." Computer file, ICPSR Version.

Jacoby, Wade. 2006. "Inspiration, coalition, and substitution. External influences on postcommunist transformations." *World Politics* 58: 623–651.

Jager, Wander and Frédéric Amblard. 2004. "Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change." *Computation & Mathematical Organization Theory* 10: 295–303.

Jaggers, K. and T.R. Gurr. 1995. "Transitions to democracy: tracking democracy's third wave with the Polity III data." *Journal of Peace Research* 32: 469–482.

Johnson, Steven. 2001. *Emergence: the connected lives of ants, brains, cities, and software*. New York: Scribner.

Johnston, Hank. 2001. "Talking the walk: speech acts and resistance in authoritarian regimes." Paper presented at the Workshop on Repression and Mobilization, University of Maryland, June 21-23, 2001.

Kalyvitis, Sarantis and Irene Vlachaki. 2006. "Democracy assistance and the democratization of recipients." Working Paper, Athens University of Economics and Business.

Kang, Soong Moon. 2007. "Equicentrality and network centralization: a micro-macro linkage." *Social Networks* 29(4): 585–601.

Karl, Terry Lynn. 1990. "Dilemmas of democratization in Latin America." *Comparative Politics* 23(1): 1–21.

Katz, Elihu. 1999. "Theorizing diffusion: Tarde and Sorokin revisited." *Annals of American Academy of Political and Social Science* 566: 144–155.

Kernighan, Brian W. and Dennis M. Ritchie. 1988. *The C programming language.* 2nd ed. Englewood Cliffs, NJ: Prentice Hall.

King, Gary. 1998. *Unifying political methodology. The likelihood theory of statistical inference.* University of Michigan Press.

Knack, Stephen. 2004. "Does foreign aid promote democracy?" *International Studies Quarterly* 48(1): 251–266.

Kopstein, Jeffrey S. and David A. Reilly. 2000. "Geographic diffusion and the transformation of the postcommunist world." *World Politics* 53: 1–37.

Koster, M., R. Lindelauf, I. Lindner and G. Owen. 2008. "Mass-mobilization with noisy conditional beliefs." *Mathematical Social Sciences* 55: 55–77.

Kuran, Timur. 1989. "Sparks and prairie fires: a theory of unanticipated political revolution." *Public Choice* 61: 41–74.

Kuran, Timur. 1990. "Private and public preferences." *Economics and Philosophy* 6: 1–26.

Kuran, Timur. 1991*a*. "The East European revolution of 1989: is it surprising that we were surprised?" *The American Economic Review* 81(2): 121–125.

Kuran, Timur. 1991*b*. "Now out of never: the element of surprise in the East European revolution of 1989." *World Politics* 44(1): 7–48.

Kuran, Timur. 1995. *Private truths, public lies. The social consequences of preference falsification.* Cambridge, Massachusetts: Harvard University Press.

Kurzman, Charles. 1998. "Waves of democratization." *Studies in Comparative International Development* 3: 42–64.

Laver, Michael. 2005. "Policy and the dynamics of political competition." *American Political Science Review* 99(2): 263–281.

Laver, Michael and Ernest Sergenti. 2007. "Do ideologically intransigent parties affect the policy positions of other parties? And ... rigorously characterizing output from computational models of party competition." Paper presented at the Annual Meeting of the American Political Science Association, Chicago, August 2007.

le Gallo, Julie. 2000. "Econometrie spatiale. 1. Autocorrélation spatiale." Working Paper no. 2000-05, Laboratoire d'Analyse et des Techniques EConomiques.

Leandro, Jose E., Hartwig Schafer and Gaspar Frontini. 1999. "Towards a more effective conditionality: an operational framework." *World Development* 27(2): 285–299.

Levitsky, Steven and Lucan A. Way. 2003. "Ties that bind? International linkage and competitive authoritarian regime change in the post-cold war era." Paper prepared for the Annual Meeting of the American Political Science Association, Philadelphia, PA, August 27-30, 2003.

Levitsky, Steven and Lucan A. Way. 2005. "International linkage and democratization." *Journal of Democracy* 16(3): 20–34.

Linz, Juan J. and Alfred Stepan. 1996. *Problems of democratic transition and consolidation: southern Europe, South America, and post-communist Europe.* Baltimore: Johns Hopkins University Press.

Lipset, Martin Seymour. 1959. "Some social requisites of democracy: economic development and political legitimacy." *American Political Science Review* 53: 69–105.

Lipset, Martin Seymour. 1960. *Political man: the social bases of politics.* Garden City, NY: Doubleday.

Lohmann, Susanne. 1994. "The dynamics of informational cascades: the Monday demonstrations in Leipzig, East Germany, 1989-91." *World Politics* 47(1): 42–101.

Lord, C.G., L. Ross and M.R. Lepper. 1979. "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence." *Journal of Personality and Social Psychology* 37(11): 2098–2109.

Lustick, Ian S., Dan Miodownik and Roy J. Eidelson. 2004. "Secessionism in multicultural states: does sharing power prevent or encourage it?" *American Political Science Review* 98(2): 209–229.

Machiavelli, Niccolò. 1531. *The discourses.* Penguin Classics. Edited with an introduction by Bernard Crick, using the translation of Leslie J. Walker, with revisions by Brian Richardson (1984).

Markoff, John. 1996. *Waves of democracy: social movements and political change.* Thousand Oaks, CA: Pine Forge Press.

Marsden, Peter V. 2002. "Egocentric and sociocentric measures of network centrality." *Social Networks* 24: 407–422.

Marshall, M.G. and K. Jaggers. 2002. "Polity IV project: political regime characteristics and transitions, 1800-2002.". *http://www.bsos.umd.edu/cidcm/polity/*

McFaul, Michael. 2002. "The fourth wave of democracy and dictatorship. Noncooperative transitions in the postcommunist world." *World Politics* 54: 212–244.

McMillen, Daniel P. 1992. "Probit with spatial autocorrelation." *Journal of Regional Science* 32(3): 335–348.

McMillen, Daniel P. 1995. Spatial effects in probit models: a Monte Carlo investigation. In *New directions in spatial econometrics*, ed. Luc Anselin and Raymond J.G.M. Florax. Berlin: Springer Verlag pp. 189–228.

Miller, A.G., J.W. McHoskey, C.M. Bane and T.G. Bowd. 1993. "The attitude polarization phenomenon: Role of response measure, attitude extremity, and behavioral consequences of reported attitude change." *Journal of Personality and Social Psychology* 64(4): 561–574.

Modelski, George and Gardner Perry. 1991. "Democratization in long perspective." *Technological Forecasting and Social Change* 39: 23–34.

Modelski, George and Gardner Perry. 2002. """Democratization in long perspective" revisited." *Technological Forecasting and Social Change* 69: 359–376.

Mok, Diana, Barry Wellman and Ranu Basu. 2007. "Did distance matter before the Internet? Interpersonal contact and support in the 1970s." *Social Networks* 29: 430–461.

Moore, Barrington. 1966. *Social origins of dictatorship and democracy: lord and peasant in the making of the modern world.* Boston: Beacon Press.

Moran, P.A.P. 1948. "Some theorems on time series: II. The significance of the serial correlation coefficient." *Biometrika* 35(3/4): 255–260.

Moran, P.A.P. 1950. "A test for the serial independence of residuals." *Biometrika* 37(1/2): 178–181.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and causal inference. Methods and principles for social research.* New York: Cambridge University Press.

Moss, Scott. 2001. "Game theory: limitations and an alternative." *Journal of Artificial Societies and Social Simulation* 4(2). *http://jasss.soc.surrey.ac.uk/4/2/2.html*

Muller, Edward N., Henry A. Dietz and Steven E. Finkel. 1991. "Discontent and the expected utility of rebellion: the case of Peru." *American Political Science Review* 85(4): 1261–1282.

Noelle-Neumann, Elisabeth. 1993. *The spiral of silence.* 2nd ed. Chicago: University of Chicago Press.

North, M.J., N.T. Collier and J.R. Vos. 2006. "Experiences creating three implementations of the Repast agent modeling toolkit." *ACM Transactions on Modeling and Computer Simulation* 16(1): 1–25.

O'Loughlin, J., M.D. Ward, C.L. Lofdahl, J.S. Cohen, D.S. Brown, D. Reilly, K.S. Gleditsch and M. Shin. 1998. "The diffusion of democracy, 1946-1994." *Annals of the Association of American Geographers* 88: 545–574.

Opp, Karl-Dieter and Christiane Gern. 1993. "Dissident groups, personal networks, and spontaneous cooperation: the East German revolution of 1989." *American Sociological Review* 58(5): 659–680.

Oversloot, Hans, Joop van Holsteyn and Gert van der Berg. 2002. "Against all: exploring the vote 'against all' in the Russian Federation's electoral system." *Journal of Communist Studies and Transition Politics* 18(4): 31–50.

Page, Lawrence and Sergey Brin. 1998. The anatomy of a large-scale hypertextual web search engine. In *Proceedings of the Seventh International Web Conference (WWW 98).*

Page, Lawrence, Sergey Brin, R. Motwani and T. Winograd. 1998. Pagerank citation ranking: bringing order to the web. Technical report Stanford University.

Papke, Leslie E. and Jeffrey M. Wooldridge. 1996. "Econometric models for fractional response variables with an application to 401 (K) plan participation rates." *Journal of Applied Econometrics* 11(6): 619–632.

Paulos, John Allen. 1995. *A mathematician reads the newspaper*. New York: Anchor Books.

Paxton, Pamela. 2000. "Women's suffrage in the measurement of democracy: problems of operationalization." *Studies in Comparative International Development* 35(3): 92–111.

Paxton, Pamela. 2002. "Social capital and democracy: an interdependent relationship." *American Sociological Review* 67(2): 254–277.

Perloff, Richard M. 2003. *The dynamics of persuasion. Communication and attitudes in the 21st century*. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates.

Petty, Richard E. and John T. Cacioppo. 1996. *Attitudes and persuasion: classic and contemporary approaches*. Boulder: Westview Press.

Pierce, Walter M. 1940. "Climbing on the bandwagon." *Public Opinion Quarterly* 4(2): 241–243.

Pinkse, Joris and Margaret E. Slade. 1998. "Contracting in space: an application of spatial statistics to discrete-choice models." *Journal of Econometrics* 85: 125–154.

Plümper, Thomas and Christian Martin. 2006. "Multi-party competition: a computational model with abstention and biased voters.". *http://ssrn.com/abstract=882894*

Plümper, Thomas and Vera E. Troeger. 2007. "Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects." *Political Analysis* 15: 124–139.

Poirier, Dale J. and Paul A. Ruud. 1988. "Probit with dependent observations." *The Review of Economic Studies* 55(4): 593–614.

Popper, Karl. 1962. *The logic scientific discovery*. London: Hutchinson.

Price, Vincent. 1992. *Public opinion*. Sage Publications.

Przeworski, Adam, Michael Alvarez, José Antonio Cheibub and Fernando Limongi. 2000. *Democracy and development: political institutions and well-being in the world, 1950-1990.* Cambridge: Cambridge University Press.

Putnam, Robert D., Robert Leonardi and Raffaella Y. Nanetti. 1993. *Making democracy work: civic traditions in modern Italy.* Princeton: Princeton University Press.

R Development Core Team. 2008. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
http://www.R-project.org

Ramirez, Francisco O., Yasemin Soysal and Suzanne Shanahan. 1997. "The changing logic of political citizenship: cross-national acquisition of women's suffrage rights, 1890 to 1990." *American Sociological Review* 62(5): 735–745.

Reynolds, Paul. 2003. "The 'democratic domino' theory." *BBC News World Edition* . Downloaded 08-09-2005.
http://news.bbc.co.uk/2/hi/middle_east/2935969.stm

Rogers, Everett M. 1995. *Diffusion of innovations.* 4th ed. New York: The Free Press.

Rohrschneider, Robert. 1996. "Institutional learning versus value diffusion: the evolution of democratic values among parliamentarians in Eastern and Western Germany." *Journal of Politics* 58(2): 422–446.

Rosenau, J.N. 1988. "Patterned chaos in global life: structure and process in the two worlds of world politics." *International Political Science Review* 9: 327–364.

Ross, Michael. 2004. "Does taxation lead to representation?" *British Journal of Political Science* 34: 229–249.

Schelling, Thomas. 1978. *Micromotives and macrobehavior.* New York: W.W. Norton.

Schumpeter, Joseph A. 1976. *Capitalism, socialism and democracy.* New York: Harper & Row.

Shadwick, Bradley S. 2007. "The role of computation in physics." *Computing in Science & Engineering* 9(1): 101–103.

Shapiro, Michael J., G. Matthew Bonham and Daniel Heradstveit. 1988. "A discursive practices approach to collective decision-making." *International Studies Quarterly* 32(4): 397–419.

Sherif, M. and C.I. Hovland. 1961. *Social judgment.* New Haven, CT: Yale University Press.

Sherif, Muzafer. 1967. Introduction. In *Attitude, ego-involvement, and change*, ed. C.W. Sherif and M. Sherif. New York: Wiley pp. 1–5.

Simmons, Beth A. and Zachary Elkins. 2004. "The globalization of liberalization: policy diffusion in the international political economy." *American Political Science Review* 98(1): 171–189.

Starr, Harvey. 1991. "Democratic dominoes: diffusion approaches to the spread of democracy in the international system." *Journal of Conflict Resolution* 35(2): 356–381.

Staten, Clark. 1996. "The Islamic domino theory." EmergencyNet News Service.
http://www.emergency.com/domino.htm

Stinnett, Douglas M., Jaroslav Tir, Philip Schafer, Paul F. Diehl and Charles Gochman. 2002. "The correlates of war project direct contiguity data, version 3." *Conflict Management and Peace Science* 19(2): 58–66.

Strang, David. 1991. "Adding social structure to diffusion models. An event history framework." *Sociological Methods and Research* 19(3): 324–353.

Stroustrup, Bjarne. 1991. *The C++ programming language.* 2nd ed. Addison-Wesley.

Swinton, Scott M. 2002. "Capturing household-level spatial influence in agricultural management using random effects regression." *Agricultural Economics* 27: 371–381.

Tallberg, Christian. 2004. "Testing centralization in random graphs." *Social Networks* 26: 205–219.

The President's News Conference of April 7, 1954. 1954.
http://coursesa.matrix.msu.edu/ hst306/documents/domino.html

Tiefelsdorf, Michael. 2000. *Modelling spatial processes.* Vol. 87 of *Lecture notes in earth sciences* Berlin: Springer.

Tobler, W.R. 1970. "A computer movie simulating urban growth in the Detroit region." *Economic Geography* 46: 234–240.

Tolstoy, Leo. 1993. *War and peace.* Wordsworth. Translated by Louise & Aylmer Maude (orig. 1920).

Treier, Shawn and Simon Jackman. 2008. "Democracy as a latent variable." *American Journal of Political Science* 52(1): 201–217.

Vanhanen, Tatu. 1997. *Prospects for democracy: a study of 172 countries.* New York: Routledge.

Vatikiotis, Michael R.J. 1998. *Indonesian politics under Suharto: the rise and fall of the New Order.* Routledge.

Volden, Craig. 2007. "Failures: diffusion, learning, and policy abandonment." Paper presented at the Annual Meeting of the American Political Science Association, Chicago, August 2007.

Ward, Michael D., John O'Loughlin, Michael Shin, Corey L. Lofdahl, Kristian S. Gleditsch and Jordin S. Cohen. 1996. "The spatial and temporal diffusion of democracy, 1946-1994." Presented at the 37th Annual Conference of the International Studies Association, San Diego, CA.

Ward, Michael D. and Kristian S. Gleditsch. 1998. "Democratizing for peace." *American Political Science Review* 92: 51–61.

Watts, Duncan J. 1999. *Small worlds: the dynamics of networks between order and randomness.* Princeton: Princeton University Press.

Watts, Duncan J. 2003. *Six degrees: the science of a connected age.* Norton.

Watts, Duncan J. and Steven H. Strogatz. 1998. "Collective dynamics of 'small-world' networks." *Nature* 393: 440–442.

Wejnert, Barbara. 2005. "Diffusion, development, and democracy, 1800-1999." *American Sociological Review* 70: 53–81.

Welsh, William A. 1984. "Inter-nation interaction and political diffusion: notes toward a conceptual framework." Paper presented at the annual meeting of the International Studies Association, Atlanta, March 28-31.

Welzel, Christian. 2006. "Democratization as an empancipative process: the neglected role of mass motivations." *European Journal of Political Research* 45: 871–896.

Whitehead, Laurence. 1996. Three international dimensions of democratization. In *The international dimensions of democratization. Europe and the Americas*, ed. Laurence Whitehead. Oxford studies in democratization Oxford: Oxford University Press pp. 3–26.

Wood, Elisabeth Jean. 2000. *Forging democracy from below: insurgent transitions in South Africa and El Salvador*. Cambridge: Cambridge University Press.

Zakaria, Fareed. 1997. "The rise of illiberal democracy." *Foreign Affairs* 76(6): 22–43.

Zhu, Jian-Hua and Stanley Rosen. 1993. "From discontent to protest: individual-level causes of the 1989 pro-democracy movement in China." *International Journal of Public Opinion Research* (5): 234–249.

APPENDIX A

Convergence

This R code implements a simple function that provides a statistic for the level of convergence. The input is a measure of the amount of change over time, the result the mean and variance of the last fifth of this variable.

$$\bar{x} = \frac{5}{n} \sum_{i=\frac{4}{5}n}^{n} x_i \tag{A.1}$$

$$\sigma_x^2 = \frac{5}{n - \frac{1}{5}} \sum_{i=\frac{4}{5}n}^{n} (x_i - \bar{x})^2 \tag{A.2}$$

```
convergence <- function(x)
  {
    x <- x[floor(length(x) * .8):length(x)]

    x.bar <- mean(x)
    x.var <- var(x)

    list(mean = x.bar, variance = x.var)
  }
```

APPENDIX B

Code for the empirical data analysis

B.1 Generating spatial lags

The following Stata do-file[1] was used to generate the spatial lags on the democracy score:

```
clear
set mem 1g
set more off
cd "~/Desktop/academic/data/"

* List the variables here - note that there is listwise deletion of
* missings across those!
local vars = "v_comp v_part v_id Dv_comp Dv_part Dv_id p_polity2
                p_polity2_bin Dp_polity2 Dp_polity2_bin"

* Read contiguity dataset
insheet using "Direct Contiguity Master File.txt", clear

* Duplicate dataset
rename statelno tmpno
rename statelab tmpab
rename statehno statelno
```

[1] In some cases, lines have been split over multiple lines, for printing purposes. The original code was run without those line breaks.

```
rename statehab statelab
rename tmpno statehno
rename tmpab statehab

save tmp, replace
insheet using "Direct Contiguity Master File.txt", clear
append using tmp

drop note version
drop if conttype > 1

* Expand dataset
gen nyears = endyear - begyear + 1
expand nyears

bys statelno statehno conttype: gen year = _n + begyear - 1

* Merge democracy indicators
gen luniqid = statelno * 10000 + year
gen huniqid = statehno * 10000 + year

rename luniqid uniqid
sort uniqid
merge uniqid using vanhanen_polity_bollen, keep(`vars') uniqusing
rename uniqid luniqid
rename _merge m_lower

foreach v of varlist `vars' {
    rename `v' l`v'
}

rename huniqid uniqid
sort uniqid
merge uniqid using vanhanen_polity_bollen, keep(`vars') uniqusing
rename uniqid huniqid
rename _merge m_higher

foreach v of varlist `vars' {
    rename `v' h`v'
}
```

```
* Listwise deletion if any variable is missing
foreach v of local vars {
    drop if l`v' == .
    drop if h`v' == .
}

* Generate spatial lags
sort statelno year
foreach v of local vars {
    by statelno year: egen S`v' = mean(h`v')
}

* Save with all dyads preserved
save spatial_lags_dyads, replace

* Collapse to country-year format
collapse S*, by(statelno year)

* Merge into democracy dataset
gen uniqid = statelno * 10000 + year
sort uniqid
save spatial_lags, replace

use vanhanen_polity_bollen, clear
merge uniqid using spatial_lags, keep(S*) uniqmaster

* Generate world-wide spatial lags
sort year
foreach v of local vars {
    bys year: egen avg_world_`v' = mean(`v')
    bys year: egen n_world_`v' = count(`v')
    gen W`v' = (avg_world_`v' * n_world_`v' - `v') / (n_world_`v' - 1)
}

* Drop all other variables
keep country year uniqid S* W* avg_world_* n_world_*

* Save data set (spatial lags)
sort uniqid
save spatial_lags, replace
```

```
* Save annual data only
collapse avg_* n_*, by(year)
save annual_data, replace

set more on
```

B.2 Calculating Moran's I

The following Stata do-file was used to calculate Moran's I:

```
* Calculating Moran's I based on COW contiguity dataset, with Vanhanen
* and Polity IV scores
*
* Jos Elkink, Dublin, September 2005

* 30/09/2005  Original version
* 10/10/2005  Added calculation of Wjis and standard error
* 23/06/2008  Switched to using spatial_lags_dyads file

* W = weight matrix
* Wij = cell (i,j) from the weight matrix
* Wijs = standardized cell (i,j) from the weight matrix

set more off

* Define variables over which to calculate Moran's I
local vars = "v_comp v_part v_id Dv_comp Dv_part Dv_id p_polity2
                          p_polity2_bin Dp_polity2 Dp_polity2_bin"

* Open data file
use "~\Desktop\academic\data\spatial_lags_dyads.dta", clear
sort year
merge year using "~\Desktop\academic\data\annual_data.dta", uniqusing
rename _merge m_annual
sort year statelno

* use only cases where conttype == 1, thus only direct land borders

drop if conttype > 1
```

```
* unstandardized, Wij is now the same as conttype (since it's always 1)
* standardizing Wij :
gen Wij = conttype
by year statelno: egen sum_Wij = sum(Wij)
gen Wijs = Wij / sum_Wij
drop sum_Wij

* calculating standardized Wjis (needed for calculation of standard error)
sort year statehno
by year statehno: egen sum_Wij = sum(Wij)
gen Wjis = Wij / sum_Wij
drop sum_Wij
sort year statelno

label variable Wij "Unstandardized Wij"
label variable Wijs "Standardized Wij"
label variable Wjis "Standardized Wji"


* notation:  Ei = sum over i; Ej = sum over j; Eij = sum over i and over j;
*            _x = average of x; xi = x for case i; etc.

* Moran's I = sum_wdxij / sum_w * N / sum_dx2
* Whereby:   dx2 = (xi - _x) * (xi - _x)
*            dxij = (xi - _x) * (xj - _x)
*            wdxij = Wijs * dxij
*            sum_wdxij = Eij(wdxij)
*            sum_w = Eij(Wijs)
*            sum_dx2 = Ei(dx2)
* Formula used: Le Gallo (2000), p. 9

by year: egen sum_w = sum(Wijs)
label variable sum_w "Total of Wijs"

gen temp = (Wijs + Wjis) * (Wijs + Wjis)
by year: egen sum_ws = sum(temp)
drop temp
label variable sum_ws "Total of (Wijs + Wjis)^2 (S2 in formula for s.e.)"

foreach indicator of local vars {
```

```
  gen dx2_`indicator' = (l`indicator' - avg_world_`indicator')
                            * (l`indicator' - avg_world_`indicator')
  gen dxij_`indicator' = (l`indicator' - avg_world_`indicator')
                            * (h`indicator' - avg_world_`indicator')
  gen wdxij_`indicator' = Wijs * dxij_`indicator'
  by year: egen sum_wdxij_`indicator' = sum(wdxij_`indicator')
  by year: egen sum_dx2_`indicator' = sum(dx2_`indicator')
  gen moran_`indicator' = sum_wdxij_`indicator' / sum_w
                           * n_world_`indicator' / sum_dx2_`indicator'

  label variable dx2_`indicator' "Deviation from average score
                                        (`indicator') squared (state 1)"
  label variable dxij_`indicator' "Covariance state 1 and state 2 (`indicator')"
  label variable wdxij_`indicator' "Covariance state 1 and state 2, weighted for
                                     standardized contiguity matrix (`indicator')"
  label variable sum_wdxij_`indicator' "Total of wdxij (`indicator')"
  label variable sum_dx2_`indicator' "Total of dx2 (`indicator')"
  label variable moran_`indicator' "Moran's I (`indicator')"

  gen se_`indicator' = (2 * n_world_`indicator' * n_world_`indicator'
                        * sum_w - n_world_`indicator' * sum_ws + 12 * sum_w * sum_
                        / (4 * sum_w * sum_w * (n_world_`indicator'
                                                   * n_world_`indicator' - 1))
  label variable se_`indicator' "Standard error of Moran's I (`indicator')"
}

* Save datafile in Stata format
save "~\Desktop\academic\data\calculate moran - output.dta", replace

* Smoothen the data over time
collapse n_world* moran_* se_*, by(year)

tsset year

foreach v of local vars {

  gen exp_moran_`v' = -1 / (n_world_`v' - 1)
  gen moran_`v'_l = moran_`v' - se_`v'
  gen moran_`v'_h = moran_`v' + se_`v'
```

```
  tssmooth ma moran_`v'_s = moran_`v', window(3 1 3)
  tssmooth ma moran_`v'_ls = moran_`v'_l, window(3 1 3)
  tssmooth ma moran_`v'_hs = moran_`v'_h, window(3 1 3)
}

save "~\Desktop\academic\data\calculate moran - smooth output.dta", replace

set more on
```

B.3 Empirical analyses and plots

The empirical regression analysis was performed using the following R code:

```
library(foreign)
library(MASS)
library(arm)

setwd("~/Desktop/academic/data/")

## Open Vanhanen data
vh <- read.dta("vanhanen.dta", convert.underscore=TRUE)
vhc <- read.dta("vanhanen_collapsed.dta", convert.underscore=TRUE)

## Open collapsed democracy data
collapsed <- read.dta("vanhanen_polity_bollen_collapsed.dta", convert.undersc
annual <- read.dta("annual_data.dta", convert.underscore=TRUE)

## Open data on correlations of democracy data
correlations <- read.dta("correlations_indicators.dta", convert.underscore=TR

## Open data on Moran's I
moran <- read.dta("calculate moran - smooth output.dta", convert.underscore=T

## Open spatial lags and Polity IV data and merge
polity.spatial <- read.dta("spatial_lags.dta", convert.underscore=TRUE)
polity <- read.dta("vanhanen_polity_bollen.dta", convert.underscore=TRUE)
gleditsch <- read.table("Gleditsch Ward IO 2006/repdata.asc", header=TRUE)
gleditsch$uniqid <- gleditsch$numid * 10000 + gleditsch$year
polity <- merge(polity, polity.spatial, by="uniqid", all=TRUE)
polity <- merge(polity, gleditsch, by="uniqid", all=TRUE)
```

```
setwd("~/Desktop/academic/main/")

## Helper function for standardization
stdze <- function(x) {

  v <- x[!is.na(x)]

  (v - mean(v)) / sqrt(var(v))
}

## Generate density plot of Vanhanen democracy scores
x <- vh$part[vh$part > 0 & vh$comp > 0]
y <- vh$comp[vh$part > 0 & vh$comp > 0]

d <- kde2d(x,y,n=250)

postscript("diss_vanhanen_participation_vs_competition.eps")
filled.contour(d, nlevels=40, xlab="Participation",
               ylab="Competition", col=gray(100:50/100), bty="n")
dev.off()

## Generate plot to compare democracy indicators
year.comp <- collapsed$year[!is.na(collapsed$v.comp)]
year.part <- collapsed$year[!is.na(collapsed$v.part)]
year.polity <- collapsed$year[!is.na(collapsed$p.polity2)]
year.bollen <- collapsed$year[!is.na(collapsed$b.DEM)]
year.fh.pr <- collapsed$year[!is.na(collapsed$f.pr)]
year.fh.cl <- collapsed$year[!is.na(collapsed$f.cl)]

## 1972 is the year that the Freedom House data starts

comp.pre1972 <- stdze(collapsed$v.comp[year.comp < 1972])
part.pre1972 <- stdze(collapsed$v.part[year.part < 1972])
polity2.pre1972 <- stdze(collapsed$p.polity2[year.polity < 1972])
comp.post1972 <- stdze(collapsed$v.comp[year.comp >= 1972])
part.post1972 <- stdze(collapsed$v.part[year.part >= 1972])
polity2.post1972 <- stdze(collapsed$p.polity2[year.polity >= 1972])

year.comp.pre1972 <- year.comp[year.comp < 1972]
year.part.pre1972 <- year.part[year.part < 1972]
```

```
year.polity.pre1972 <- year.polity[year.polity < 1972]
year.comp.post1972 <- year.comp[year.comp >= 1972]
year.part.post1972 <- year.part[year.part >= 1972]
year.polity.post1972 <- year.polity[year.polity >= 1972]

comp <- stdze(collapsed$v.comp)
part <- stdze(collapsed$v.part)
polity2 <- stdze(collapsed$p.polity2)
bollen <- stdze(collapsed$b.DEM)
fh.pr <- stdze(collapsed$f.pr.rev)
fh.cl <- stdze(collapsed$f.cl.rev)

postscript("diss_compare_measures.eps")

par(mfrow=c(2,1))

plot(comp.pre1972~year.comp.pre1972, type="l",
     ylim=c(-2,2.5), xlim=c(1800,1972), xlab="Year",
     ylab="Standardized score", bty="n")
lines(part.pre1972~year.part.pre1972, col="grey")
lines(polity2.pre1972~year.polity.pre1972, lty="dashed")
points(bollen~year.bollen, pch=19, cex=1, col="grey")
legend("topleft", legend=c("Vanhanen competition", "Vanhanen participation",
                    "Polity IV score", "Bollen's Political Democracy Index"),
       col=c("black","grey","black","grey"),
       lty=c("solid","solid","dashed","blank"),
       lwd=c(1,1,1,0), pch=c(NA,NA,NA,19), bty="n")


plot(comp.post1972~year.comp.post1972, type="l",
     ylim=c(-2,2.5), xlim=c(1972,2005), xlab="Year",
     ylab="Standardized score", bty="n")
lines(part.post1972~year.part.post1972, col="grey")
lines(polity2.post1972~year.polity.post1972, lty="dashed")
points(bollen~year.bollen, pch=19, cex=1, col="grey")
lines(fh.pr~year.fh.pr, col="black", lwd=2, lty="longdash")
lines(fh.cl~year.fh.cl, col="grey", lwd=2, lty="longdash")
legend("bottomright", legend=c("Freedom House - political rights",
                    "Freedom House - civil liberties"),
       col=c("black","grey"),
       lty=c("longdash","longdash"),
```

```
       lwd=c(2,2), pch=c(NA,NA), bty="n")


par(mfrow=c(1,1))

dev.off()

## Generate plot of correlations of democracy indicators
postscript("diss_corr_measures.eps")

plot(correlations$ma.c.part.DEM ~ correlations$year, type="l", xlab="Year",
     ylab="Correlation", ylim=c(0.1,1), lwd=2, bty="n")
lines(correlations$ma.c.comp.DEM ~ correlations$year, col="grey", lwd=2)
lines(correlations$ma.c.part.polity ~ correlations$year, lty="dotted", lwd=2)
lines(correlations$ma.c.comp.polity ~ correlations$year, lty="dotted",
      col="grey", lwd=2)
lines(correlations$ma.c.part.arat ~ correlations$year, lty="dashed", lwd=2)
lines(correlations$ma.c.comp.arat ~ correlations$year, lty="dashed",
      col="grey", lwd=2)
lines(correlations$ma.c.part.fhpr ~ correlations$year, lty="longdash", lwd=2)
lines(correlations$ma.c.comp.fhpr ~ correlations$year, lty="longdash",
      col="grey", lwd=2)
lines(correlations$ma.c.part.fhcl ~ correlations$year, lty="longdash", lwd=1)
lines(correlations$ma.c.comp.fhcl ~ correlations$year, lty="longdash",
      col="grey", lwd=1)
legend("bottomleft", col=c("black","grey","black","grey","black","grey",
                     "black","grey","black","grey"),
       lty=c("solid","solid","dotted","dotted","dashed","dashed","longdash",
         "longdash","longdash","longdash"),
       legend=c("Participation - Bollen","Competition - Bollen",
         "Participation - Polity","Competition - Polity","Participation - Arat'
         "Competition - Arat","Participation - FH Political Rights",
         "Competition - FH Political Rights",
         "Participation - FH Civil Liberties",
         "Competition - FH Civil Liberties"), bty="n",
       lwd=c(2,2,2,2,2,2,2,2,1,1))

dev.off()

## Generate plot of Vanhanen indicators as time path
postscript("diss_vanhanen_participation_vs_competition_path.eps")
```

```
plot(comp~part, data=vhc, type="l", xlab="Participation", ylab="Competition",
     xlim=c(0,36), bty="n")

for (i in 1:length(vhc$year)) {

  if (vhc$year[i] %% 20 == 0) {

    points(vhc$part[i],vhc$comp[i],cex=1,pch=19)
    text(vhc$part[i],vhc$comp[i],vhc$year[i],pos=4)
  }
}
dev.off()

## Plot Moran's I over time for competition and participation
plot.moran <- function(moran, var, name, legend.loc) {

  m <- moran[,sprintf("moran.%s.s", var)]
  ml <- moran[,sprintf("moran.%s.ls", var)]
  mh <- moran[,sprintf("moran.%s.hs", var)]
  mexp <- moran[,sprintf("exp.moran.%s", var)]

  year <- moran$year

  plot(m~year, ylab=sprintf("Moran's I (%s)", name),
       xlab="Year", cex=0, type="l", ylim=c(-0.4,1), bty="n")
  lines(mh~year, col="grey")
  lines(ml~year, col="grey")
  lines(m~year)
  lines(mexp~year, lty="dashed")
  legend(legend.loc, legend=c("Moran's I","Confidence interval",
                              "Expected value"),
         lty=c("solid", "solid", "dashed"), col=c("black","grey","black"),
         bty="n")
}

postscript("diss_moran_time.eps")

par(mfcol=c(2,2))

plot.moran(moran, "v.comp", "Vanhanen's competition", "topleft")
plot.moran(moran, "p.polity2.bin", "Polity IV, binary", "topright")
```

```
plot.moran(moran, "v.part", "Vanhanen's participation", "topright")
plot.moran(moran, "Dp.polity2.bin", "Polity IV, binary, 1st difference",
           "topright")

dev.off()

## Regression analyses
lag <- function(x, panel, time) {

  lag.within <- function(x, time) {

    xn <- rep(NA, length(x))

    for (i in 1:length(x))
      if (sum(time == (time[i]-1)) == 1)
        xn[i] = x[time == (time[i]-1)]

    xn
  }

  xn <- rep(NA, length(x))
  for (i in na.omit(unique(panel)))
    xn[which(panel == i)] <- lag.within(x[which(panel == i)],
               time[which(panel == i)])

  xn
}

polity$ipcents <- polity$ipyears / 100

for (v in c("v.comp", "Wv.comp", "Sv.comp", "SDv.comp",
            "v.part", "Wv.part", "Sv.part", "SDv.part",
            "p.polity2.bin", "Wp.polity2.bin",
            "Sp.polity2.bin", "SDp.polity2.bin",
            "energy2", "cwar", "ipcents"))
  polity[,sprintf("L%s", v)] <- lag(polity[,v], polity$v.ssno, polity$v.year)

polity$Lp.polity2.bin.inverse <- 1 - polity$Lp.polity2.bin

summary(lmer(v.comp ~ Lv.comp + LWv.comp + LSv.comp + LSDv.comp
            + (1 | v.ssno), data=polity))
```

```
summary(lmer(v.part ~ Lv.part + LWv.part + LSv.part + LSDv.part
            + (1 | v.ssno), data=polity))
summary(lmer(p.polity2.bin ~ 0 + Lp.polity2.bin.inverse + Lp.polity2.bin
            + Lp.polity2.bin.inverse:LWp.polity2.bin
            + Lp.polity2.bin.inverse:LSp.polity2.bin
            + Lp.polity2.bin.inverse:LSDp.polity2.bin
            + Lp.polity2.bin:LWp.polity2.bin
            + Lp.polity2.bin:LSp.polity2.bin
            + Lp.polity2.bin:LSDp.polity2.bin
            + (1 | p.ccode),
            family=binomial(link="logit"), data=polity))

summary(lmer(v.comp ~ Lv.comp + LWv.comp + LSv.comp + LSDv.comp
            + log(Lenergy2) + Lcwar + Lipcents
            + (1 | v.ssno), data=polity))
summary(lmer(v.part ~ Lv.part + LWv.part + LSv.part + LSDv.part
            + log(Lenergy2) + Lcwar + Lipcents
            + (1 | v.ssno), data=polity))
summary(lmer(p.polity2.bin ~ 0 + Lp.polity2.bin.inverse + Lp.polity2.bin
            + Lp.polity2.bin.inverse:LWp.polity2.bin
            + Lp.polity2.bin.inverse:LSp.polity2.bin
            + Lp.polity2.bin.inverse:LSDp.polity2.bin
            + Lp.polity2.bin.inverse:log(Lenergy2)
            + Lp.polity2.bin.inverse:Lcwar
            + Lp.polity2.bin.inverse:Lipcents
            + Lp.polity2.bin:LWp.polity2.bin
            + Lp.polity2.bin:LSp.polity2.bin
            + Lp.polity2.bin:LSDp.polity2.bin
            + Lp.polity2.bin:log(Lenergy2)
            + Lp.polity2.bin:Lcwar
            + Lp.polity2.bin:Lipcents
            + (1 | p.ccode),
            family=binomial(link="logit"), data=polity))
```

APPENDIX C

Replication code for Jager and Amblard (2004)

This code replicates the agent-based model presented in Jager and Amblard (2004), which demonstrates how different threshold values in the social judgment theory (see §3.1.2) can lead to uniformity, bipolarization, and pluriformity in attitudes, without the necessity of "opinion leaders". The code is written for the statistical package R (R Development Core Team 2008). The function below implements the configurable simulation.

```
att.simulation <- function(name, nagents = 1000, niterations = 350000, delta
                           use.normal = FALSE, init = 10, init.std = 20,
                           u.mean = 20, u.std = 0, t.mean = 30, t.std = 0,
                           save.plot = NULL, save.col=TRUE)
  {
    # Parameters
    nlevels <- 100

    u <- rnorm(nagents, u.mean, u.std)
    t <- rnorm(nagents, t.mean, t.std)

    t[t < u] <- u[t < u]

    # Initialization
    if (use.normal)
      {
        attitudes <- rnorm(nagents, init, init.std)
```

```
    attitudes[attitudes < 0] <- 0
    attitudes[attitudes > nlevels] <- nlevels
  }
else
  attitudes <- runif(nagents, 0, 100)

# Storage to enable more fancy graphs
if (diff(range(u)) > 0)
  u.std <- (u - min(u)) / diff(range(u)) * 100
else
  u.std <- u

if (diff(range(t)) > 0)
  t.std <- (t - min(t)) / diff(range(t)) * 100
else
  t.std <- t

# Store initial state
init.state <- attitudes

# Set plot parameters when saving the plots
if (!is.null(save.plot))
  {
    postscript(sprintf("~/Desktop/att1.%s.eps", name),
               height=200, width=200*(length(save.plot)+1))
    par(mfrow=c(1,length(save.plot)+1))
    if (save.col)
      cols <- terrain.colors(100)[u.std]
    else
      cols <- "black"
  }

# Storage for convergence information
change = 0
change.track <- NULL

# Iterations
for (i in 1:niterations)
  {
    # Sample two communication agents and implement SJT model
    actors <- sample(1:nagents, 2, replace=FALSE)
```

```
    a1 <- attitudes[actors[1]]
    a2 <- attitudes[actors[2]]
    diff <- abs(a2 - a1)

    if (diff <= u[actors[1]])
      attitudes[actors[1]] <- max(min(nlevels,
                                      a1 + sign(a2-a1) * delta), 0)
    else if (diff >= t[actors[1]])
      attitudes[actors[1]] <- max(min(nlevels,
                                      a1 - sign(a2-a1) * delta), 0)

    # Store the amount of change in attitudes over 1000 iterations
    change <- change + abs(attitudes[actors[1]] - a1)

    if (i %% 1000 == 0)
      {
        change.track <- rbind(change.track, c(i, change))
        change <- 0
      }

    # If not saving the plots, show a plot every 1000 iteration
    if (i %% 1000 == 0 & is.null(save.plot))
      {
        plot(attitudes ~ init.state, bty="n",
             col=terrain.colors(100)[u.std], ylim=c(0,100), xlim=c(0,100)
        text(20,90,sprintf("Correlation: %9.4f  Iteration: %d",
                           cor(attitudes, init.state), i))
      }
    # If saving the plots, save for each requested iteration
    # plus the final state
    else if (i %in% save.plot | i == niterations)
      {
        plot(attitudes ~ init.state, bty="n",
             col=cols, ylim=c(0,100), xlim=c(0,100),
             xlab="", ylab="")
        text(0,90,sprintf("Iteration:\n%d", i), cex=2, pos=4)
      }
  }

if (!is.null(save.plot))
```

```
    {
      dev.off()
      par(mfrow=c(1,1))
    }

  colnames(change.track) <- c("iteration","change")

  list(init.state = init.state, final.state = attitudes,
       change.track = as.data.frame(change.track))
}
```

The replication of Jager and Amblard (2004) is implemented as follows (see §4.3.1):

```
# Replication of the parameter settings of Jager & Amblard (2004)
as1 <- att.simulation("rep1", u.mean=20, t.mean=30,
                      save.plot=c(1,2,3)*30000, save.col=FALSE)
as2 <- att.simulation("rep2", u.mean=60, t.mean=80,
                      save.plot=c(1,2,3)*30000, save.col=FALSE)
as3 <- att.simulation("rep3", u.mean=30, t.mean=60,
                      save.plot=c(1,2,3)*30000, save.col=FALSE)
as4 <- att.simulation("rep4", u.mean=10, t.mean=80,
                      save.plot=c(1,2,3)*40000, save.col=FALSE)
```

The convergence of the changes in attitudes are validated using the code from Appendix A.

```
# Check convergence
convergence(as1$change.track[,2])
convergence(as2$change.track[,2])
convergence(as3$change.track[,2])
convergence(as4$change.track[,2])

postscript("~/Desktop/rep.conv.eps")
plot(as1$change.track[,2] ~ as1$change.track[,1],
     type="l", lwd=2, xlab="Iteration",
     ylab="Change in attitudes per 1000 iterations", ylim=c(0,1000))
```

```
lines(as2$change.track[,2] ~ as2$change.track[,1],
      lty="dotted", lwd=2)
lines(as3$change.track[,2] ~ as3$change.track[,1],
      lty="dashed", lwd=2)
lines(as4$change.track[,2] ~ as4$change.track[,1],
      lty="longdash", lwd=2)
legend("topright", lwd=2, lty=c("solid","dotted","dashed","longdash"),
       legend=c("u=20,t=30","u=60,t=80","u=30,t=60","u=10,t=80"))
dev.off()
```

The equivalent two steps, the simulation and convergence measures, are calculated for the replication with normally distributed initial values for alpha ($\alpha_0 \sim N_{[0,100]}(10, 20)$):

```
# Replication of the parameter settings of Jager & Amblard (2004), with
# normally distributed initial attitudes
as1a <- att.simulation("rep1na", u.mean=20, t.mean=30, use.normal=TRUE,
                       save.plot=c(1,2,3)*30000, save.col=FALSE)
as2a <- att.simulation("rep2na", u.mean=60, t.mean=80, use.normal=TRUE,
                       save.plot=c(1,2,3)*30000, save.col=FALSE)
as3a <- att.simulation("rep3na", u.mean=30, t.mean=60, use.normal=TRUE,
                       save.plot=c(1,2,3)*30000, save.col=FALSE)
as4a <- att.simulation("rep4na", u.mean=10, t.mean=80, use.normal=TRUE,
                       save.plot=c(1,2,3)*40000, save.col=FALSE)

# Check convergence
convergence(as1a$change.track[,2])
convergence(as2a$change.track[,2])
convergence(as3a$change.track[,2])
convergence(as4a$change.track[,2])

postscript("~/Desktop/repna.conv.eps")
plot(as1a$change.track[,2] ~ as1a$change.track[,1],
     type="l", lwd=2, xlab="Iteration",
     ylab="Change in attitudes per 1000 iterations", ylim=c(0,1000))
lines(as2a$change.track[,2] ~ as2a$change.track[,1],
      lty="dotted", lwd=2)
lines(as3a$change.track[,2] ~ as3a$change.track[,1],
      lty="dashed", lwd=2)
```

```
lines(as4a$change.track[,2] ~ as4a$change.track[,1],
      lty="longdash", lwd=2)
legend("topright", lwd=2, lty=c("solid","dotted","dashed","longdash"),
       legend=c("u=20,t=30","u=60,t=80","u=30,t=60","u=10,t=80"))
dev.off()
```

The code used for the replication of Jager and Amblard (2004) with normally distributed threshold values is as follows:

```
# Replication of the parameter settings of Jager & Amblard (2004), with
# normally distributed threshold values
u.std <- t.std <- 20

as1b <- att.simulation("rep1nt", u.mean=20, u.std=u.std, t.mean=30, t.std=t.std,
                       save.plot=c(1,2,3)*30000, save.col=FALSE)
as2b <- att.simulation("rep2nt", u.mean=60, u.std=u.std, t.mean=80, t.std=t.std,
                       save.plot=c(1,2,3)*30000, save.col=FALSE)
as3b <- att.simulation("rep3nt", u.mean=30, u.std=u.std, t.mean=60, t.std=t.std,
                       save.plot=c(1,2,3)*30000, save.col=FALSE)
as4b <- att.simulation("rep4nt", u.mean=10, u.std=u.std, t.mean=80, t.std=t.std,
                       save.plot=c(1,2,3)*40000, save.col=FALSE)

# Check convergence
convergence(as1b$change.track[,2])
convergence(as2b$change.track[,2])
convergence(as3b$change.track[,2])
convergence(as4b$change.track[,2])

postscript("~/Desktop/repnt.conv.eps")
plot(as1b$change.track[,2] ~ as1b$change.track[,1],
     type="l", lwd=2, xlab="Iteration",
     ylab="Change in attitudes per 1000 iterations", ylim=c(0,1000))
lines(as2b$change.track[,2] ~ as2b$change.track[,1],
      lty="dotted", lwd=2)
lines(as3b$change.track[,2] ~ as3b$change.track[,1],
      lty="dashed", lwd=2)
lines(as4b$change.track[,2] ~ as4b$change.track[,1],
      lty="longdash", lwd=2)
legend("topright", lwd=2, lty=c("solid","dotted","dashed","longdash"),
```

```
       legend=c("u=20,t=30","u=60,t=80","u=30,t=60","u=10,t=80"))
dev.off()
```

APPENDIX D

Convergence statistics of replication of Jager and Amblard (2004)

u	t	Mean of $\Delta\alpha$	Variance of $\Delta\alpha$
20	30	95	124
60	80	1000	0
30	60	17	14
10	80	149	113

Table D.1: Convergence in replication of Jager and Amblard (2004). The maximum value of $\bar{\Delta\alpha}$ is 1000.

u	t	Mean of $\Delta\alpha$	Variance of $\Delta\alpha$
20	30	0	0
60	80	964	31
30	60	924	67
10	80	403	254

Table D.2: Convergence with normally distributed initial values in Jager and Amblard (2004) with $\alpha_0 \sim N_{[0,100]}(10, 20)$. The maximum value of $\bar{\Delta\alpha}$ is 1000.

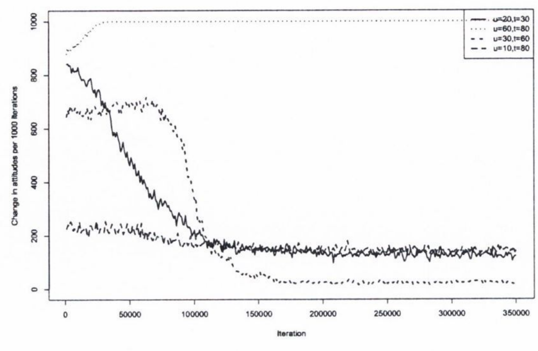

Figure D.1: Convergence in replication of Jager and Amblard (2004).

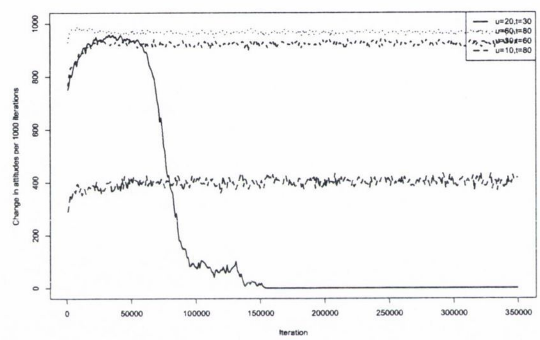

Figure D.2: Convergence with normally distributed initial values in Jager and Amblard (2004) with $\alpha_0 \sim N_{[0,100]}(10, 20)$.

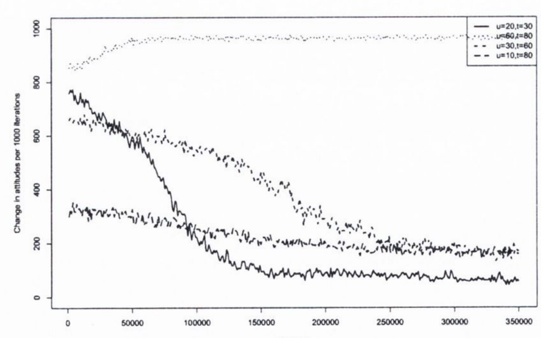

Figure D.3: Convergence with normally distributed thresholds in Jager and Amblard (2004).

$\bar{u}$	$\bar{t}$	Mean of $\Delta\alpha$	Variance of $\Delta\alpha$
20	30	64	107
60	80	963	35
30	60	167	277
10	80	168	138

Table D.3: Convergence with normally distributed thresholds in Jager and Amblard (2004) Jager and Amblard (2004). The maximum value of $\bar{\Delta\alpha}$ is 1000.

APPENDIX E

Regressions of Monte Carlo simulation results

	Model 1		Model 2		Model 3	
Constant	0.625	*	0.657	*	0.560	*
	(0.083)		(0.081)		(0.026)	
Broadcast effect (B)	0.027		0.030		0.041	*
	(0.030)		(0.030)		(0.005)	
Chance of cross-border communication (τ)	-0.187		-0.197		0.003	
	(0.136)		(0.136)		(0.022)	
Communication effect ($delta$)	-0.371	*	-0.376	*	-0.304	*
	(0.068)		(0.067)		(0.009)	
Random chance of coups (K)	-157.533		-210.546		421.202	*
	(727.357)		(727.041)		(110.033)	
$I(B > 0)$	0.042				0.046	
	(0.025)				(0.025)	
$B \times \tau$	0.050		0.053			
	(0.052)		(0.052)			
$B \times \delta$	0.006		0.007			
	(0.024)		(0.024)			
$\tau \times \delta$	0.176		0.185			
	(0.116)		(0.116)			
$B \times K$	179.107		194.464			
	(258.417)		(258.385)			
$\tau \times K$	1188.970		1255.463			
	(1187.912)		(1187.845)			
$\delta \times K$	898.368		946.300			
	(598.335)		(597.961)			
$B \times \tau \times \delta$	-0.021		-0.023			
	(0.042)		(0.042)			
$B \times \tau \times K$	-300.021		-321.729			
	(435.161)		(435.185)			
$B \times \delta \times K$	-211.422		-222.443			
	(208.940)		(208.941)			
$\tau \times \delta \times K$	-1577.266		-1649.126			
	(992.281)		(991.862)			
$B \times \tau \times \delta \times K$	249.151		267.795			
	(356.490)		(356.495)			
R^2	0.39		0.39		0.39	

Table E.1: Regression results for Monte Carlo simulation analysis with as dependent variable the global level of democracy in the last 10% of iterations. Standard errors in parentheses. Based on 2000 simulation runs. $I(x)$ equals 1 if condition x is true, 0 otherwise.

	Model 1	Model 2	Model 3
Constant	0.043 *	0.042 *	0.010
	(0.020)	(0.019)	(0.006)
Broadcast effect (B)	0.011	0.011	0.015 *
	(0.007)	(0.007)	(0.001)
Chance of cross-border communication (τ)	-0.018	-0.017	0.006
	(0.033)	(0.033)	(0.005)
Communication effect ($delta$)	-0.040 *	-0.040 *	-0.001
	(0.016)	(0.016)	(0.002)
Random chance of coups (K)	-367.067 *	-364.808 *	-79.275 *
	(175.031)	(174.830)	(26.671)
$I(B > 0)$	-0.002		-0.001
	(0.006)		(0.006)
$B \times \tau$	0.007	0.007	
	(0.012)	(0.012)	
$B \times \delta$	0.009	0.009	
	(0.006)	(0.006)	
$\tau \times \delta$	0.028	0.028	
	(0.028)	(0.028)	
$B \times K$	23.098	22.444	
	(62.186)	(62.133)	
$\tau \times K$	358.271	355.438	
	(285.860)	(285.638)	
$\delta \times K$	278.962	276.920	
	(143.984)	(143.790)	
$B \times \tau \times \delta$	-0.014	-0.013	
	(0.010)	(0.010)	
$B \times \tau \times K$	-112.634	-111.709	
	(104.717)	(104.648)	
$B \times \delta \times K$	-41.346	-40.877	
	(50.279)	(50.243)	
$\tau \times \delta \times K$	-286.313	-283.251	
	(238.783)	(238.510)	
$B \times \tau \times \delta \times K$	131.515	130.720	
	(85.786)	(85.725)	
R^2	0.12	0.12	0.10

Table E.2: Regression results for Monte Carlo simulation analysis with as dependent variable the average deviation from the expected value in spatial clustering, Moran's I. Standard errors in parentheses. Based on 2000 simulation runs. $I(x)$ equals 1 if condition x is true, 0 otherwise.

APPENDIX F

Code for network analysis on countries

For the analysis of the network of countries as a (social) network, most of the code is written in C++ (Stroustrup 1991). This code is imported a dynamic library in R, after which the simulations were set up.

F.1 Conquering algorithm source code

The C++ code consists of a number of source files, each of which will be listed below.

main.cpp

```
#include <iostream>
#include <exception>

extern "C" {
#include <R.h>
#include <Rinternals.h>
#include <R_ext/Rdynload.h>
}

#include "model.h"
#include "exception.h"
```

```
using std::cout;
using std::endl;
using std::exception;

extern "C" {

void runSimulation(int *p_nFieldWidth,
           int *p_nFieldHeight,
           int *p_nBorderMultiplier,
           double *out_ConnMatrix,
           int *out_TimeTaken,
           int *p_nCountries,
           int *p_nVerbose)
{
  time_t tt = time(NULL);

  try
    {
      CModel Model(*p_nFieldWidth, *p_nFieldHeight, *p_nBorderMultiplier,
           *p_nVerbose);

      Model.RRun(out_ConnMatrix, p_nCountries);
    }
  catch (CException& e)
    {
      cout << "Uncaught exception: " << e.GetMessage() << endl;
    }
  catch (exception& e)
    {
      Rprintf("Uncaught exception: %s\n", e.what());
    }

  *out_TimeTaken = (int) (tt - time(NULL));
}


  static R_NativePrimitiveArgType runSimulation_t[] = {INTSXP, INTSXP,
              INTSXP, REALSXP, REALSXP, INTSXP, INTSXP, INTSXP};

// Stuff for dynamic loading in R
R_CMethodDef cMethods[] = {
```

```
  {"runSimulation", (DL_FUNC) &runSimulation, 7, runSimulation_t},
  {NULL, NULL, 0}
};

void R_init_mylib(DllInfo *info)
{
  R_registerRoutines(info, cMethods, NULL, NULL, NULL);
}

void R_unload_mylib(DllInfo *info)
{
}

} // extern "C"
```

model.h

```
#ifndef _MODEL_H
#define _MODEL_H

#include <vector>
#include "country.h"
#include "province.h"
#include "normal.h"

#define NORTH 0
#define WEST 1
#define SOUTH 2
#define EAST 3

using std::vector;

class CReporter;

class CModel
{
 public:
  CModel(int p_nFieldWidth, int p_nFieldHeight, int p_nBorderMultiplier,
     int p_nVerbose);
  virtual ~CModel();
```

```
  void RRun(double *out_ConnMatrix, int *n_Countries);

  CProvince& GetProvince(int x, int y);
  CProvince& GetNeighbour(int x, int y, int k);

 private:
  void Setup();
  void ClearDataStorage();
  void CreateProvinces();
  void CreateCountries();
  void CreateConnectionMatrix();
  void SaveConnectionMatrix();

  // Run parameters
  bool m_bVerbose;

  // Model parameters
  int m_nFieldWidth;
  int m_nFieldHeight;
  int m_nBorderMultiplier;

  // Reporters
  double* m_R_out_ConnMatrix;

  // Internal data storage
  vector<CCountry*> m_vCountries;
  vector<CProvince*> m_vProvinces;
  double* m_pdConnectionMatrix;
};

#endif
```

model.cpp

```
#include "model.h"

#include <cstdlib>
#include <ctime>
#include <iostream>
#include <iomanip>
#include <sstream>
```

```
#include <fstream>
#include <exception>
#include "exception.h"
#include "uniform.h"

using std::cout;
using std::endl;
using std::ostringstream;
using std::fstream;
using std::ios;
using std::fixed;
using std::setprecision;
using std::exception;

extern "C" {
#include <R.h>
}

CModel::CModel(int p_nFieldWidth, int p_nFieldHeight,
           int p_nBorderMultiplier, int p_nVerbose)
  : m_nFieldWidth(p_nFieldWidth),
    m_nFieldHeight(p_nFieldHeight),
    m_nBorderMultiplier(p_nBorderMultiplier)
{
  //srand((unsigned) time(0));

  m_bVerbose = p_nVerbose == 1;

  if (m_bVerbose)
    Rprintf("Starting with parameters:\n  Width %d\n  Height %d\n
              Multiplier %d\n  Verbose %d\n", m_nFieldWidth, m_nFieldHeight,
        m_nBorderMultiplier, p_nVerbose);
}

CModel::~CModel()
{
}

void CModel::RRun(double *out_ConnMatrix, int *n_Countries)
{
  if (m_bVerbose)
```

```
    Rprintf("Running with addresses:\n  Connection matrix %x\n
             Number of countries %x\n", out_ConnMatrix, n_Countries);

  if (m_bVerbose) Rprintf("Setup...\n");

  if (out_ConnMatrix == NULL)
    Rprintf("Null pointer passed to RRun() (%x)!\n", out_ConnMatrix);

  m_R_out_ConnMatrix = out_ConnMatrix;

  Setup();

  if (m_bVerbose) Rprintf("Storing connection matrix...\n");
  SaveConnectionMatrix();

  if (m_bVerbose) Rprintf("Saving number of countries... (%d)\n",
              m_vCountries.size());
  *n_Countries = m_vCountries.size();
}

void CModel::Setup()
{
  int i;

  ClearDataStorage();

  if (m_bVerbose) Rprintf("Creating provinces...\n");
  CreateProvinces();
  if (m_bVerbose) Rprintf("Creating countries...\n");
  CreateCountries();
  if (m_bVerbose) Rprintf("Creating connection matrix...\n");
  CreateConnectionMatrix();
}

void CModel::ClearDataStorage()
{
  int i;

  for (i = 0; i < m_vCountries.size(); ++i)
    delete m_vCountries[i];
```

```
  for (i = 0; i < m_vProvinces.size(); ++i)
    delete m_vProvinces[i];

  m_vCountries.clear();
  m_vProvinces.clear();
}

void CModel::CreateProvinces()
{
  int x,y,i = 0;
  CProvince *pNewProvince;

  for (x = 0; x < m_nFieldWidth; ++x)
    for (y = 0; y < m_nFieldHeight; ++y)
      {
    pNewProvince = new CProvince(this, x, y, ++i);
    pNewProvince->CreateSovereign();
    m_vProvinces.push_back(pNewProvince);
      }
}

void CModel::CreateCountries()
{
  int i, x, y, k;

  for (i = 0; i < m_nFieldWidth * m_nFieldHeight * m_nBorderMultiplier; ++i)
    {
      x = rand() % m_nFieldWidth;
      y = rand() % m_nFieldHeight;
      k = rand() % 4;

      GetNeighbour(x,y,k).GetCountry().Conquer(GetProvince(x,y));
    }

  for (i = 0; i < m_vProvinces.size(); ++i)
    if (m_vProvinces[i]->IsCapital())
    m_vCountries.push_back(&m_vProvinces[i]->GetCountry());

  for (i = 0; i < m_vCountries.size(); ++i)
    m_vCountries[i]->SetID(i);
}
```

```
void CModel::CreateConnectionMatrix()
{
  int i,j,k,id1,id2,s;
  int nCountries = m_vCountries.size();

  // Create empty connection matrix
  m_pdConnectionMatrix = new double[nCountries * nCountries];
  for (i = 0; i < nCountries; ++i)
    for (j = 0; j < nCountries; ++j)
      m_pdConnectionMatrix[i * nCountries + j] = 0.0;

  // Set all borders to one
  for (i = 0; i < m_vProvinces.size(); ++i)
    {
      id1 = m_vProvinces[i]->GetCountry().GetID();

      id2 = m_vProvinces[i]->GetNeighbour(NORTH).GetCountry().GetID();
      if (id1 != id2)
    {
      m_pdConnectionMatrix[id1 * nCountries + id2] = 1;
      m_pdConnectionMatrix[id2 * nCountries + id1] = 1;
    }

      id2 = m_vProvinces[i]->GetNeighbour(WEST).GetCountry().GetID();
      if (id1 != id2)
    {
      m_pdConnectionMatrix[id1 * nCountries + id2] = 1;
      m_pdConnectionMatrix[id2 * nCountries + id1] = 1;
    }
    }
}

void CModel::SaveConnectionMatrix()
{
  int i,j;
  int nCountries = m_vCountries.size();

  for (i = 0; i < nCountries * nCountries; ++i)
    m_R_out_ConnMatrix[i] = m_pdConnectionMatrix[i];
}
```

```
CProvince& CModel::GetProvince(int x, int y)
{
  if (x > m_nFieldWidth - 1 || x < 0
      || y > m_nFieldHeight - 1 || y < 0)
    {
      ostringstream strMessage;
      strMessage << "Attempt to get non-existing province at ("
         << x << "," << y << ")";
      throw CException(strMessage.str());
    }

  return *m_vProvinces.at(x * m_nFieldHeight + y);
}

CProvince& CModel::GetNeighbour(int x, int y, int k)
{
  if (k > 3 || k < 0
      || x > m_nFieldWidth - 1 || x < 0
      || y > m_nFieldHeight - 1 || y < 0)
    {
      ostringstream strMessage;
      strMessage << "Neighbour requested of impossible province - x,y,k: "
         << x << "," << y << "," << k;
      throw CException(strMessage.str());
    }

  switch (k)
    {
    case NORTH:
      if (y == 0) y = m_nFieldHeight - 1;
      else --y;
      break;

    case SOUTH:
      if (y == m_nFieldHeight - 1) y = 0;
      else ++y;
      break;

    case WEST:
      if (x == 0) x = m_nFieldWidth - 1;
```

```
      else --x;
      break;

    case EAST:
      if (x == m_nFieldWidth - 1) x = 0;
      else ++x;
      break;
    }

  return GetProvince(x,y);
}
```

country.h

```
#ifndef _COUNTRY_H
#define _COUNTRY_H

#include <vector>

using std::vector;

class CProvince;
class CModel;

class CCountry
{
 public:
  CCountry(CModel* p_pModel, CProvince* p_pProvince);
  virtual ~CCountry();

  void Conquer(CProvince& p_Province);
  bool OwnsProvince(CProvince& p_Province);
  bool IsAtom() { return m_vProvinces.size() == 1; }
  void RemoveProvince(CProvince* p_pProvince);
  CProvince& GetCapital();
  void SetCapital(CProvince& p_Capital);
  int GetSize() { return m_vProvinces.size(); }

  int GetID() { return m_nID; }
  vector<CProvince*>& GetProvinces() { return m_vProvinces; }
```

```
  void SetID(int p_nID) { m_nID = p_nID; }

 private:
  bool CheckConnectednessWithoutProvince(CProvince& p_RemovedProvince);
  void CheckCWPRecursive(CProvince& p_CurrentProvince,
           CProvince& p_RemovedProvince,
           vector<int>& p_vCheckedProvinces);

  int m_nID;
  vector<CProvince*> m_vProvinces;
  CProvince* m_pCapital;
  CModel* m_pModel;
};

#endif
```

country.cpp

```
#include "country.h"

#include <algorithm>
#include <iostream>
#include <iomanip>
#include <cmath>
#include "province.h"
#include "model.h"
#include "uniform.h"

#define MAX(a,b) ((a)>(b)?(a):(b))
#define MIN(a,b) ((a)>(b)?(b):(a))

using std::find;
using std::cout;
using std::endl;
using std::left;
using std::right;
using std::setw;

CCountry::CCountry(CModel* p_pModel, CProvince* p_pProvince)
  : m_pModel(p_pModel)
{
```

```
  // Countries are created by the province which becomes the capital
  m_vProvinces.push_back(p_pProvince);
  m_pCapital = p_pProvince;
  m_nID = p_pProvince->GetID();
}

CCountry::~CCountry()
{
}

CProvince& CCountry::GetCapital()
{
  return *m_pCapital;
}

void CCountry::SetCapital(CProvince& p_Capital)
{
  m_pCapital = &p_Capital;
}

void CCountry::Conquer(CProvince& p_Province)
{
  // If the province to be conquered is not already part of this
  // country and the country of the province would still be
  // connected after conquering then set the country of the
  // province to the this one and add the province to this
  // country's provinces
  if (p_Province.GetCountry().GetID() != m_nID
      && p_Province.GetCountry().CheckConnectednessWithoutProvince(p_Province))
    {
      p_Province.SetCountry(this);
      m_vProvinces.push_back(&p_Province);
    }
}

void CCountry::RemoveProvince(CProvince* p_pProvince)
{
  vector<CProvince*>::iterator iter;

  // Find this province and remove it
  iter = find(m_vProvinces.begin(), m_vProvinces.end(), p_pProvince);
```

```
  if (iter != m_vProvinces.end())
    m_vProvinces.erase(iter);

  // If the province to be removed is the capital,
  // and there are other provinces left, set
  // a new, random capital
  if (m_pCapital == p_pProvince && m_vProvinces.size())
    m_pCapital = m_vProvinces[rand() % m_vProvinces.size()];
}

bool CCountry::CheckConnectednessWithoutProvince(CProvince& p_RemovedProvince
{
  vector<int> vCheckedProvinces;

  // If this is the only province in the country, no further checking needed
  if (IsAtom())
    return true;

  // Recursively find out whether provinces remaining are still connected
  CheckCWPRecursive(*m_pCapital, p_RemovedProvince, vCheckedProvinces);

  return vCheckedProvinces.size() == m_vProvinces.size() - 1;
}

void CCountry::CheckCWPRecursive(CProvince& p_CurrentProvince,
                 CProvince& p_RemovedProvince,
                 vector<int>& p_vCheckedProvinces)
{
  int k, nCurrentID;

  nCurrentID = p_CurrentProvince.GetID();

  // If the current province is a compatriot of the capital
  // and it is not the province to be removed
  // and it is not among the provinces already checked
  // then add to provinces checked
  // and check all four neighbouring provinces
  if (m_pCapital->IsCompatriot(p_CurrentProvince)
      && nCurrentID != p_RemovedProvince.GetID()
      && find(p_vCheckedProvinces.begin(),
          p_vCheckedProvinces.end(), nCurrentID)
```

```
      == p_vCheckedProvinces.end())
    {
      p_vCheckedProvinces.push_back(nCurrentID);

      // Check all four neighbouring provinces,
      // unless all provinces of this country have
      // already been checked
      for (k = 0; k < 4 &&
          p_vCheckedProvinces.size() < m_vProvinces.size() - 1; ++k)
    CheckCWPRecursive(p_CurrentProvince.GetNeighbour(k),
              p_RemovedProvince, p_vCheckedProvinces);
    }
}
```

province.h

```
#ifndef _PROVINCE_H
#define _PROVINCE_H

#include <vector>

using std::vector;

class CModel;
class CCountry;

class CProvince
{
 public:
  CProvince(CModel* p_pModel, int p_nX, int p_nY, int p_nID);
  virtual ~CProvince();

  void CreateSovereign();
  bool IsCompatriot(CProvince& p_Province);
  bool IsCapital();
  CProvince& GetNeighbour(int k);

  CCountry& GetCountry() { return *m_pCountry; }
  int GetID() { return m_nID; }
  int GetX() { return m_nX; }
  int GetY() { return m_nY; }
```

```
  void SetCountry(CCountry* p_pCountry);

 private:
  int m_nX, m_nY;
  int m_nID;
  CCountry* m_pCountry;
  CModel* m_pModel;
};

#endif
```

province.cpp

```
#include "province.h"

#include "country.h"
#include "uniform.h"
#include "model.h"

#define MAX(a,b) ((a)>(b)?(a):(b))

CProvince::CProvince(CModel* p_pModel, int p_nX, int p_nY, int p_nID)
  : m_nX(p_nX),
    m_nY(p_nY),
    m_nID(p_nID),
    m_pModel(p_pModel)
{
}

CProvince::~CProvince()
{
}

void CProvince::CreateSovereign()
{
  m_pCountry = new CCountry(m_pModel, this);
}

bool CProvince::IsCompatriot(CProvince& p_Province)
{
```

```
  return m_pCountry->GetID() == p_Province.GetCountry().GetID();
}

bool CProvince::IsCapital()
{
  return m_pCountry->GetCapital().GetID() == m_nID;
}

CProvince& CProvince::GetNeighbour(int k)
{
  return m_pModel->GetNeighbour(m_nX, m_nY, k);
}

void CProvince::SetCountry(CCountry* p_pCountry)
{
  m_pCountry->RemoveProvince(this);
  m_pCountry = p_pCountry;
}
```

uniform.h

```
#ifndef _UNIFORM_H
#define _UNIFORM_H

class CUniform
{
 public:
  static double GetDouble();
  static bool GetBoolean(double p_dProbability);
  static int GetNextIntFromTo(int p_nFrom, int p_nTo);
};

#endif
```

uniform.cpp

```
#include "uniform.h"

#include <cstdlib>

double CUniform::GetDouble()
```

```
{
  return (double) rand() / (double) RAND_MAX;
}

bool CUniform::GetBoolean(double p_dProbability)
{
  return CUniform::GetDouble() < p_dProbability;
}

int CUniform::GetNextIntFromTo(int p_nFrom, int p_nTo)
{
  return rand() % (p_nTo - p_nFrom + 1) + p_nFrom;
}
```

normal.h

```
#ifndef _NORMAL_H
#define _NORMAL_H

class CNormal
{
 public:
  CNormal(double p_dMean, double p_dStandardDeviation);
  virtual ~CNormal();

  int GetNextIntFromTo(int p_nFrom, int p_nTo);
  int GetNextIntFrom(int p_nFrom);
  double GetNextDouble();

  double GetMean() { return m_dMean; }
  double GetStandardDeviation() { return m_dStandardDeviation; }

 private:
  double m_dMean;
  double m_dStandardDeviation;
  bool m_bHasStoredValue;
  double m_dStoredValue;
};

#endif
```

normal.cpp

The below code makes use of Box and Muller (1958).

```
#include "normal.h"

#define MAX(a,b) ((a)>(b)?(a):(b))
#define MIN(a,b) ((a)>(b)?(b):(a))

#include <cmath>
#include "uniform.h"

CNormal::CNormal(double p_dMean, double p_dStandardDeviation)
  : m_dMean(p_dMean),
    m_dStandardDeviation(p_dStandardDeviation),
    m_bHasStoredValue(false)
{
}

CNormal::~CNormal()
{
}

int CNormal::GetNextIntFromTo(int p_nFrom, int p_nTo)
{
  int nValue = (int) floor(GetNextDouble());
  return MIN(MAX(p_nFrom, nValue), p_nTo);
}

int CNormal::GetNextIntFrom(int p_nFrom)
{
  int nValue = (int) floor(GetNextDouble());
  return MAX(p_nFrom, nValue);
}

double CNormal::GetNextDouble()
{
  if (m_bHasStoredValue)
    {
      m_bHasStoredValue = false;
      return m_dStoredValue;
    }
```

```
  else
    {
      // See G.E.P. Box and Mervin E. Muller,
      // "A note on the generation of random normal deviates", The Annals
      // of Mathematical Statistics, 1958 (29:2), pp. 610-611.

      double u1 = CUniform::GetDouble();
      double u2 = CUniform::GetDouble();

      m_dStoredValue = pow(-2 * log(u1), .5) * sin(6.28318530717959 * u2)
                             * m_dStandardDeviation + m_dMean;
      m_bHasStoredValue = true;

      return pow(-2 * log(u2), .5) * cos(6.28318530717959 * u1)
                                 * m_dStandardDeviation + m_dMean;
    }
}
```

exception.h

```
#ifndef _EXCEPTION_H
#define _EXCEPTION_H

#include <string>

using std::string;

class CException
{
 public:
  CException(string p_strMessage) { m_strMessage = p_strMessage; }
  virtual ~CException() {};

  string& GetMessage() { return m_strMessage; }

 private:
  string m_strMessage;
};

#endif
```

F.2 The simulation analyses

The main analysis is done in R. First, the simulation code is imported as a dynamic library.

```
dyn.load("country.so")

simulation <- function(width=20, height=20, multiplier=6, verbose=TRUE)
{
  ## The maximum number of countries is one for each province; the connection
  ## matrix is nCountries x nCountries, so the storage space should be
  ## (width * height) ^ 2.
  outConnMatrix <- as.double(rep(0, (width * height) ^ 2))
  outTimeTaken <- as.integer(0)
  nCountries <- as.integer(0)

  out <- .C("runSimulation",
            as.integer(width),
            as.integer(height),
            as.integer(multiplier),
            outConnMatrix = outConnMatrix,
            outTimeTaken = outTimeTaken,
            nCountries = nCountries,
            as.integer(verbose))

  if (verbose)
    cat(sprintf("Reached end of simulation after %d seconds\n",
                out$outTimeTaken));

  connMatrix <- t(matrix(out$outConnMatrix[1:(out$nCountries^2)],
                         nrow=out$nCountries, ncol=out$nCountries))

  list(nCountries=out$nCountries, W=connMatrix,
       time.taken=out$outTimeTaken)
}
```

Once the dynamic library has been loaded, a set of simulations has been run to provide the results discussed in §4.1.1. The first simulations concern the comparison between the network generating algorithm as implemented in the main model with the real world system of states.

```
source("network_analysis.R")
```

```
source("tallberg.R")

library(sna)
library(foreign)

## ===== Real world network statistics =====

system("say Starting with real world network")

## Read DCM data from Correlates of War project
dcm <- read.csv("~/Desktop/academic/data/Direct Contiguity Master File.txt")

state.ids <- rownames(table(c(dcm$StateLNo, dcm$StateHNo)))
world.data <- NULL
postscript("~/Desktop/academic/main/world_network.eps")
par(mfrow = c(1,2))
for (year in min(dcm$BegYear):max(dcm$EndYear)) {

  W <- matrix(0, length(state.ids), length(state.ids))
  rownames(W) <- state.ids
  colnames(W) <- state.ids

  for (state in state.ids) {

    edges <- dcm[dcm$StateLNo == state & dcm$BegYear <= year & dcm$EndYear >=
                 & dcm$ContType == 1, "StateHNo"]

    if (length(edges) > 0)
      W[state, as.character(edges)] <- W[as.character(edges), state] <- 1
  }

  # Drop countries that are not connected; in most cases these will be countr
  # that did not exist in that year, but also islands will be dropped
  sel <- rowSums(W) > 0
  W <- W[sel,sel]

  # Calculate network statistics
  world.data <- rbind(world.data, c(year, dim(W)[1],
                                    centralization(W, degree,
                                                   mode="graph"),
                                    centralization(W, betweenness,
```

```
                                              mode="graph"),
                                centralization(W, tallberg.closeness,
                                              mode="graph"),
                                mean(colSums(W)),
                                sd(colSums(W))))

  # Make some plots
  if (year == 1990) {

    us.index <- which(rownames(W) == "2")
    g <- gplot(W, usearrows=F, vertex.cex=1.5, vertex.sides=20,
               edge.lwd=2, xlab="1990")
    points(g[us.index,1], g[us.index,2], cex=1, col="black", pch=19)
  } else if (year == 2000) {

    us.index <- which(rownames(W) == "2")
    g <- gplot(W, usearrows=F, vertex.cex=1.5, vertex.sides=20,
               edge.lwd=2, xlab="2000")
    points(g[us.index,1], g[us.index,2], cex=1, col="black", pch=19)
  }
}
par(mfrow = c(1,1))
dev.off()
colnames(world.data) <- c("year","ncountries",
                          "degree.centrality","betweenness.centrality",
                          "closeness.centrality","average.degree","std.degree")
world.data <- as.data.frame(world.data)

## Create plot of last year we have data for
postscript(sprintf("~/Desktop/academic/main/world_network%s.eps", year))
us.index <- which(rownames(W) == "2")
g <- gplot(W, usearrows=F, vertex.cex=1.5, vertex.sides=20, edge.lwd=2)
text(g[us.index,1],g[us.index,2],"USA",cex=1,col="blue",pos=4)
dev.off()

## ===== The conquering algorithm network =====

system("say Starting with conquering network")

## Model parameters
sizes <- c(10,20,30)
```

```
multipliers <- c(3,6,10,20)
iterations <- 20

## Simulation iterations
conquer.data <- NULL
for (s in sizes) {
  for (m in multipliers) {
    for (i in 1:iterations) {

      # Create artificial country setup
      sim.output <- simulation(width=s, height=s, multiplier=m, verbose=FALSE

      # Calculate network statistics
      conquer.data <- rbind(conquer.data, c(i, s, m, sim.output$nCountries,
                                      centralization(sim.output$W,
                                                     degree,
                                                     mode="graph"),
                                      centralization(sim.output$W,
                                                     betweenness,
                                                     mode="graph"),
                                      centralization(sim.output$W,
                                                     tallberg.closeness
                                                     mode="graph"),
                                      mean(colSums(sim.output$W)),
                                      sd(colSums(sim.output$W))))

    }
  }
}
colnames(conquer.data) <- c("iteration","width","multiplier","ncountries",
                             "degree.centrality","betweenness.centrality",
                             "closeness.centrality","average.degree","std.degr
conquer.data <- as.data.frame(conquer.data)

## ===== Random networks =====

system("say Starting with Bernoulli networks")

## Model parameters
reconnections <- c(10,50,100,1000)
probabilities <- c(.01,.02,.05,.1)
```

```
par(mfrow=c(1,1))
dev.off()

system("say Finished with network analysis")
```

The closeness measure of centrality as suggested by Tallberg (2004) is implemented below. Note that only the features required for our analyses are included, not the full set of a proper centrality mearsure.

```
library(sna)

tallberg.closeness <- function(dat, g=1, diag=NULL, gmode=NULL, tmaxdev=FALSE)
{
  v <- dim(dat)[1]

  if (tmaxdev)
    {
      res <- 1/2 * v - 1
    }
  else
    {
      d <- geodist(dat)$gdist
      d.inv <- 1/d
      diag(d.inv) <- 0

      res <- 1/(v-1) * rowSums(d.inv)
    }

  res
}
```

APPENDIX G

Code for the agent-based model

For the main simulation, the primary code is written in C++ (Stroustrup 1991), with an interface in R (R Development Core Team 2008). The analysis was done using R setup code.

G.1 Main simulation code

The C++ code consists of a number of source files, each of which will be listed below. The code is compiled with:

```
R CMD SHLIB -o model.so *.cpp
```

main.cpp

```
#include <iostream>
#include <exception>

extern "C" {
#include <R.h>
#include <Rinternals.h>
#include <R_ext/Rdynload.h>
}

#include "model.h"
#include "exception.h"

using std::cout;
```

```
using std::endl;
using std::exception;

extern "C" {

  void runSimulation(int *p_nIterations,
             int *p_nThinning,
             int *p_nDetailThinning,
             int *p_nNumberOfLevels,
             double *p_dInitialProportionDemocratic,
             double *p_dRandomRevolutionChance,
             double *p_dCrossBorderChance,
             int *p_nRegimeDelay,
             int *p_nBroadcastEffect,
             int *p_nCommunicationEffect,
             int *p_nFieldWidth,
             int *p_nFieldHeight,
             int *p_nBorderMultiplier,
             int *p_bUseDecay,
             double *p_dDecayStart,
             double *p_dDecayStrength,
             int *p_nIsoMean,
             int *p_nIsoStd,
             int *p_nPopMean,
             int *p_nPopStd,
             int *p_nAttMean,
             int *p_nAttStd,
             int *p_nInGroupMean,
             int *p_nInGroupStd,
             int *p_nOutGroupMean,
             int *p_nOutGroupStd,
             int *p_nBatchID,
             int *p_nRunID,
             double *out_Moran,
             double *out_Proportion,
             double *out_ConnMatrix,
             double *out_AttMap,
             int *out_Population,
             int *out_Details,
             int *out_TimeTaken,
             int *p_nCountries,
```

```
            int *p_nVerbose)
{
  time_t tt = time(NULL);

  try
    {
  CModel Model(*p_nIterations,
            *p_nThinning, *p_nDetailThinning,
            *p_nNumberOfLevels,
            *p_dInitialProportionDemocratic,
            *p_dRandomRevolutionChance,
            *p_dCrossBorderChance,
            *p_nRegimeDelay,
            *p_nBroadcastEffect, *p_nCommunicationEffect,
            *p_nFieldWidth, *p_nFieldHeight,
            *p_nBorderMultiplier,
            *p_bUseDecay,
            *p_dDecayStart, *p_dDecayStrength,
            *p_nIsoMean, *p_nIsoStd,
            *p_nPopMean, *p_nPopStd,
            *p_nAttMean, *p_nAttStd,
            *p_nInGroupMean, *p_nInGroupStd,
            *p_nOutGroupMean, *p_nOutGroupStd,
            *p_nBatchID, *p_nRunID,
            *p_nVerbose);

  Model.Run(out_Moran, out_Proportion, out_ConnMatrix,
        out_AttMap, out_Population, out_Details,
        p_nCountries);
    }
  catch (CException& e)
    {
  cout << "Uncaught exception: " << e.GetMessage() << endl;
    }
  catch (exception& e)
    {
  Rprintf("Uncaught exception: %s\n", e.what());
    }

  *out_TimeTaken = (int) (tt - time(NULL));
}
```

```
  static R_NativePrimitiveArgType runSimulation_t[]
  = {INTSXP, INTSXP, INTSXP, INTSXP, REALSXP, REALSXP, REALSXP,
     INTSXP, INTSXP, INTSXP, INTSXP, INTSXP, INTSXP, INTSXP, REALSXP,
     REALSXP, INTSXP, INTSXP, INTSXP, INTSXP, INTSXP, INTSXP, INTSXP,
     INTSXP, INTSXP, INTSXP, INTSXP, INTSXP, REALSXP, REALSXP, REALSXP,
     REALSXP, INTSXP, INTSXP, INTSXP, INTSXP, INTSXP};

  // Stuff for dynamic loading in R
  R_CMethodDef cMethods[] = {
    {"runSimulation", (DL_FUNC) &runSimulation, 37, runSimulation_t},
    {NULL, NULL, 0}
  };

  void R_init_mylib(DllInfo *info)
  {
    R_registerRoutines(info, cMethods, NULL, NULL, NULL);
  }

  void R_unload_mylib(DllInfo *info)
  {
  }

} // extern "C"
```

model.h

```
#ifndef _MODEL_H
#define _MODEL_H

#include <vector>
#include "citizen.h"
#include "country.h"
#include "province.h"
#include "normal.h"

#define NORTH 0
#define WEST 1
#define SOUTH 2
#define EAST 3
```

```
#define ACTION_OPEN 0
#define ACTION_CLOSE 1
#define ACTION_WRITE 2

using std::vector;

class CModel
{
 public:
  CModel(int p_nIterations,
     int p_nThinning, int p_nDetailThinning,
     int p_nNumberOfLevels,
     double p_dInitialProportionDemocratic,
     double p_dRandomRevolutionChance,
     double p_dCrossBorderChance,
     int p_nRegimeDelay,
     int p_nBroadcastEffect, int p_nCommunicationEffect,
     int p_nFieldWidth, int p_nFieldHeight,
     int p_nBorderMultiplier,
     int p_bUseDecay,
     double p_dDecayStart, double p_dDecayStrength,
     int p_nIsoMean, int p_nIsoStd,
     int p_nPopMean, int p_nPopStd,
     int p_nAttMean, int p_nAttStd,
     int p_nInGroupMean, int p_nInGroupStd,
     int p_nOutGroupMean, int p_nOutGroupStd,
     int p_nBatchID, int p_nRunID,
     int p_nVerbose);
  virtual ~CModel();

  void Run(double *out_Moran, double *out_Proportion,
      double *out_ConnMatrix, double *out_AttMap,
      int *out_Population, int *out_Details,
      int *n_Countries);

  CProvince& GetProvince(int x, int y);
  CProvince& GetNeighbour(int x, int y, int k);
  void AddRevolution(bool p_bCoup);

  int GetCurrentRun() { return m_nRunID; } // deprecated
  int GetRunID() { return m_nRunID; }
```

```
  int GetCurrentIteration() { return m_nCurrentIteration; }
  int GetCurrentSession() { return m_nBatchID; } // deprecated
  int GetBatchID() { return m_nBatchID; }
  int GetNumberOfLevels() { return m_nNumberOfLevels; }
  int GetRegimeDelay() { return m_nRegimeDelay; }
  int GetBroadcastEffect() { return m_nBroadcastEffect; }
  int GetCommunicationEffect() { return m_nCommunicationEffect; }
  double GetInitialProportionDemocratic() {
    return m_dInitialProportionDemocratic; }
  double GetRandomRevolutionChance() { return m_dRandomRevolutionChance; }
  double GetCrossBorderChance() { return m_dCrossBorderChance; }
  CNormal& GetIsolationNormal() { return *m_pIsolationNormal; }
  bool UseDecay() { return m_bUseDecay; }
  double GetDecayStart() { return m_dDecayStart; }
  double GetDecayStrength() {return m_dDecayStrength; }

private:
  void Setup();
  void Step();
  void ShuffleCitizens();
  void ClearDataStorage();
  void CreateProvinces();
  void CreateCountries();
  void CountCitizens();
  void CreateCitizens();
  void CreateConnectionMatrix();
  void SaveConnectionMatrix();
  void PrintMap();
  double GetAverageAttitude();
  double GetProportionDemocratic();
  double GetProportionRevolutions();
  double GetMoranI();
  void PrintProgress();
  void WriteCountryState(int action);
  void ReportToR();

  // Run parameters
  int m_nRuns;
  int m_nIterations;
  int m_nBatchID;
  int m_nRunID;
```

```
int m_nCurrentIteration;
bool m_bVerbose;
int m_nThinning;
int m_nMapThinning;
int m_nDetailThinning;
int m_nDotProgress;
int m_nLineProgress;

FILE *m_fCountryFile;

// Model parameters
int m_nFieldWidth;
int m_nFieldHeight;
int m_nBorderMultiplier;
int m_nNumberOfLevels;
int m_nRegimeDelay;
int m_nBroadcastEffect;
int m_nCommunicationEffect;
double m_dInitialProportionDemocratic;
double m_dRandomRevolutionChance;
double m_dCrossBorderChance;
bool m_bUseDecay;
double m_dDecayStart;
double m_dDecayStrength;

// Model statistics
int m_nCoups;
int m_nRevolutions;
double m_nrgDemocracyMemory[80];

// R output
double* m_R_out_Moran;
double* m_R_out_Proportion;
double* m_R_out_ConnMatrix;
double* m_R_out_AttMap;
int* m_R_out_Population;
int* m_R_out_Details;

// Distribution
CNormal* m_pIsolationNormal;
CNormal* m_pPopulationNormal;
```

```
  CNormal* m_pInGroupNormal;
  CNormal* m_pOutGroupNormal;
  CNormal* m_pAttitudeNormal;

  // Internal data storage
  vector<CCitizen*> m_vCitizens;
  vector<CCountry*> m_vCountries;
  vector<CProvince*> m_vProvinces;
  double* m_pdConnectionMatrix;
};

#endif
```

model.cpp

```
#include "model.h"

#include <cstdlib>
#include <ctime>
#include <iostream>
#include <iomanip>
#include <sstream>
#include <fstream>
#include <exception>
#include "exception.h"
#include "uniform.h"

using std::cout;
using std::endl;
using std::ostringstream;
using std::fstream;
using std::ios;
using std::fixed;
using std::setprecision;
using std::exception;

extern "C" {
#include <R.h>
}

CModel::CModel(int p_nIterations,
```

```
            int p_nThinning, int p_nDetailThinning,
            int p_nNumberOfLevels,
            double p_dInitialProportionDemocratic,
            double p_dRandomRevolutionChance,
            double p_dCrossBorderChance,
            int p_nRegimeDelay,
            int p_nBroadcastEffect, int p_nCommunicationEffect,
            int p_nFieldWidth, int p_nFieldHeight,
            int p_nBorderMultiplier,
            int p_bUseDecay,
            double p_dDecayStart, double p_dDecayStrength,
            int p_nIsoMean, int p_nIsoStd,
            int p_nPopMean, int p_nPopStd,
            int p_nAttMean, int p_nAttStd,
            int p_nInGroupMean, int p_nInGroupStd,
            int p_nOutGroupMean, int p_nOutGroupStd,
            int p_nBatchID, int p_nRunID,
            int p_nVerbose)
  : m_nRuns(1),
    m_nIterations(p_nIterations),
    m_nNumberOfLevels(p_nNumberOfLevels),
    m_dInitialProportionDemocratic(p_dInitialProportionDemocratic),
    m_dRandomRevolutionChance(p_dRandomRevolutionChance),
    m_dCrossBorderChance(p_dCrossBorderChance),
    m_nRegimeDelay(p_nRegimeDelay),
    m_nBroadcastEffect(p_nBroadcastEffect),
    m_nCommunicationEffect(p_nCommunicationEffect),
    m_nFieldWidth(p_nFieldWidth),
    m_nFieldHeight(p_nFieldHeight),
    m_nBorderMultiplier(p_nBorderMultiplier),
    m_bUseDecay(p_bUseDecay),
    m_dDecayStart(p_dDecayStart),
    m_dDecayStrength(p_dDecayStrength),
    m_nThinning(p_nThinning),
    m_nMapThinning(100),
    m_nDetailThinning(p_nDetailThinning),
    m_nBatchID(p_nBatchID),
    m_nRunID(p_nRunID)
{
  srand((unsigned) time(0));
```

```
  m_bVerbose = p_nVerbose == 1;

  m_pIsolationNormal = new CNormal(p_nIsoMean, p_nIsoStd);
  m_pPopulationNormal = new CNormal(p_nPopMean, p_nPopStd);
  m_pAttitudeNormal = new CNormal(p_nAttMean, p_nAttStd);
  m_pInGroupNormal = new CNormal(p_nInGroupMean, p_nInGroupStd);
  m_pOutGroupNormal = new CNormal(p_nOutGroupMean, p_nOutGroupStd);

  m_nDotProgress = p_nIterations / 100;
  m_nLineProgress = p_nIterations / 10;
}

CModel::~CModel()
{
  int i;

  delete m_pIsolationNormal;
  delete m_pPopulationNormal;
  delete m_pAttitudeNormal;
  delete m_pInGroupNormal;
  delete m_pOutGroupNormal;

  for (i = 0; i < m_vProvinces.size(); ++i)
    delete m_vProvinces[i];

  for (i = 0; i < m_vCitizens.size(); ++i)
    delete m_vCitizens[i];

  for (i = 0; i < m_vCountries.size(); ++i)
    delete m_vCountries[i];

  delete m_pdConnectionMatrix;

  m_vCitizens.clear();
  m_vCountries.clear();
  m_vProvinces.clear();

  WriteCountryState(ACTION_CLOSE);
}

void CModel::Run(double *out_Moran, double *out_Proportion,
```

```
          double *out_ConnMatrix, double *out_AttMap,
          int *out_Population, int *out_Details,
          int *n_Countries)
{
  if (m_bVerbose) Rprintf("Setup...\n");

  if (out_Moran == NULL || out_Proportion == NULL || out_ConnMatrix == NULL
      || out_Population == NULL || out_Details == NULL)
    Rprintf("Null pointer passed to RRun() (%x; %x; %x; %x; %x)!\n",
        out_Moran, out_Proportion, out_ConnMatrix, out_Population,
        out_Details);

  m_R_out_Moran = out_Moran;
  m_R_out_Proportion = out_Proportion;
  m_R_out_ConnMatrix = out_ConnMatrix;
  m_R_out_AttMap = out_AttMap;
  m_R_out_Population = out_Population;
  m_R_out_Details = out_Details;

  Setup();

  if (m_bVerbose) Rprintf("Saving number of countries... (%d)\n",
            m_vCountries.size());
  *n_Countries = m_vCountries.size();

  if (m_bVerbose) Rprintf("Start simulation... (%d iterations)\n", m_nIterati

  for (m_nCurrentIteration = 1;
       m_nCurrentIteration <= m_nIterations;
       ++m_nCurrentIteration)
    Step();
}

void CModel::Step()
{
  int i,t,n;

  WriteCountryState(ACTION_WRITE);

  ReportToR();
```

```
  for (i = 0; i < m_vCountries.size(); ++i)
    m_vCountries[i]->UpdateIsolation();

  for (i = 0; i < m_vCitizens.size() / 10; ++i)
    m_vCitizens[CUniform::GetNextIntFromTo(0, m_vCitizens.size()-1)]
      ->DetermineCommunication();

  ShuffleCitizens();

  for (i = 0; i < m_vCitizens.size(); ++i)
    m_vCitizens[i]->DetermineProtest();

  m_vCountries[CUniform::GetNextIntFromTo(0, m_vCountries.size()-1)]
    ->GetCapital().Broadcast();

  for (i = 0; i < m_vCountries.size(); ++i)
    m_vCountries[i]->TestRevolution();

  PrintProgress();
}

void CModel::ShuffleCitizens()
{
  int i,a,b;
  CCitizen* pSwap;
  int nCitizens = m_vCitizens.size() - 1;

  for (i = 0; i < nCitizens / 2; ++i)
    {
      a = CUniform::GetNextIntFromTo(0, nCitizens);
      b = CUniform::GetNextIntFromTo(0, nCitizens);

      pSwap = m_vCitizens[a];
      m_vCitizens[a] = m_vCitizens[b];
      m_vCitizens[b] = pSwap;
    }
}

void CModel::Setup()
{
  int i;
```

```
  ClearDataStorage();
  m_nCoups = 0;
  m_nRevolutions = 0;

  if (m_bVerbose) Rprintf("Creating provinces...\n");
  CreateProvinces();
  if (m_bVerbose) Rprintf("Creating countries...\n");
  CreateCountries();
  if (m_bVerbose) Rprintf("Creating citizens...\n");
  CreateCitizens();
  if (m_bVerbose) Rprintf("Counting citizen...\n");
  CountCitizens();

  if (m_bVerbose) Rprintf("Creating connection matrix...\n");
  CreateConnectionMatrix();

  if (m_bVerbose) Rprintf("Storing connection matrix...\n");
  SaveConnectionMatrix();

  if (m_bVerbose) Rprintf("Resetting democracy memory...\n");
  for (i = 0; i < 80; ++i)
    m_nrgDemocracyMemory[i] = 0;

  WriteCountryState(ACTION_OPEN);
}

void CModel::ClearDataStorage()
{
  int i;

  for (i = 0; i < m_vCitizens.size(); ++i)
    delete m_vCitizens[i];

  for (i = 0; i < m_vCountries.size(); ++i)
    delete m_vCountries[i];

  for (i = 0; i < m_vProvinces.size(); ++i)
    delete m_vProvinces[i];

  m_vCitizens.clear();
```

```
  m_vCountries.clear();
  m_vProvinces.clear();
}

void CModel::CreateProvinces()
{
  int x,y,i = 0;
  CProvince *pNewProvince;

  for (x = 0; x < m_nFieldWidth; ++x)
    for (y = 0; y < m_nFieldHeight; ++y)
      {
    pNewProvince = new CProvince(this, x, y, ++i);
    pNewProvince->CreateSovereign();
    m_vProvinces.push_back(pNewProvince);
      }
}

void CModel::CreateCountries()
{
  int i, x, y, k;

  for (i = 0; i < m_nFieldWidth * m_nFieldHeight * m_nBorderMultiplier; ++i)
    {
      x = rand() % m_nFieldWidth;
      y = rand() % m_nFieldHeight;
      k = rand() % 4;

      GetNeighbour(x,y,k).GetCountry().Conquer(GetProvince(x,y));
    }

  for (i = 0; i < m_vProvinces.size(); ++i)
    if (m_vProvinces[i]->IsCapital())
      m_vCountries.push_back(&m_vProvinces[i]->GetCountry());

  for (i = 0; i < m_vCountries.size(); ++i)
    m_vCountries[i]->SetID(i);
}

void CModel::CountCitizens()
{
```

```
  int i;

  for (i = 0; i < m_vCountries.size(); ++i)
    m_R_out_Population[i] = m_vCountries[i]->GetPopulation();
}

void CModel::CreateCitizens()
{
  int i, j, nPopulation;
  CCitizen* pNewCitizen;

  for (i = 0; i < m_vProvinces.size(); ++i)
    {
      nPopulation = m_pPopulationNormal->GetNextIntFrom(1);

      for (j = 0; j < nPopulation; ++j)
    {
      pNewCitizen = new CCitizen(this, m_vProvinces[i]);
      pNewCitizen->SetAttitude(m_pAttitudeNormal
                  ->GetNextIntFromTo(0,m_nNumberOfLevels));
      pNewCitizen->SetOutGroupThreshold(m_pOutGroupNormal
                      ->GetNextIntFrom(0));
      pNewCitizen->SetInGroupThreshold(m_pInGroupNormal
                     ->GetNextIntFrom(0));
      m_vProvinces[i]->GetCountry().AddCitizen(pNewCitizen);
      m_vProvinces[i]->AddCitizen(pNewCitizen);
      m_vCitizens.push_back(pNewCitizen);
    }
    }
}

void CModel::CreateConnectionMatrix()
{
  int i,j,k,id1,id2,s;
  int nCountries = m_vCountries.size();

  // Create empty connection matrix
  m_pdConnectionMatrix = new double[nCountries * nCountries];
  for (i = 0; i < nCountries; ++i)
    for (j = 0; j < nCountries; ++j)
      m_pdConnectionMatrix[i * nCountries + j] = 0.0;
```

```
  // Set all borders to one
  for (i = 0; i < m_vProvinces.size(); ++i)
    {
      id1 = m_vProvinces[i]->GetCountry().GetID();

      id2 = m_vProvinces[i]->GetNeighbour(NORTH).GetCountry().GetID();
      if (id1 != id2)
    {
      m_pdConnectionMatrix[id1 * nCountries + id2] = 1;
      m_pdConnectionMatrix[id2 * nCountries + id1] = 1;
    }

      id2 = m_vProvinces[i]->GetNeighbour(WEST).GetCountry().GetID();
      if (id1 != id2)
    {
      m_pdConnectionMatrix[id1 * nCountries + id2] = 1;
      m_pdConnectionMatrix[id2 * nCountries + id1] = 1;
    }
    }

  // Standardize to have rows add up to one
  for (i = 0; i < nCountries; ++i)
    {
      s = 0;

      for (j = 0; j < nCountries; ++j)
    s += (int) m_pdConnectionMatrix[i * nCountries + j];

      for (j = 0; j < nCountries; ++j)
    m_pdConnectionMatrix[i * nCountries + j] /= (double) s;
    }
}

void CModel::SaveConnectionMatrix()
{
  int i,j;
  int nCountries = m_vCountries.size();

  for (i = 0; i < nCountries * nCountries; ++i)
    m_R_out_ConnMatrix[i] = m_pdConnectionMatrix[i];
```

```
}

void CModel::PrintMap()
{
  int x,y,i,t;

  for (y = 0; y < m_nFieldHeight; ++y)
    {
      for (x = 0; x < m_nFieldWidth; ++x)
    if (m_vProvinces.at(x * m_nFieldHeight + y)
        ->GetCountry().IsDemocracy())
      //cout << (char) 178 << (char) 178;
      cout << "##";
    else
      //cout << (char) 219 << (char) 219;
      cout << "..";

      cout << endl;
    }

  i = m_nCurrentIteration / 100;

  m_nrgDemocracyMemory[i-1] = GetProportionDemocratic();

  for (t = 0; t < i; t += 2)
    {
      for (x = 0; x < m_nrgDemocracyMemory[t] * 40; ++x)
    cout << " ";
      cout << "*" << endl;
    }

  for (x = 0; x < m_nFieldWidth; ++x)
    cout << "--";

  cout << " " << fixed << setprecision(2) <<
    ((double) m_nCurrentIteration * 100 / 8000) << " %";

  cout << endl;
}

CProvince& CModel::GetProvince(int x, int y)
```

```
{
  if (x > m_nFieldWidth - 1 || x < 0
      || y > m_nFieldHeight - 1 || y < 0)
    {
      ostringstream strMessage;
      strMessage << "Attempt to get non-existing province at ("
         << x << "," << y << ")";
      throw CException(strMessage.str());
    }

  return *m_vProvinces.at(x * m_nFieldHeight + y);
}

CProvince& CModel::GetNeighbour(int x, int y, int k)
{
  if (k > 3 || k < 0
      || x > m_nFieldWidth - 1 || x < 0
      || y > m_nFieldHeight - 1 || y < 0)
    {
      ostringstream strMessage;
      strMessage << "Neighbour requested of impossible province - x,y,k: "
         << x << "," << y << "," << k;
      throw CException(strMessage.str());
    }

  switch (k)
    {
    case NORTH:
      if (y == 0) y = m_nFieldHeight - 1;
      else --y;
      break;

    case SOUTH:
      if (y == m_nFieldHeight - 1) y = 0;
      else ++y;
      break;

    case WEST:
      if (x == 0) x = m_nFieldWidth - 1;
      else --x;
      break;
```

```
    case EAST:
      if (x == m_nFieldWidth - 1) x = 0;
      else ++x;
      break;
    }

  return GetProvince(x,y);
}

void CModel::AddRevolution(bool p_bCoup)
{
  p_bCoup ? m_nCoups++ : m_nRevolutions++;
}

double CModel::GetAverageAttitude()
{
  int i;
  double p = 0.0;

  for (i = 0; i < m_vCitizens.size(); ++i)
    p += m_vCitizens[i]->GetAttitude();

  return p / (double) (m_vCitizens.size() * m_nNumberOfLevels);
}

double CModel::GetProportionDemocratic()
{
  int i;
  double p = 0.0;

  for (i = 0; i < m_vCountries.size(); ++i)
    p += (m_vCountries[i]->IsDemocracy() ? 1 : 0)
      * m_vCountries[i]->GetCitizens().size();

  return p / (double) m_vCitizens.size();
}

double CModel::GetProportionRevolutions()
{
  if (m_nRevolutions + m_nCoups > 0)
```

```
    return (double) m_nRevolutions
      / (double) (m_nRevolutions + m_nCoups);
  else
    return 0.0;
}

double CModel::GetMoranI()
{
  int i,j;
  double num1 = 0, num2 = 0, nom1 = 0, nom2 = 0;
  double avg, demi, demj;
  int nCountries = m_vCountries.size();

  avg = GetProportionDemocratic();

  for (i = 0; i < nCountries; ++i)
    {
      demi = (m_vCountries[i]->IsDemocracy() ? 1 : 0);
      nom2 += (demi - avg) * (demi - avg);

      for (j = 0; j < nCountries; ++j)
    {
      demj = (m_vCountries[j]->IsDemocracy() ? 1 : 0);
      nom1 += m_pdConnectionMatrix[i * nCountries + j];
      num1 += m_pdConnectionMatrix[i * nCountries + j]
        * (demi - avg) * (demj - avg);
    }
    }

  num2 = nCountries;

  if (nom2 < 0.0000001) // All countries are either democratic or non-democratic
    nom2 = 0.0000001;

  return num1 / nom1 * num2 / nom2;
}

void CModel::PrintProgress()
{
  if (m_nCurrentIteration % m_nDotProgress == 0 && m_bVerbose)
    if (m_nCurrentIteration % m_nLineProgress == 0)
```

```
      Rprintf("(%d)", m_nCurrentIteration);
    else
      Rprintf(".");
}

void CModel::WriteCountryState(int action)
{
  int i;
  char czFilename[50];

  switch (action)
    {
    case ACTION_OPEN:
      sprintf(czFilename, "output/batch%d_%d_countries.data",
          m_nBatchID, m_nRunID);
      if (m_bVerbose)
    Rprintf("Opening country state file (%s)...\n", czFilename);
      m_fCountryFile = fopen(czFilename, "w");
      if (m_bVerbose && m_fCountryFile == NULL)
    Rprintf("WARNING! Country state file was not opened.\n");
      break;

    case ACTION_CLOSE:
      fclose(m_fCountryFile);
      break;

    case ACTION_WRITE:
      if (m_fCountryFile != NULL)
    for (i = 0; i < m_vCountries.size(); ++i)
      fputs((m_vCountries[i]->IsDemocracy() ? "1" : "0"), m_fCountryFile);
      fputs("\n", m_fCountryFile);
      break;
    }
}

void CModel::ReportToR()
{
  int i, t, n;
  int nVariables = 8;

  try
```

```
{
  if (m_nCurrentIteration % m_nThinning == 0)
{
  m_R_out_Moran[m_nCurrentIteration / m_nThinning] = GetMoranI();
  m_R_out_Proportion[m_nCurrentIteration / m_nThinning]
    = GetProportionDemocratic();
}

  if (m_nCurrentIteration % m_nDetailThinning == 0)
{
  t = m_nCurrentIteration / m_nDetailThinning - 1;
  n = m_vCountries.size();

  for (i = 0; i < n; ++i)
    {
      m_R_out_Details[t * n * nVariables + i]
    = (int) m_vCountries[i]->IsDemocracy();
      m_R_out_Details[t * n * nVariables + n + i]
    = m_vCountries[i]->GetProtesting();
      m_R_out_Details[t * n * nVariables + 2 * n + i]
    = m_vCountries[i]->GetAvgAttitude();
      m_R_out_Details[t * n * nVariables + 3 * n + i]
    = m_vCountries[i]->GetCoupCount();
      m_R_out_Details[t * n * nVariables + 4 * n + i]
    = m_vCountries[i]->GetRevolutionCount();
      m_R_out_Details[t * n * nVariables + 5 * n + i]
    = m_vCountries[i]->GetRegimeAge();
      m_R_out_Details[t * n * nVariables + 6 * n + i]
    = m_vCountries[i]->GetLastRegimeChange();
      m_R_out_Details[t * n * nVariables + 7 * n + i]
    = m_vCountries[i]->GetInitialRegime();
    }
}

  if (m_nCurrentIteration % m_nMapThinning == 0)
{
  int nProvinces = m_vProvinces.size();
  for (i = 0; i < nProvinces; ++i)
    m_R_out_AttMap[nProvinces*(m_nCurrentIteration/m_nMapThinning)
          + i - nProvinces]
      = m_vProvinces[i]->GetAverageAttitude();
```

```
    }
    }
  catch (exception& e)
    {
      Rprintf("Exception %s caught\n", e.what());
    }
}
```

province.h

```
#ifndef _PROVINCE_H
#define _PROVINCE_H

#include <vector>

using std::vector;

class CModel;
class CCountry;
class CCitizen;

class CProvince
{
 public:
  CProvince(CModel* p_pModel, int p_nX, int p_nY, int p_nID);
  virtual ~CProvince();

  void CreateSovereign();
  bool IsCompatriot(CProvince& p_Province);
  bool IsCapital();
  CProvince& GetNeighbour(int k);
  void AddProtester();
  void RemoveProtester();
  bool AllProtesting();
  double GetProtestProportion();
  void AddCitizen(CCitizen* p_pCitizen);
  void Broadcast();
  void ReceiveBroadcast(bool p_bDemocracy);

  CCountry& GetCountry() { return *m_pCountry; }
  int GetID() { return m_nID; }
```

```
  int GetX() { return m_nX; }
  int GetY() { return m_nY; }
  double GetAverageAttitude();
  vector<CCitizen*>& GetCitizens() { return m_vCitizens; }

  void SetCountry(CCountry* p_pCountry);

 private:
  int m_nX, m_nY;
  int m_nID;
  int m_nProtesters;
  vector<CCitizen*> m_vCitizens;
  CCountry* m_pCountry;
  CModel* m_pModel;
};

#endif
```

province.cpp

```
#include "province.h"

#include "country.h"
#include "uniform.h"
#include "model.h"

#define MAX(a,b) ((a)>(b)?(a):(b))

CProvince::CProvince(CModel* p_pModel, int p_nX, int p_nY, int p_nID)
  : m_nX(p_nX),
    m_nY(p_nY),
    m_nID(p_nID),
    m_pModel(p_pModel),
    m_nProtesters(0)
{
}

CProvince::~CProvince()
{
}
```

```
void CProvince::CreateSovereign()
{
  m_pCountry = new CCountry(m_pModel, this);
}

bool CProvince::IsCompatriot(CProvince& p_Province)
{
  return m_pCountry->GetID() == p_Province.GetCountry().GetID();
}

bool CProvince::IsCapital()
{
  return m_pCountry->GetCapital().GetID() == m_nID;
}

CProvince& CProvince::GetNeighbour(int k)
{
  return m_pModel->GetNeighbour(m_nX, m_nY, k);
}

void CProvince::SetCountry(CCountry* p_pCountry)
{
  m_pCountry->RemoveProvince(this);
  m_pCountry = p_pCountry;
}

void CProvince::AddCitizen(CCitizen* p_pCitizen)
{
  m_vCitizens.push_back(p_pCitizen);
}

void CProvince::AddProtester()
{
  m_nProtesters++;
}

void CProvince::RemoveProtester()
{
  m_nProtesters--;
}
```

```
bool CProvince::AllProtesting()
{
  return m_nProtesters == m_vCitizens.size();
}

double CProvince::GetProtestProportion()
{
  return (double) m_nProtesters / (double) m_vCitizens.size();
}

void CProvince::Broadcast()
{
  if (!m_pCountry->IsDemocracy())
    return;

  int i, iso1, iso2;
  vector<CProvince*> vpNeighbours;

  vpNeighbours.push_back(&GetNeighbour(WEST));
  vpNeighbours.push_back(&GetNeighbour(NORTH).GetNeighbour(WEST));
  vpNeighbours.push_back(&GetNeighbour(NORTH));
  vpNeighbours.push_back(&GetNeighbour(NORTH).GetNeighbour(EAST));
  vpNeighbours.push_back(&GetNeighbour(EAST));
  vpNeighbours.push_back(&GetNeighbour(SOUTH).GetNeighbour(EAST));
  vpNeighbours.push_back(&GetNeighbour(SOUTH));
  vpNeighbours.push_back(&GetNeighbour(SOUTH).GetNeighbour(WEST));
  vpNeighbours.push_back(this);

  for (i = 0; i < vpNeighbours.size(); ++i)
    {
      if (&vpNeighbours[i]->GetCountry() != m_pCountry)
    {
      iso1 = m_pCountry->GetIsolation();
      iso2 = vpNeighbours[i]->GetCountry().GetIsolation();

      iso1 = MAX(iso1,iso2);
    }
      else
    iso1 = 0;

      if (CUniform::GetNextIntFromTo(0,100) > iso1)
```

```
    vpNeighbours[i]->ReceiveBroadcast(m_pCountry->IsDemocracy());
    }
}

void CProvince::ReceiveBroadcast(bool p_bDemocracy)
{
  int i;

  for (i = 0; i < m_vCitizens.size(); ++i)
    m_vCitizens[i]->ReceiveBroadcast(p_bDemocracy);
}

double CProvince::GetAverageAttitude()
{
  int i, s = 0;

  for (i = 0; i < m_vCitizens.size(); ++i)
    s += m_vCitizens[i]->GetAttitude();

  return (double) s / m_vCitizens.size();
}
```

country.h

```
#ifndef _COUNTRY_H
#define _COUNTRY_H

#include <vector>

using std::vector;

class CProvince;
class CCitizen;
class CModel;

class CCountry
{
 public:
  CCountry(CModel* p_pModel, CProvince* p_pProvince);
  virtual ~CCountry();
```

```
 void Conquer(CProvince& p_Province);
 void UpdateIsolation();
 bool OwnsProvince(CProvince& p_Province);
 bool IsAtom() { return m_vProvinces.size() == 1; }
 void RemoveProvince(CProvince* p_pProvince);
 CProvince& GetCapital();
 void SetCapital(CProvince& p_Capital);
 int GetSize() { return m_vProvinces.size(); }
 void AddCitizen(CCitizen* p_pCitizen) { m_vCitizens.push_back(p_pCitizen); }
 void TestRevolution();

 int GetID() { return m_nID; }
 bool IsDemocracy() { return m_bDemocracy; }
 int GetIsolation() { return m_nIsolation; }
 int GetDelay() { return m_nDelay; }
 vector<CProvince*>& GetProvinces() { return m_vProvinces; }
 vector<CCitizen*>& GetCitizens() { return m_vCitizens; }

 void SetID(int p_nID) { m_nID = p_nID; }
 void SetDemocracy(bool p_bDemocracy) { m_bDemocracy = p_bDemocracy; }
 void SetIsolation(int p_nIsolation) { m_nIsolation = p_nIsolation; }

 int GetPopulation();
 int GetProtesting();
 int GetAvgAttitude();
 int GetCoupCount() { return m_nCoupCount; }
 int GetRevolutionCount() { return m_nRevolutionCount; }
 int GetRegimeAge() { return m_nRegimeAge; }
 int GetLastRegimeChange() { return m_nLastRevolution; }
 int GetInitialRegime() { return m_bOriginalRegime ? 1 : 0; }

private:
 bool CheckConnectednessWithoutProvince(CProvince& p_RemovedProvince);
 void CheckCWPRecursive(CProvince& p_CurrentProvince,
            CProvince& p_RemovedProvince,
            vector<int>& p_vCheckedProvinces);

 int m_nID;
 bool m_bDemocracy;
 bool m_bOriginalRegime;
 int m_nIsolation;
```

```
  int m_nDelay;
  int m_nLastRevolution;
  int m_nRegimeAge;
  int m_nCoupCount;
  int m_nRevolutionCount;
  double m_dCoupChance;
  vector<CProvince*> m_vProvinces;
  vector<CCitizen*> m_vCitizens;
  CProvince* m_pCapital;
  CModel* m_pModel;
};

#endif
```

country.cpp

```
#include "country.h"

#include <algorithm>
#include <iostream>
#include <iomanip>
#include <cmath>
#include "province.h"
#include "model.h"
#include "uniform.h"

#define MAX(a,b) ((a)>(b)?(a):(b))
#define MIN(a,b) ((a)>(b)?(b):(a))

using std::find;
using std::cout;
using std::endl;
using std::left;
using std::right;
using std::setw;

CCountry::CCountry(CModel* p_pModel, CProvince* p_pProvince)
  : m_bOriginalRegime(true),
    m_nDelay(0),
    m_nLastRevolution(0),
    m_nRegimeAge(0),
```

```
    m_dCoupChance(0),
    m_nCoupCount(0),
    m_nRevolutionCount(0),
    m_pModel(p_pModel)
{
  // Countries are created by the province which becomes the capital
  m_vProvinces.push_back(p_pProvince);
  m_pCapital = p_pProvince;
  m_nID = p_pProvince->GetID();
  m_nIsolation = p_pModel->GetIsolationNormal().GetNextIntFromTo(0,100);
  m_bDemocracy = CUniform::GetBoolean(p_pModel->
                    GetInitialProportionDemocratic());
}

CCountry::~CCountry()
{
}

void CCountry::UpdateIsolation()
{
  if (m_bDemocracy)
    m_nIsolation = 0;
  else
    {
      m_nIsolation = m_nIsolation + CUniform::GetNextIntFromTo(-1,1);
      m_nIsolation = MAX(0,MIN(100,m_nIsolation));
    }
}

void CCountry::TestRevolution()
{
  m_nRegimeAge++;

  if (m_pModel->UseDecay() && !m_bOriginalRegime)
    {
      m_dCoupChance = m_pModel->GetDecayStart()
                    * exp(m_pModel->GetDecayStrength()
                  * m_nRegimeAge);
      m_dCoupChance = MAX(m_dCoupChance,
              m_pModel->GetRandomRevolutionChance());
    }
```

```
  else
    m_dCoupChance = m_pModel->GetRandomRevolutionChance();

  bool bChange = false;

  if (m_pCapital->AllProtesting() && !m_nDelay)
    {
      bChange = true;
      m_pModel->AddRevolution(false);
      m_nLastRevolution = 1;
      m_nRevolutionCount++;
    }
  else if (CUniform::GetBoolean(m_dCoupChance))
    {
      bChange = true;
      m_pModel->AddRevolution(true);
      m_nLastRevolution = 2;
      m_nCoupCount++;
    }

  if (bChange)
    {
      int i;

      m_bDemocracy = !m_bDemocracy;

      for (i = 0; i < m_vCitizens.size(); ++i)
    m_vCitizens[i]->SetProtesting(false);

      m_nDelay = m_pModel->GetRegimeDelay();
      m_bOriginalRegime = false;
      m_nRegimeAge = 0;
    }

  if (m_nDelay)
    --m_nDelay;
}

CProvince& CCountry::GetCapital()
{
  return *m_pCapital;
```

```
}

void CCountry::SetCapital(CProvince& p_Capital)
{
  m_pCapital = &p_Capital;
}

void CCountry::Conquer(CProvince& p_Province)
{
  // If the province to be conquered is not already part of this country
  // and the country of the province would still be connected after conquering
  // then set the country of the province to the this one
  // and add the province to this country's provinces
  if (p_Province.GetCountry().GetID() != m_nID
      && p_Province.GetCountry().CheckConnectednessWithoutProvince(p_Province))
    {
      if (p_Province.GetCountry().GetProvinces().size() == 1)
    delete &p_Province.GetCountry();

      p_Province.SetCountry(this);
      m_vProvinces.push_back(&p_Province);
    }
}

void CCountry::RemoveProvince(CProvince* p_pProvince)
{
  vector<CProvince*>::iterator iter;

  // Find this province and remove it
  iter = find(m_vProvinces.begin(), m_vProvinces.end(), p_pProvince);
  if (iter != m_vProvinces.end())
    m_vProvinces.erase(iter);

  // If the province to be removed is the capital, and there are other
  // provinces left, set a new, random capital
  if (m_pCapital == p_pProvince && m_vProvinces.size())
    m_pCapital = m_vProvinces[rand() % m_vProvinces.size()];
}

bool CCountry::CheckConnectednessWithoutProvince(CProvince& p_RemovedProvince)
{
```

```
  vector<int> vCheckedProvinces;

  // If this is the only province in the country, no further checking needed
  if (IsAtom())
    return true;

  // Recursively find out whether provinces remaining are still connected
  CheckCWPRecursive(*m_pCapital, p_RemovedProvince, vCheckedProvinces);

  return vCheckedProvinces.size() == m_vProvinces.size() - 1;
}

void CCountry::CheckCWPRecursive(CProvince& p_CurrentProvince,
                  CProvince& p_RemovedProvince,
                  vector<int>& p_vCheckedProvinces)
{
  int k, nCurrentID;

  nCurrentID = p_CurrentProvince.GetID();

  // If the current province is a compatriot of the capital
  // and it is not the province to be removed
  // and it is not among the provinces already checked
  // then add to provinces checked
  // and check all four neighbouring provinces
  if (m_pCapital->IsCompatriot(p_CurrentProvince)
      && nCurrentID != p_RemovedProvince.GetID()
      && find(p_vCheckedProvinces.begin(), p_vCheckedProvinces.end(),
          nCurrentID) == p_vCheckedProvinces.end())
    {
      p_vCheckedProvinces.push_back(nCurrentID);

      // Check all four neighbouring provinces, unless all provinces of
      // this country have already been checked
      for (k = 0;
       k < 4 && p_vCheckedProvinces.size() < m_vProvinces.size() - 1;
       ++k)
    CheckCWPRecursive(p_CurrentProvince.GetNeighbour(k),
             p_RemovedProvince, p_vCheckedProvinces);
    }
}
```

```
int CCountry::GetPopulation()
{
  int i, s = 0;

  for (i = 0; i < m_vProvinces.size(); ++i)
    s += m_vProvinces[i]->GetCitizens().size();

  return s;
}

int CCountry::GetProtesting()
{
  int i, s = 0;

  for (i = 0; i < m_vCitizens.size(); ++i)
    s += (int) m_vCitizens[i]->IsProtesting();

  return s;
}

int CCountry::GetAvgAttitude()
{
  int i, s = 0;

  for (i = 0; i < m_vCitizens.size(); ++i)
    s += m_vCitizens[i]->GetAttitude();

  return (int) round((double) s / (double) m_vCitizens.size());
}
```

citizen.h

```
#ifndef _CITIZEN_H
#define _CITIZEN_H

class CProvince;
class CModel;

class CCitizen
{
```

```
 public:
  CCitizen(CModel* p_pModel, CProvince* p_pProvince);
  virtual ~CCitizen();

  void DetermineCommunication();
  void Communicate(CCitizen* p_pCitizen);
  void ReceiveBroadcast(bool p_bDemocracy);
  void DetermineProtest();
  void SetProtesting(bool p_bProtesting);

  bool IsProtesting() { return m_bProtesting; }
  int GetAttitude() { return m_nAttitude; }
  int GetInGroupThreshold() { return m_nInGroupThreshold; }
  int GetOutGroupThreshold() { return m_nOutGroupThreshold; }
  CProvince* GetProvince() { return m_pProvince; }

  void SetAttitude(int p_nAttitude) { m_nAttitude = p_nAttitude; }
  void SetInGroupThreshold(int p_nInGroupThreshold);
  void SetOutGroupThreshold(int p_nOutGroupThreshold);
  void SetProvince(CProvince& p_Province) { m_pProvince = &p_Province; }

 private:
  bool m_bProtesting;
  int m_nAttitude;
  int m_nInGroupThreshold, m_nOutGroupThreshold;
  CProvince* m_pProvince;
  CModel* m_pModel;
};

#endif
```

citizen.cpp

```
#include "citizen.h"

#include <cmath>
#include "province.h"
#include "model.h"
#include "country.h"
#include "uniform.h"
```

```
extern "C" {
#include <R.h>
}

using std::abs;

#define MAX(a,b) ((a)>(b)?(a):(b))
#define MIN(a,b) ((a)<(b)?(a):(b))

CCitizen::CCitizen(CModel* p_pModel, CProvince* p_pProvince)
  : m_pModel(p_pModel),
    m_pProvince(p_pProvince),
    m_nInGroupThreshold(0),
    m_nOutGroupThreshold(0),
    m_bProtesting(false)
{
}

CCitizen::~CCitizen()
{
}

void CCitizen::DetermineCommunication()
{
  CProvince* pProvince = 0;
  CProvince* pNeighbour;
  CCitizen* pCitizen;
  int i,iso1,iso2;
  double k,t;
  int nSumIsolation = 0;

  k = CUniform::GetDouble();

  for (i = 0; i < 4 && pProvince == 0; ++i)
    {
      pNeighbour = &m_pProvince->GetNeighbour(i);
      if (&pNeighbour->GetCountry() != &m_pProvince->GetCountry())
    {
      iso1 = pNeighbour->GetCountry().GetIsolation();
      iso2 = m_pProvince->GetCountry().GetIsolation();
```

```
      t = MAX(iso1,iso2) * m_pModel->GetCrossBorderChance() / 400;
    }
      else
    t = m_pModel->GetCrossBorderChance() / 4;

      if (k < t)
    pProvince = pNeighbour;
      else
    k -= t;
    }

  if (pProvince == 0)
    pProvince = m_pProvince;

  if (pProvince->GetCitizens().size() > 0)
    {
      pCitizen = pProvince->GetCitizens()
    [CUniform::GetNextIntFromTo(0, pProvince->GetCitizens().size() - 1)];

      Communicate(pCitizen);
    }
}

void CCitizen::Communicate(CCitizen* p_pCitizen)
{
  int nDistance, nAttitude1, nAttitude2, nMaxLevel;
  int nCommEffect = m_pModel->GetCommunicationEffect();

  nAttitude1 = GetAttitude();
  nAttitude2 = p_pCitizen->GetAttitude();
  nMaxLevel = m_pModel->GetNumberOfLevels() - 1;

  nDistance = abs(nAttitude1 - nAttitude2);

  if (nDistance < m_nInGroupThreshold)
    nAttitude2 > nAttitude1 ? nAttitude1 += nCommEffect
      : nAttitude1 -= nCommEffect;
  else if (nDistance > m_nOutGroupThreshold)
    nAttitude2 > nAttitude1 ? nAttitude1 -= nCommEffect
      : nAttitude1 += nCommEffect;
```

```
  SetAttitude(MAX(0,MIN(nMaxLevel,nAttitude1)));
}

void CCitizen::ReceiveBroadcast(bool p_bDemocracy)
{
  if (p_bDemocracy)
    m_nAttitude += m_pModel->GetBroadcastEffect();
  else
    m_nAttitude -= m_pModel->GetBroadcastEffect();

  if (m_nAttitude < 0)
    m_nAttitude = 0;
  else
    {
      int nMaxLevel = m_pModel->GetNumberOfLevels() - 1;

      if (m_nAttitude > nMaxLevel)
    m_nAttitude = nMaxLevel;
    }
}

void CCitizen::DetermineProtest()
{
  double dProtestProportion, dAttitudeProportion;
  bool bDemocracy;

  dProtestProportion = m_pProvince->GetProtestProportion();
  dAttitudeProportion = (double) m_nAttitude
    / (double) (m_pModel->GetNumberOfLevels() - 1);
  bDemocracy = m_pProvince->GetCountry().IsDemocracy();

  if (bDemocracy && dAttitudeProportion <= dProtestProportion)
    SetProtesting(true);
  else if (!bDemocracy && (1.0 - dAttitudeProportion) <= dProtestProportion)
    SetProtesting(true);
  else
    SetProtesting(false);
}

void CCitizen::SetInGroupThreshold(int p_nInGroupThreshold)
{
```

```
  if (p_nInGroupThreshold > 0)
    if (p_nInGroupThreshold < m_nOutGroupThreshold)
      m_nInGroupThreshold = p_nInGroupThreshold;
    else
      m_nInGroupThreshold = m_nOutGroupThreshold;
  else
    m_nInGroupThreshold = 0;
}

void CCitizen::SetOutGroupThreshold(int p_nOutGroupThreshold)
{
  if (p_nOutGroupThreshold > m_nInGroupThreshold)
    if (p_nOutGroupThreshold > 0)
      m_nOutGroupThreshold = p_nOutGroupThreshold;
    else
      m_nOutGroupThreshold = 0;
  else
    m_nOutGroupThreshold = m_nInGroupThreshold;
}

void CCitizen::SetProtesting(bool p_bProtesting)
{
  if (p_bProtesting && !m_bProtesting)
    {
      m_bProtesting = true;
      m_pProvince->AddProtester();
    }
  else if (!p_bProtesting && m_bProtesting)
    {
      m_bProtesting = false;
      m_pProvince->RemoveProtester();
    }
}
```

uniform.h

```
#ifndef _UNIFORM_H
#define _UNIFORM_H

class CUniform
{
```

```
 public:
  static double GetDouble();
  static bool GetBoolean(double p_dProbability);
  static int GetNextIntFromTo(int p_nFrom, int p_nTo);
};

#endif
```

uniform.cpp

```
#include "uniform.h"

#include <cstdlib>

double CUniform::GetDouble()
{
  return (double) rand() / (double) RAND_MAX;
}

bool CUniform::GetBoolean(double p_dProbability)
{
  return CUniform::GetDouble() < p_dProbability;
}

int CUniform::GetNextIntFromTo(int p_nFrom, int p_nTo)
{
  return rand() % (p_nTo - p_nFrom + 1) + p_nFrom;
}
```

normal.h

```
#ifndef _NORMAL_H
#define _NORMAL_H

class CNormal
{
 public:
  CNormal(double p_dMean, double p_dStandardDeviation);
  virtual ~CNormal();

  int GetNextIntFromTo(int p_nFrom, int p_nTo);
```

```
  int GetNextIntFrom(int p_nFrom);
  double GetNextDouble();

  double GetMean() { return m_dMean; }
  double GetStandardDeviation() { return m_dStandardDeviation; }

 private:
  double m_dMean;
  double m_dStandardDeviation;
  bool m_bHasStoredValue;
  double m_dStoredValue;
};

#endif
```

normal.cpp

The below code makes use of Box and Muller (1958).

```
#include "normal.h"

#define MAX(a,b) ((a)>(b)?(a):(b))
#define MIN(a,b) ((a)>(b)?(b):(a))

#include <cmath>
#include "uniform.h"

CNormal::CNormal(double p_dMean, double p_dStandardDeviation)
  : m_dMean(p_dMean),
    m_dStandardDeviation(p_dStandardDeviation),
    m_bHasStoredValue(false)
{
}

CNormal::~CNormal()
{
}

int CNormal::GetNextIntFromTo(int p_nFrom, int p_nTo)
{
  int nValue = (int) round(GetNextDouble());
```

```
  return MIN(MAX(p_nFrom, nValue), p_nTo);
}

int CNormal::GetNextIntFrom(int p_nFrom)
{
  int nValue = (int) round(GetNextDouble());
  return MAX(p_nFrom, nValue);
}

double CNormal::GetNextDouble()
{
  if (m_bHasStoredValue)
    {
      m_bHasStoredValue = false;
      return m_dStoredValue;
    }
  else
    {
      // See G.E.P. Box and Mervin E. Muller, "A note on the generation
      // of random normal deviates", The Annals of Mathematical Statistics,
      // 1958 (29:2), pp. 610-611.

      double u1 = CUniform::GetDouble();
      double u2 = CUniform::GetDouble();

      m_dStoredValue = pow(-2 * log(u1), .5) * sin(6.28318530717959 * u2)
    * m_dStandardDeviation + m_dMean;
      m_bHasStoredValue = true;

      return pow(-2 * log(u2), .5) * cos(6.28318530717959 * u1)
    * m_dStandardDeviation + m_dMean;
    }
}
```

exception.h

```
#ifndef _EXCEPTION_H
#define _EXCEPTION_H

#include <string>
```

```
using std::string;

class CException
{
 public:
  CException(string p_strMessage) { m_strMessage = p_strMessage; }
  virtual ~CException() {};

  string& GetMessage() { return m_strMessage; }

 private:
  string m_strMessage;
};

#endif
```

exception.cpp

G.2 R base code

import.R

```
dyn.load("model.so")

factor.to.numeric <- function(x)
  {
    if (is.null(dim(x)))
      x <- as.numeric(levels(x))[x]
    else
      for (i in 1:dim(x)[2])
        x[,i] <- as.numeric(levels(x[,i]))[x[,i]]

    x
  }

simulation <- function(iter=8000, thin=10, nlevels=100, propDemoc=.01,
                       randRev=.0001, crossB=.5, regDelay=50, broadcast=NA,
                       comm.effect=1, width=20, height=20, multiplier=6,
                       useDecay=TRUE, decayStart=.05, decayStrength=-.15,
```

```
                    isoMean=50, isoStd=10, popMean=50, popStd=10,
                    attMean=10, attStd=5, inGMean=10, inGStd=4,
                    outGMean=90, outGStd=4, verbose=TRUE, detailThin=100,
                    batchID=0, runID=0) {

if (is.na(broadcast))
  broadcast = floor(nlevels / 10)

nCountryVariables <- 8

outMoran <- as.double(rep(0, iter / thin))
outProportion <- as.double(rep(0, iter / thin))
outConnMatrix <- as.double(rep(0, (width*height)^2))
outAttMap <- as.double(rep(0, (width*height)*iter/100))
outPopulation <- as.integer(rep(0, width*height))
outDetails <- as.integer(rep(0, nCountryVariables * width*height
                             * (iter / detailThin)))
outTimeTaken <- as.integer(0)
nCountries <- as.integer(0)

out <- .C("runSimulation",
          as.integer(iter),
          as.integer(thin),
          as.integer(detailThin),
          as.integer(nlevels),
          as.double(propDemoc),
          as.double(randRev),
          as.double(crossB),
          as.integer(regDelay),
          as.integer(broadcast),
          as.integer(comm.effect),
          as.integer(width),
          as.integer(height),
          as.integer(multiplier),
          as.integer(useDecay),
          as.double(decayStart),
          as.double(decayStrength),
          as.integer(isoMean),
          as.integer(isoStd),
          as.integer(popMean),
          as.integer(popStd),
```

```
            as.integer(attMean),
            as.integer(attStd),
            as.integer(inGMean),
            as.integer(inGStd),
            as.integer(outGMean),
            as.integer(outGStd),
            as.integer(batchID),
            as.integer(runID),
            outMoran = outMoran,
            outProportion = outProportion,
            outConnMatrix = outConnMatrix,
            outAttMap = outAttMap,
            outPopulation = outPopulation,
            outDetails = outDetails,
            outTimeTaken = outTimeTaken,
            nCountries = nCountries,
            as.integer(verbose))

  cat("\nReached end of simulation - returning data\n");

  nData <- iter / detailThin
  country <- rep(1:out$nCountries, nCountryVariables * nData)
  iteration <- rep(1:nData, each = nCountryVariables * out$nCountries)
  variable <- rep(rep(c("democracy","protest","avg.att","ncoups","nrevs",
                        "reg.age","last.change","init.reg"),
                      each = out$nCountries), nData)
  details <- data.frame(cbind(country, iteration, variable,
                   out$outDetails[1:length(country)]))
  colnames(details) <- c("country","iteration","variable","value")
  details <- reshape(details, direction="wide", idvar=c("country","iteration"
                     timevar="variable")
  colnames(details) <- c("country","iteration","democracy","protest","avg.att
                         "ncoups","nrevs","reg.age","last.change","init.reg")
  details[,-1] <- factor.to.numeric(details[,-1])
  details$population <- rep(out$outPopulation[1:out$nCountries], nData)

  connMatrix <- t(matrix(out$outConnMatrix[1:(out$nCountries^2)],
                         nrow=out$nCountries, ncol=out$nCountries))

  list(moran=out$outMoran, propDemoc=out$outProportion, attMap=out$outAttMap,
       nCountries=out$nCountries, countryDetails=details, Wstd=connMatrix,
```

```
        iter=1:iter, time.taken=out$outTimeTaken, out=out)
}
```

G.3 Simulation scripts

Parameter sweep

For the parameter sweeps, the following code was used, with an equivalent run to get the longer versions (more iterations):

```
#setwd("~/Desktop/academic/main/model/")
source("import.R")

doubleEqual <- function(x,y,margin=.000001)
    (x - margin) < y & (x + margin) > y

nIterations <- 100000
thinning <- 50
detailThinning <- 50
fileThinning <- 1
nPerSetting <- 100
seriesID <- 12
startRun <- 2420
endRun <- 2300

outMoran <- NULL
propDemoc <- NULL
nCountries <- NULL
attMap <- NULL
Wstd <- list()
countryDetails <- NULL

add.params <- function(params, add)
  cbind(params %x% matrix(1, length(add), 1), add)

params <- c(10) # in-group mean
params <- add.params(params, c(70)) # out-group/in-group difference
params <- add.params(params, c(0,1,5)) # broadcast effect
params <- add.params(params, c(0,.5)) # seq(from=.1, to=.9, by=.4)) # crossB
params <- add.params(params, c(0,1)) # communication effect
```

```
params <- add.params(params, c(0,.0001,.0002)) # random revolution chance
params <- add.params(params, 10) # initial mean attitude
params <- add.params(params, 20) # initial mean standard deviation
params <- add.params(params, .1) # initial proportion democratic
params <- add.params(params, 1:nPerSetting) # run

params

colnames(params) <- c("threshold", "out.add", "broadcast", "crossB", "comm.ef

init.time <- proc.time()[3]

I <- dim(params)[1]
if (endRun == 0)
  endRun <- 1

runs.duration <- abs(endRun - startRun) + 1

for (i in startRun:endRun) {

  runs.proportion.done <- abs(i - startRun) / runs.duration

  if (runs.proportion.done == 0)
    runs.proportion.done <- 1

  cat(sprintf("\nSTARTING RUN %d OF [%d -> %d] (approximately %d minutes left

  batch <- floor(i / fileThinning)

##  cat(sprintf("Memory use report before simulation: %d max obtained - %d in

  res <- simulation(iter=nIterations,
                    broadcast=params[i,"broadcast"],
                    crossB=params[i,"crossB"],
                    comm.effect=params[i,"comm.eff"],
                    attMean=params[i,"att"],
                    attStd=params[i,"att.std"],
                    inGMean=params[i,"threshold"],
                    inGStd=0,
                    outGMean=params[i,"threshold"] + params[i,"out.add"],
                    outGStd=0, propDemoc=params[i,"init.dem"],
```

```
                          detailThin=detailThinning, thin=thinning,
                          batchID=seriesID, runID=i)

##  cat(sprintf("Memory use report after simulation:  %d max obtained - %d in use\n"

  cat("Reached end of simulation - storing data\n")

  outMoran <- cbind(outMoran, res$moran)
  propDemoc <- cbind(propDemoc, res$propDemoc)
  nCountries <- c(nCountries, res$nCountries)
  Wstd <- c(Wstd, res$Wstd)
  attMap <- cbind(attMap, res$attMap) # only works because width, height, and iter
  res$countryDetails$run <- i
  countryDetails <- rbind(countryDetails, res$countryDetails)

  if (i %% fileThinning == 0) {

    save(outMoran, propDemoc, nCountries, attMap, params, countryDetails, Wstd, file

    outMoran <- NULL
    propDemoc <- NULL
    nCountries <- NULL
    Wstd <- list()
    attMap <- NULL
    countryDetails <- NULL
  }
}
```

Monte Carlo parameter settings

For the Monte Carlo version of the simulation runs, the following code was used:

```
#setwd("~/Desktop/academic/main/model/")
source("import.R")

doubleEqual <- function(x,y,margin=.000001)
    (x - margin) < y & (x + margin) > y

nIterations <- 100000
thinning <- 50
```

```
detailThinning <- 50
fileThinning <- 1
nRuns <- 1000
seriesID <- 10

outMoran <- NULL
propDemoc <- NULL
nCountries <- NULL
attMap <- NULL
Wstd <- list()
countryDetails <- NULL

params <- cbind(round(runif(nRuns, 10, 30)), # width
                round(runif(nRuns, 10, 30)), # height
                round(runif(nRuns, 0, 20)), # multiplier
                round(runif(nRuns, 0, 100)), # isomean
                round(runif(nRuns, 1, 30)), # isostd
                runif(nRuns, 0, .5), # initdem
                round(runif(nRuns, 20, 100)), # popmean
                round(runif(nRuns, 1, 40)), # popstd
                runif(nRuns), # crossB
                round(runif(nRuns, 0, 100)), # attmean
                round(runif(nRuns, 1, 50)), # attstd
                round(runif(nRuns, 0, 80)), # ingmean
                round(runif(nRuns, 1, 20)), # ingstd
                round(runif(nRuns, 0, 80)), # outgmean (minus ingmean)
                round(runif(nRuns, 1, 20)), # outgstd
                round(runif(nRuns, 0, 100)), # broadcast
                round(runif(nRuns, 0, 2)), # comm.eff
                runif(nRuns, 0, 1/5000), # chance of coup
                round(runif(nRuns, 50, 100)), # delay
                runif(nRuns, 0, 1/5), # decay start
                runif(nRuns, -1/3, 0)) # decay strength

colnames(params) <- c("width","height","multiplier","iso","iso.std","init.dem

params

init.time <- proc.time()[3]

I <- dim(params)[1]
```

```
for (i in I:1) {

  cat(sprintf("\nSTARTING RUN %d OF %d (approximately %d minutes left):\n\n", i, di

  batch <- floor(i / fileThinning)

##  cat(sprintf("Memory use report before simulation: %d max obtained - %d in use\n

  res <- simulation(iter=nIterations,
                    broadcast=params[i,"broadcast"],
                    crossB=params[i,"crossB"],
                    comm.effect=params[i,"comm.eff"],
                    attMean=params[i,"att"],
                    attStd=params[i,"att.std"],
                    inGMean=params[i,"threshold"],
                    inGStd=params[i,"threshold.std"],
                    outGMean=params[i,"threshold"] + params[i,"out.add"],
                    outGStd=params[i,"out.add.std"],
                    propDemoc=params[i,"init.dem"],
                    width=params[i,"width"],
                    height=params[i,"height"],
                    multiplier=params[i,"multiplier"],
                    isoMean=params[i,"iso"],
                    isoStd=params[i,"iso.std"],
                    popMean=params[i,"pop"],
                    popStd=params[i,"pop.std"],
                    randRev=params[i,"randrev"],
                    regDelay=params[i,"delay"],
                    decayStart=params[i,"decay.start"],
                    decayStrength=params[i,"decay.strength"],
                    detailThin=detailThinning, thin=thinning,
                    batchID=seriesID, runID=i)

##  cat(sprintf("Memory use report after simulation:  %d max obtained - %d in use\n

  cat("Reached end of simulation - storing data\n")

  outMoran <- cbind(outMoran, res$moran)
  propDemoc <- cbind(propDemoc, res$propDemoc)
  nCountries <- c(nCountries, res$nCountries)
  Wstd <- c(Wstd, res$Wstd)
```

```
  attMap <- cbind(attMap, res$attMap) # only works because width, height, and
  res$countryDetails$run <- i
  countryDetails <- rbind(countryDetails, res$countryDetails)

  if (i %% fileThinning == 0) {

    save(outMoran, propDemoc, nCountries, attMap, params, countryDetails, Wst

    outMoran <- NULL
    propDemoc <- NULL
    nCountries <- NULL
    Wstd <- list()
    attMap <- NULL
    countryDetails <- NULL
  }
}
```

G.4 Analysis scripts

convergence.R

```
convergence <- function(x, precision = NULL)
  {
    halfway <- floor(length(x) / 2)
    end <- length(x)
    last.ten <- floor(length(x) / 10) * 9

    if (is.null(precision))
      precision <- diff(range(x)) / 100

    end.value <- mean(x[last.ten:end], na.rm=TRUE)
    n.value <- sum(!is.na(x[last.ten:end]))
    halfway.mean <- mean(x[halfway:end], na.rm=TRUE)
    is.converged <- (halfway.mean - precision) < end.value & end.value < (hal

    list(value = end.value, n = n.value, conv = is.converged)
  }
```

analysis.R

```
source("convergence.R")
```

```
temporal.cor <- function(x, lags=5)
{
    n <- length(x)

    c <- rep(NA, lags)

    for (l in 1:lags)
        c[l] <- cor(x[-(1:l)], x[-c((n-l+1):n)])

    c
}

getCountryState <- function(batch, run, file.thin)
  {
    curr <- (run-1) %/% file.thin + 1
    i <- (run-1) %% file.thin + 1

    load(sprintf("output/batch%d_%d.Rdata", batch, curr))

    countryState <- as.integer(unlist(strsplit(
          readLines(sprintf("output/batch%d_%d_countries.data",
                              batch, run)), fixed=TRUE, split=NULL)))
    countryState <- matrix(countryState, nCountries[i])

    countryState.diff <- countryState[,-1] - countryState[,-dim(countryState)[2]]
    trans.dem <- apply(countryState.diff == 1, 2, sum)
    trans.aut <- apply(countryState.diff == -1, 2, sum)

    list(countryState=countryState, countryState.diff=countryState.diff, trans.dem=t
  }

analyse <- function(batch, run, file.thin)
{
  prev <- 0
  nlags <- 10

  for (r in run)
    {
      cat("Analysing run", r, "\n")
```

```
## Only load files when needed (ie. when crossing the file thinning
## border)
curr <- (r-1) %/% file.thin + 1
if (curr != prev)
  {
    load(sprintf("output/batch%d_%d.Rdata", batch, curr))

    ## Only the first time, load parameter settings
    if (prev == 0)
      {
        par.data <- as.data.frame(params)
        par.data$avg.moran <- NA
        par.data$n.moran <- NA
        par.data$conv.moran <- NA
        par.data$conv.value.moran <- NA
        par.data$conv.n.moran <- NA
        par.data$conv.democ <- NA
        par.data$conv.value.democ <- NA
        par.data$conv.n.democ <- NA
        par.data$ncountries <- NA

        trans.stats <- NULL
      }

    prev <- curr
  }
i <- (r-1) %% file.thin + 1

## Calculate clustering statistics
par.data$avg.moran[r] <- mean(outMoran[,i], na.rm=TRUE)
par.data$n.moran[r] <- sum(!is.na(outMoran[,i]))

## Calculate equilibrium statistics
conv <- convergence(outMoran[,i], .05)
par.data$conv.moran[r] <- conv$conv
par.data$conv.value.moran[r] <- conv$value
par.data$conv.n.moran[r] <- conv$n

conv <- convergence(propDemoc[,i], .05)
par.data$conv.democ[r] <- conv$conv
par.data$conv.value.democ[r] <- conv$value
```

```
      par.data$conv.n.democ[r] <- conv$n

      ## Store number of countries
      par.data$ncountries[r] <- nCountries[i]

      ## Calculate transition statistics
      cs <- getCountryState(batch, curr, file.thin)
      tc.dem <- suppressWarnings(temporal.cor(cs$trans.dem, nlags))
      tc.aut <- suppressWarnings(temporal.cor(cs$trans.aut, nlags))
      trans.stats <- rbind(trans.stats, c(tc.dem, tc.aut))
    }

  colnames(trans.stats) <- rep("", 2*nlags)
  for (i in 1:nlags)
    {
      colnames(trans.stats)[i] <- sprintf("dem.lag.%d", i)
      colnames(trans.stats)[i + nlags] <- sprintf("aut.lag.%d", i)
    }

  cbind(par.data[run,], trans.stats)
}

calculate.autocor <- function(batch, run, file.thin)
{
  prev <- 0

  for (r in run)
    {
      cat("Analysing run", r, "\n")

      if (batch == 14) {
        save(r, file="last.Rdata")
        cat(r, file="last.txt")
      }

      ## Only load files when needed (ie. when crossing the file thinning
      ## border)
      curr <- (r-1) %/% file.thin + 1
      if (curr != prev)
        {
          load(sprintf("output/batch%d_%d.Rdata", batch, curr))
```

```
          ## Only the first time, load parameter settings
          if (prev == 0)
            {
              par.data <- as.data.frame(params)
              par.data$autocor.dem <- NA
              par.data$autocor.aut <- NA
            }

          prev <- curr
        }
      i <- (r-1) %% file.thin + 1

      ## Calculate autocorrelation statistics
      cs <- getCountryState(batch, curr, file.thin)
      dem <- tapply(cs$trans.dem, (1:length(cs$trans.dem) - 1) %/% 100, sum)
      aut <- tapply(cs$trans.aut, (1:length(cs$trans.aut) - 1) %/% 100, sum)
      par.data$autocor.dem[r] <- cor(dem[-1], dem[-length(dem)])
      par.data$autocor.aut[r] <- cor(aut[-1], aut[-length(aut)])
    }

  par.data[run,]
}

calculate.moranI <- function(x, W)
{
    res <- lm(x ~ 1)$residuals
    ## And then ... ? :)
}

calculate.trans.moran <- function(batch, run, prefix="output/")
{
  lag <- function(x)
  {
    rbind(NA, x[-1,])
  }

  for (r in run)
    {
      fn <- sprintf("%sbatch%d_%d.Rdata", prefix, batch, r)
      cat("Loading", fn, "\n")
```

```
    load(fn)

    fn <- sprintf("%sbatch%d_%d_countries.data", prefix, batch, run)
    cat("Loading", fn, "\n")
  countryState <- as.integer(unlist(strsplit(readLines(fn), fixed=TRUE, split=NULL
    countryState <- matrix(countryState, ncol=nCountries)
    cat("Dimensions countryState:", dim(countryState), "\n")

    transCountryState <- rbind(NA, countryState[-1,] - countryState[-nrow(countryS
    blocks <- rep(1:100, each=1000)
    transToDem <- transCountryState == 1
    cat("Dimensions transToDem:", dim(transToDem), "\n")
    transToDem <- apply(transToDem, 2, tapply, blocks, sum)
    cat("Dimensions transToDem:", dim(transToDem), "\n")
    transToAut <- transCountryState == -1
    transToAut <- apply(transToAut, 2, tapply, blocks, sum)
    mode(transToAut) <- "numeric"

    cat("Preparing Wstd\n")
    Wstd <- matrix(Wstd, nrow=nCountries)
    mode(Wstd) <- "numeric"
    cat("Dimensions Wstd:", dim(Wstd), "\n")
  }
}


calculate.trans.moran2 <- function(batch, run, prefix="output/")
{
  lag <- function(x)
  {
   rbind(NA, x[-1,])
  }

  for (r in run)
   {
    fn <- sprintf("%sbatch%d_%d.Rdata", prefix, batch, r)
    cat("Loading", fn, "\n")
    load(fn)

    fn <- sprintf("%sbatch%d_%d_countries.data", prefix, batch, run)
    cat("Loading", fn, "\n")
```

```
    countryState <- as.integer(unlist(strsplit(readLines(fn), fixed=TRUE, spl
      countryState <- matrix(countryState, ncol=nCountries)
      cat("Dimensions countryState:", dim(countryState), "\n")

      cat("Preparing Wstd\n")
      Wstd <- matrix(Wstd, nrow=nCountries)
      mode(Wstd) <- "numeric"
      cat("Dimensions Wstd:", dim(Wstd), "\n")

      cat("Calculating W %*% countryState\n")
      lWy <- as.vector(lag(countryState %*% Wstd))

      cat("Setting up y\n")
      y <- as.vector(countryState)
      y.inv <- 1 - y

      cat("Setting up lag(y)\n")
      ly <- as.vector(lag(countryState))
      ly.inv <- 1 - ly

      cat("Setting up population\n")
      population <- rep(countryDetails$population[countryDetails$iteration ==

      cat("Setting up time variable\n")
      t <- rep(1:nrow(countryState), each=nCountries)

      cat("Estimate model\n")
      m <- glm(y ~ 0 + ly + ly.inv + lWy:ly.inv + lWy:ly + population:ly.inv
           family=binomial(link="logit"), subset=(t %% 1000 == 0))

      cat(table(y[t %% 1000 == 0], ly[t %% 1000 == 0]))
    }

  m
}
```

Results tables: parameter sweep

The following code was used to generate tables 5.2, 5.4, 5.5 and 5.6:

```
source("analysis.R")
```

```
library(arm)

## Step 1: process data files in batches of 100

for (i in 1:36) {
     data <- analyse(13, ((i-1)*100+1):(i*100), 1)
     save(data, file=sprintf("dataKnoppix13_%d.Rdata", i))
}

## Step 2: merge batches into one data file

d2 <- NULL

for (i in 1:36) {
    load(sprintf("dataKnoppix13_%d.Rdata", i))
    d2 <- rbind(d2, data)
}
data <- d2
save(data, file="dataKnoppix13.Rdata")

## Step 3: process data files for autocorrelation measures

for (i in 1:36) {
  data.autocor <- calculate.autocor(13, ((i-1)*100+1):(i*100), 1)
  save(data.autocor, file=sprintf("dataKnoppix13_autocor_%d.Rdata", i))
}

d2 <- NULL
for (i in 1:36) {
  load(sprintf("dataKnoppix13_autocor_%d.Rdata", i))
  d2 <- rbind(d2, data.autocor)
}
data.autocor <- d2
save(data.autocor, file="dataKnoppix13_autocor.Rdata")

## Step 4: actual analysis

load("dataKnoppix13.Rdata")
load("dataKnoppix13_autocor.Rdata")
data$autocor.dem <- data.autocor$autocor.dem
```

```
data$autocor.aut <- data.autocor$autocor.aut
attach(data)

## Table 6.2

tapply(conv.value.democ,
       list(communication=comm.eff, broadcast=broadcast,
            cross.border=crossB),
       mean)
tapply(conv.value.democ,
       list(communication=comm.eff, broadcast=broadcast,
            cross.border=crossB),
       sd)

## Table 6.3

moran.expected <- -1 / (ncountries - 1)
dev.moran <- avg.moran - moran.expected

tapply(dev.moran,
       list(communication=comm.eff, broadcast=broadcast,
            cross.border=crossB),
       mean)
tapply(dev.moran,
       list(communication=comm.eff, broadcast=broadcast,
            cross.border=crossB),
       sd)

## Table 6.4

tapply(autocor.dem,
       list(communication=comm.eff, broadcast=broadcast),
       mean)
tapply(autocor.dem,
       list(communication=comm.eff, broadcast=broadcast),
       sd)

tapply(autocor.dem,
       list(communication=comm.eff, broadcast=broadcast,
            randRev=randRev),
       mean)
```

```
tapply(autocor.dem,
       list(communication=comm.eff, broadcast=broadcast,
            randRev=randRev),
       sd)

## Table 6.5

tapply(autocor.aut,
       list(communication=comm.eff, broadcast=broadcast),
       mean)
tapply(autocor.aut,
       list(communication=comm.eff, broadcast=broadcast),
       sd)
```

Results tables: spatio-temporal clustering

The following C code (Kernighan and Ritchie 1988) was used to generate the regression results presented in Section 5.2.3, which makes use of the LAPACK library (Anderson et al. 1999) for the regression analysis:

```
#include <stdlib.h>
#include <stdio.h>
#include "f2c.h"
/* #include "cblas.h" */
#include "clapack.h"

const int block_size = 10;
const int k = 7;

int main(int argc, char **argv)
{
  FILE *input_W = stdin, *input_data = stdin, *output = stdout;
  int col = 0, row = 0, i, j, p, q, row_sum, ccol, ncountries, id;
  double world_mean_0, world_mean_1, country_mean, country_diff_mean;
  char c;
  unsigned char first_line = 1;
  unsigned char *preprev, *prev, *curr, *tmp;
  double *W;
  double W_first[500]; /* Consequence: can handle maximum of 500x500 W matrix */
  char buf[200];
```

```
unsigned int n = 0;
double XX[k*k]; /* X'X matrix */
double Xy[k];   /* X'y matrix */
double betas[k];
double x[k];    /* one row of data */

if (argc > 1)
  id = atoi(argv[1]);
else
  {
    fputs("ID required\n", stderr);
    exit(1);
  }

if (argc > 2)
  if ((input_W = fopen(argv[2], "r")) == NULL)
    {
  fprintf(stderr, "Could not open W file %s\n", argv[2]);
  return 1;
    }
  else
    fprintf(stderr, "Opened %s\n", argv[2]);

if (argc > 3)
  if ((input_data = fopen(argv[3], "r")) == NULL)
    {
  fprintf(stderr, "Could not open data file %s\n", argv[3]);
  return 1;
    }
  else
    fprintf(stderr, "Opened %s\n", argv[3]);

if (argc > 4)
  if ((output = fopen(argv[4], "w")) == NULL)
    {
  fprintf(stderr, "Could not open output file %s\n", argv[4]);
  return 1;
    }
  else
    fprintf(stderr, "Opened %s\n", argv[4]);
```

```
/* Read W from top of file
   Assumes W is written with a command like
   write.table(matrix(Wstd, nrow=nCountries), file="batch20_5352_W.csv",
               col.names=FALSE, row.names=FALSE)
*/
ncountries = 1000; /* just a high value ... */
col = ccol = row = 0;
while ((row < ncountries) && ((c = fgetc(input_W)) != EOF))
  if (c == '\n')
    {
  /* Before handling the line as a whole, handle the last number before \n */
  col++;
  buf[ccol] = '\0';
  W_first[col] = atof(buf);

  /* If first line, allocate memory for W */
  if (!row)
    W = (double *) malloc(col * col * sizeof(double));

  /* copy W line buffer to W */
  for (i = 0; i < col; i++)
    W[row * col + i] = W_first[i];

  ncountries = col;

  ccol = col = 0;
  row++;
    }
  else if (c == ' ')
    {
  buf[ccol] = '\0';

  W_first[col] = atof(buf);
  ccol = 0;
  col++;
    }
  else
    {
  buf[ccol] = c;
  ccol++;
    }
```

```
fprintf(stderr, "W matrix formed (%d countries); turning to data\n",
    ncountries);

/* fputs("country,year,y,ly,Gly,dGly,Wly,dWly\n", output); */

/* Clear regression matrices */
for (i = 0; i < k; i++)
  {
    betas[i] = 0.0;
    Xy[i] = 0.0;

    for (j = 0; j < k; j++)
  XX[i * k + j] = 0.0;
  }

/* Read data file */
while ((c = fgetc(input_data)) != EOF)
  if (first_line)
    if (c == '\n') /* First line, and end of line reached */
  {
    first_line = 0;
    col = 0;
    row++;
  }
    else /* First line, not yet at the end */
  {
    if (col % block_size == 0)
      {
        if (col) free(curr);
        curr = prev;
        prev = (unsigned char *) malloc((col + 1) * block_size
                        * sizeof(unsigned char));

        for (i = 0; i < col; i++)
      prev[i] = curr[i];

        if (col) free(curr);
        curr = (unsigned char *) malloc((col + 1) * block_size
                        * sizeof(unsigned char));
        if (col) free(preprev);
```

```
        preprev = (unsigned char *) malloc((col + 1) * block_size
                      * sizeof(unsigned char));

        for (i = 0; i < col; i++)
      curr[i] = preprev[i] = prev[i];
      }

    preprev[col] = prev[col] = curr[col] = c - 48;

    col++;
  }
  else
    if (c == '\n') /* Not first line, end of line reached */
  {
    if (row > 1) /* On the second line, we can still not
            look at lagged first difference */
      {
        /* Calculate world mean level of democracy (Gy) */
        row_sum = 0;
        for (i = 0; i < col; i++)
      row_sum += prev[i];

        world_mean_0 = ((double) row_sum) / ((double) (col - 1));
        world_mean_1 = ((double) row_sum - 1) / ((double) (col - 1));

        for (i = 0; i < col; i++)
      {
        /* Calculate Wly */
        country_mean = 0;
        for (j = 0; j < col; j++)
          country_mean += W[i*col+j] * prev[j];

        /* Calculate Wdly */
        country_diff_mean = 0;
        for (j = 0; j < col; j++)
          country_diff_mean += W[i*col+j] * (prev[j] - preprev[j]);

        /* Fill in regression matrices */
        x[0] = 1;  /* constant */
        x[1] = (prev[i] ? 0 : world_mean_0); /* Gy lagged, aut */
        x[2] = (prev[i] ? world_mean_1 : 0); /* Gy lagged, dem */
```

```
        x[3] = (prev[i] ? 0 : country_mean); /* Wy lagged, aut */
        x[4] = (prev[i] ? country_mean : 0); /* Wy lagged, dem */
        x[5] = (prev[i] ? 0 : country_diff_mean); /* Wdy lagged, aut */
        x[6] = (prev[i] ? country_diff_mean : 0); /* Wdy lagged, dem */

        for (p = 0; p < k; p++)
          {
            Xy[p] += x[p] * curr[i];
            for (q = 0; q < k; q++)
          XX[p*k+q] += x[p] * x[q];
          }

        /* Print out data in .CSV format */
        /* fprintf(output, "%d,%d,%d,%d,%f,%f,%f,%f\n", */
        /* country */          /* i+1, */
        /* year */             /* row, */
        /* y */                /* curr[i], */
        /* y lagged */         /* prev[i], */
        /* Gy lagged */        /* (prev[i] ? world_mean_1 : world_mean_0), *
        /* Gy lagged, dem */   /* (prev[i] ? world_mean_1 : 0), */
        /* Wy lagged */        /* country_mean, */
        /* Wy lagged, dem */   /* (prev[i] ? country_mean : 0) */
        /* ); */
      }
      }

    /* Set previous to current line (through rotation) */
    tmp = preprev;
    preprev = prev;
    prev = curr;
    curr = tmp;

    col = 0;
    row++;

    //      if (row > 100) return 2;
  }
    else /* Not first line, not yet end of line */
  {
    curr[col] = c - 48;
    col++;
```

```
  }

/* At the end of the main data loop, calculate regression coefficients
   (Using LAPACK for the calculation of the inverse of the X'X matrix)
*/
doublereal *XXr = (doublereal *) malloc(k * k * sizeof(doublereal));

for (i = 0; i < k; i++)
  for (j = 0; j < k; j++)
    XXr[i*k+j] = (doublereal) XX[i*k+j];

/* Inverse of XX */
integer ls_n = k;  /* Number of columns in matrix to be inverted */
integer ls_m = k; /* Number of rows in matrix to be inverted */
integer ls_lda = k; /* Leading dimension (?) */
integer ls_ipiv[k];  /* Will store pivot indices */
integer ls_info;   /* Will store result code */

/* Decompose matrix */
dgetrf_(&ls_m, &ls_n, XXr, &ls_lda, ls_ipiv, &ls_info);

/* Calculate inverse */
doublereal *WORK = (doublereal *) malloc(k * k * sizeof(doublereal));
dgetri_(&ls_n, XXr, &ls_lda, ls_ipiv, WORK, &ls_n, &ls_info);

/* Print inv(XX) and Xy */
for (i = 0; i < k; i++)
  {
    for (j = 0; j < k; j++)
  fprintf(stderr, "%11.9f ", XXr[i*k+j]);

    fprintf(stderr, "| %10.2f\n", Xy[i]);
  }

/* Calculate betas */
fprintf(output, "%d", id);

for (i = 0; i < k; i++)
  {
    betas[i] = 0.0;
    for (j = 0; j < k; j++)
```

```
    {
      betas[i] += (double) XXr[i*k+j] * Xy[j];
    }

      fprintf(stderr, "Beta %d: %10.2f\n", i, betas[i]);

      fprintf(output, ",%.5f", betas[i]);
    }

  fputs("\n", output);

  return 0;
}
```

Results tables: Monte Carlo parameters

The following code was used to generate tables 5.9 and 5.10:

```
source("analysis.R")
source("display_coefficients.R")

library(arm)

options(scipen=3)

## Step 1: process data files in batches of 100

for (i in 200:200) {
      todo <- ((i-1)*10+1):(i*10)
      todo <- todo[(todo != 1702) & (todo != 2000)]
      data <- analyse(14, todo, 1)
      save(data, file=sprintf("dataKnoppix14_%d.Rdata", i))
}

## Step 2: merge batches into one data file

d2 <- NULL

for (i in 1:200) {
    load(sprintf("dataKnoppix14_%d.Rdata", i))
    d2 <- rbind(d2, data)
```

```
}
data <- d2
save(data, file="dataKnoppix14.Rdata")

## Step 3: process data files for autocorrelation measures

for (i in 1:200) {
  todo <- ((i-1)*10+1):(i*10)
  data.autocor <- calculate.autocor(14, todo, 1)
  save(data.autocor, file=sprintf("dataKnoppix14_autocor_%d.Rdata", i))
}

d2 <- NULL
for (i in 1:200) {
  load(sprintf("dataKnoppix14_autocor_%d.Rdata", i))
  d2 <- rbind(d2, data.autocor)
}
data.autocor <- d2
save(data.autocor, file="dataKnoppix14_autocor.Rdata")

## Step 4: actual analysis

load("dataKnoppix14.Rdata")
load("dataKnoppix14_autocor.Rdata")
data$autocor.dem <- data.autocor$autocor.dem
data$autocor.aut <- data.autocor$autocor.aut
attach(data)


## Monte Carlo equivalent of table based on democracy convergence

m1 <- lm(conv.value.democ ~ broadcast*crossB*comm.eff*randrev)
means <- apply(data, 2, mean)
m2 <- lm(conv.value.democ ~ broadcast*crossB*comm.eff*randrev
                            + I(broadcast > 0))
m3 <- lm(conv.value.democ ~ broadcast + crossB + comm.eff
                            + randrev + I(broadcast > 0))

means <- apply(data, 2, mean)
x.broadcast <- rep(c(0,1,5), each=2)
x.comm.eff <- rep(c(0,1), 3)
```

```
x <- cbind(x.broadcast, x.comm.eff, matrix(means[c("crossB","randrev")],
           ncol=2, nrow=6, byrow=TRUE))
colnames(x) <- c("broadcast","comm.eff","crossB","randrev")
## Table of predicted values given model with all interactions
tapply(predict(m1, as.data.frame(x)), list(x.comm.eff, x.broadcast), mean)
## Table of predicted values given model
##  with all interactions and broadcast > 0
tapply(predict(m2, as.data.frame(x)), list(x.comm.eff, x.broadcast), mean)
## Table of predicted values given model with broadcast > 0
tapply(predict(m3, as.data.frame(x)), list(x.comm.eff, x.broadcast), mean)

mall <- list(m2,m1,m3)
line.names <- c("Constant",
                "Broadcast effect ($B$)",
                "Chance of cross-border communication ($\\tau$)",
                "Communication effect ($\delta$)",
        "Random chance of coups ($K$)",
        "$I(B>0)$",
        "$B \\times \\tau$",
        "$B \\times \\delta$",
        "$\\tau \\times \\delta$",
        "$B \\times K$",
        "$\\tau \\times K$",
        "$\\delta \\times K$",
        "$B \\times \\tau \\times \\delta$",
        "$B \\times \\tau \\times K$",
        "$B \\times \\delta \\times K$",
        "$\\tau \\times \\delta \\times K$",
        "$B \\times \\tau \\times \\delta \\times K$")

display.m(mall, line.names)

s <- sim(m3, 1000)
x <- cbind(1, x[,c(1,3,2,4)], as.integer(x[,"broadcast"] > 0))
p <- x %*% t(s$beta)

tapply(apply(p, 1, mean), list(x.comm.eff, x.broadcast), mean)
tapply(apply(p, 1, sd), list(x.comm.eff, x.broadcast), mean)

## Monte Carlo equivalent of table based on deviations from expected Moran's
```

```
moran.expected <- -1 / (ncountries - 1)
dev.moran <- avg.moran - moran.expected

m1 <- lm(dev.moran ~ broadcast*crossB*comm.eff*randrev)
means <- apply(data, 2, mean)
m2 <- lm(dev.moran ~ broadcast*crossB*comm.eff*randrev + I(broadcast > 0))
m3 <- lm(dev.moran ~ broadcast + crossB + comm.eff + randrev + I(broadcast > 0))

means <- apply(data, 2, mean)
x.broadcast <- rep(c(0,1,5), each=2)
x.comm.eff <- rep(c(0,1), 3)
x <- cbind(x.broadcast, x.comm.eff, matrix(means[c("crossB","randrev")],
           ncol=2, nrow=6, byrow=TRUE))
colnames(x) <- c("broadcast","comm.eff","crossB","randrev")
## Table of predicted values given model with all interactions
tapply(predict(m1, as.data.frame(x)), list(x.comm.eff, x.broadcast), mean)
## Table of predicted values given model
##  with all interactions and broadcast > 0
tapply(predict(m2, as.data.frame(x)), list(x.comm.eff, x.broadcast), mean)
## Table of predicted values given model with broadcast > 0
tapply(predict(m3, as.data.frame(x)), list(x.comm.eff, x.broadcast), mean)

mall <- list(m2,m1,m3)
display.m(mall, line.names)

s <- sim(m3, 1000)
x <- cbind(1, x[,c(1,3,2,4)], as.integer(x[,"broadcast"] > 0))
p <- x %*% t(s$beta)

tapply(apply(p, 1, mean), list(x.comm.eff, x.broadcast), mean)
tapply(apply(p, 1, sd), list(x.comm.eff, x.broadcast), mean)
```

Equilibrium plots

The following code was used to generate Figure 5.1:

```
library(foreign)

setwd("~/Desktop/academic/main/model")

##load("dataKnoppix7.Rdata")
```

```
##unsel <- data$comm.eff == 1 & data$broadcast == 5
##data7 <- cbind(1:(dim(data)[1]), data)
##data7 <- cbind(7, data7[!unsel,])
##names(data7)[1:2] <- c("sim","file")

##load("dataKnoppix8.Rdata")
##data8 <- cbind(1:(dim(data)[1]), data)
##data8 <- cbind(8, data8)
##names(data8)[1:2] <- c("sim","file")

##data <- rbind(data7, data8)

load("dataKnoppix12.Rdata")
data$sim <- 12
data$file <- 1:3600

postscript("../diss_equilibrium_patterns.eps")
par(mfrow=c(2,3))
for (ce in c(0,1))
  for (b in c(0,1,5))
  {
    runs <- which(data$broadcast == b & data$comm.eff == ce)

    d <- NULL

    for (r in runs)
      {
        cnt <- read.csv(sprintf("output/count%d_%d.csv",
                               data$sim[r], data$file[r]),
                        header=FALSE)
        names(cnt) <- c("run","iteration","ndem")
        d <- cbind(d, cnt$ndem / data$ncountries[r])
      }

    cat(b, ce, ":", data$sim[runs], data$file[runs], "\n")

    sel <- rep(c(T,F,F,F), length.out = length(cnt$ndem))

    m <- apply(d[sel,], 1, mean)
    s <- apply(d[sel,], 1, sd)
```

```
    t <- (1:length(sel))[sel]

    plot(m ~ t, type="l", bty="n", ylim=c(0,1),
         ylab="Proportion democratic", xlab="Iteration")
    lines(I(m - 2 * s) ~ t, col="gray")
    lines(I(m + 2 * s) ~ t, col="gray")
  }
dev.off()
```