



TILL DEATH DO US PART
A comparative study
of government instability
in 28 European democracies

Daniel Walther



UMEÅ UNIVERSITY

Till death do us part

A comparative study of government
instability in 28 European democracies

Daniel Walther



Department of Political Science
Umeå 2017

Copyright © Daniel Walther

This work is protected by the Swedish Copyright Legislation (Act 1960:729)

ISBN: 978-91-7601-699-2

ISSN: 0349-0831 Research Report 2017:2

An electronic version of this thesis is available at <http://umu.diva-portal.org/>

Printed by UmU Print Service, Umeå University

Umeå, Sweden, 2017

To my parents, Inga and Bengt

Table of Contents

Table of Contents	iii
List of Figures	v
List of Tables	v
List of Articles	vii
Abstract	ix
Enkel sammanfattning på svenska	xiii

1. Introduction and overview	1
1.1 What is early cabinet termination and how common is it?	5
1.2 A brief history of the field	9
1.2.1 <i>Browne et al, Strøm and the stochastic nature of terminations</i>	12
1.2.2 <i>'The unified approach' and a formalization of critical events</i>	14
1.2.3 <i>Separating different termination types</i>	17
1.2.4 <i>Turning to the east</i>	19
2 Scope and contribution	21
3 Theory	29
3.1 A brief background: why rational choice is popular	29
3.2 Applying rational choice to party behaviour	30
3.3 Later theoretical developments in the study of government stability	33
4 Methodology and data	35
4.1 How to analyse duration data and why the Cox model was chosen	39
4.2 Polls and cabinet stability	43
5 Summary of the articles	47
5.1 Paper 1: Cabinet survival in Central and Eastern Europe – what do we know after 25 years of democracy?	47
5.2 Paper 2: Government instability and the state	49
5.3 Paper 3: How is government termination affected by the state of economy? Payoff structures, type of government and economic changes	52
5.4 Paper 4: Picking the winner(s): Forecasting elections in multiparty systems	55
5.5 Paper 5: The verdict in the polls – how government stability is affected by popular support	58
6 Where do we go from here?	61
7 Bibliography	65

PAPER I
PAPER II
PAPER III
PAPER IV
PAPER V

List of Figures

Figure 1.1. The government life cycle	4
Figure 1.2. Share of early terminations in Europe since WWII	6
Figure 1.3. Relationship between early termination and number of parties	10
Figure 1.4. Share of early terminations for different government types	10
Figure 1.5. Share of early terminations for different institutional contexts	14
Figure 1.6. How the likelihood of different termination types develops over time	18
Figure 2.1. Difference between model estimations and actual government duration	24
Figure 4.1. Comparison between different ways of handling polling data	44
Figure 5.1. Consistency of predictors of cabinet stability in CEE	48
Figure 5.2. Marginal effect of unemployment for different values of state capacity	51
Figure 5.3. Change in economic circumstances for governments that terminate prematurely and ones that do not	53
Figure 5.4. Main results for how economic effects depend on government type	54
Figure 5.5. How the DLM estimates the development of party support in Sweden between 2010 and 2016	56
Figure 5.6. Different ways of measuring popular support and how it impacts on government duration	59

List of Tables

Table 4.1. Countries and government duration in ERDDA	35
---	----

List of Articles

Walther, D (2017) “Cabinet survival in Central and Eastern Europe – what do we know after 25 years of democracy?”. *Under review*.

Walther, D & Bergman, T (2017). “Government instability and the state”. *Under review*

Hellström, J & Walther, D (2017). “How is government termination affected by the state of economy? Payoff structures, type of government and economic changes.” *Accepted for publication in Government & Opposition, January 2017*

Walther, D. (2015). “Picking the winner (s): Forecasting elections in multiparty systems.” *Electoral Studies, 40, 1-13*.

Walther, D (2017). “The verdict in the polls – how government stability is affected by popular support.” *Under review*

Abstract

This thesis is rooted in the research tradition known as coalition politics, where governments, political parties and political institutions are the central focus. The main emphasis here is on government instability and the question of why governments in modern parliamentary democracies often come to an end before the next regular election. In five distinct but interrelated papers, the thesis explores the issue of early government termination and how it is affected by public support, economic developments and the functioning of the state apparatus. The studies included in this thesis generally take a quantitative approach and make use of a dataset that contains 640 governments in 29 European democracies. Their joint goal is to improve our understanding of when early termination happens by introducing and testing new explanatory factors as well as by improving how previously identified factors are modelled.

The first paper focuses on Central and Eastern Europe. It shows that the stability of governments in that region is affected by slightly different factors than those that impact on governments in Western Europe. In particular, ideological factors and political institutions are found to be less important in Central and Eastern Europe while the formal power basis of the government and the country's economic performance matter more. In the second paper, co-authored with Professor Torbjörn Bergman, the state is brought into government stability research. The paper shows that countries with a lower quality of governance and a less efficient public sector have less stable governments. This is mainly because government parties struggle to achieve their policy goals when the state apparatus is inefficient and corrupt.

Paper 3, co-written with Associate Professor Johan Hellström, looks at how different types of governments respond to economic challenges. In particular, this paper demonstrates that the same changes in economic circumstances (e.g. increases in unemployment or inflation) have different effects on cabinet stability depending on which type of government is in charge. Single party governments are better equipped to deal with economic changes, because they are better positioned to devise new policy responses without having to compromise with other parties. Coalition governments, in contrast, become significantly more likely to terminate early when the economy takes a turn for the worse.

Finally, over the course of two papers I first explore new techniques for analysing polling data and then use them to empirically test whether governments sometimes choose termination as a way to cope with bad poll numbers. Most of the existing techniques for pooling polls and forecasting elections were explicitly designed with two party systems in mind. In Paper 4, I test some of these techniques to determine their usefulness in complex, multiparty systems, and I develop some improvements that enable us to take

advantage of more of the information in the data. In the final paper, I combine the two themes of polling and government stability by looking at how changes in government popularity affect the likelihood of premature dissolution. I find that governments, particularly single party governments, do, in fact, use terminations as a strategic response to changes in their popularity among the public. When support is high, governments tend to opportunistically call an early election, whereas they tend to abandon or reshuffle the government when support is low.

Keywords

Government instability, early termination, polling, coalition studies, comparative politics, duration modelling, Europe, cabinet turnover, cabinet dissolution, parliamentary democracies

Acknowledgements

This thesis would not have been completed punctually, or at all, without the assistance and guidance of a large group of people. Foremost among them are my two supervisors, Johan Hellström and Torbjörn Bergman. Johan taught me many important tricks of the trade, offered continuous support and guidance, and also gave me a lot of leeway to pursue my own research interests within the scope of the project. This was a supervising strategy that suited me very well. Johan's insistence that I spend two boiling hot summer months in Ann Arbor, Michigan learning statistics is also something for which I am very grateful. My co-supervisor Torbjörn offered constant support, thoughtful comments on all my work and also harbours an enthusiasm for the craft that is contagious. And Torbjörn and his family's offer to host me in their home for two months when I needed a place to stay was above and beyond the call of duty and something that I am immensely grateful for.

There are also many other people at the political science department in Umeå who deserve my thanks. Magnus, David, Jessika and Ann-Sofi for helping me find my way as a budding academic and for all the lunches at Universum. Max, Abrak, Veronika, and the other PhD students from my cohort for the coffee, the support groups and all the fun times. I am also grateful to the whole department for taking me in, giving me responsibilities early on and for pretending to understand my dialect before I learned to mask my southern roots. All of that made the transition to life in Umeå much easier. While I was working on my thesis, Umeå was privileged to host two international scholars in my field – Thomas Saalfeld and Holger Döring. I had very educational conversations with both, and I especially want to thank Holger for all of his comments on an earlier version of my thesis manuscript.

During the last 18 months, I spent most of my time at the political science department in Lund. I want to thank the people there for providing with me a place to work. This gave me a reason to leave my apartment every day (and an opportunity for some free coffee). I am particularly grateful to Hanna Bäck for all her assistance, Johannes Lindvall for inviting me to his research meetings, and to Per and Carlo for the Italia lunches, the statistics, and the discussions.

Last but not least, my family – Inga, Bengt, Andreas, and Therese – deserve special thanks for all their support and encouragement throughout my life. Andreas should be credited for helping me with my schoolwork and for the personal example he set. My parents helped me tremendously, and from very early on encouraged curiosity, discipline and intellectual ambition. From a family of manual labourers firmly rooted in the working class of Malmö, I feel very fortunate to be looking at a completed PhD thesis. It is safe to say that without their unwavering support that would not have been possible. Finally, a heartfelt thanks to Fredrika, my love, for all her support, emotional as well as practical, that helped me cross the finish line.

Enkel sammanfattning på svenska

Varför avgår regeringar ofta innan nästa ordinarie val? Hur viktigt är egentligen opinionsstöd för regeringspartier när de bestämmer om de ska fortsätta vara med i regeringen och hur kan detta stöd bäst mätas? Den här avhandlingen är uppdelad i fem artiklar som tillsammans undersöker dessa två övergripande ämnen.

Låt oss börja med regeringsinstabilitet. Frågan om varför partier ofta väljer att lämna en regering innan de måste, som de har gjort i nästan 60% av fallen i parlamentariska demokratier i Europa sedan andra världskriget, är teoretiskt intressant. Om partier nu vill vinna val och påverka policy, varför skulle man välja att ge upp makten innan man är tvungen? Att lyckas få en plats i regeringen är sällsynt för de flesta partier, så varför inte fortsätta utnyttja maktinnehavet så länge man kan?

Här tar jag avstamp i koalitionsforskning och tidigare studier av regeringsstabilitet. Huvudargumentet för varför regeringspartier ofta föredrar att lämna regeringen är att partier har fler mål än maktinnehav. De vill påverka policy i enlighet med vad deras väljare föredrar och de vill också vinna röster i kommande val för att försäkra sig om sin egen långsiktiga överlevnad. Frågan om varför regeringar ofta avgår i förtid blir därför en fråga om när partiets andra mål blir så pass starka att det är värt att ge upp sitt nuvarande maktinnehav för att istället försöka uppnå andra målsättningar.

Artiklarna i den här avhandlingen tar avstamp i tidigare fynd men försöker också bryta ny mark genom att introducera nya förklaringsfaktorer och genom att förbättra hur tidigare identifierade förklaringsfaktorer mäts och modelleras. För att göra detta utnyttjas ERDDA, som är ett dataset med information om 640 regeringar i 29 europeiska demokratier. Datasetet innehåller bl.a. information om regeringssituationen, parlamentet, politiska institutioner och ekonomisk utveckling. Statistiska tekniker, så kallad durationsmodellering, används sedan för att undersöka vilka faktorer som är viktigast för att förutsäga hur länge en viss regering sannolikt kommer att överleva.

Avhandlingens forskningsresultat är utspridda över fem artiklar. Varje artikel gör sina egna unika bidrag och de viktigaste fynden presenteras kort nedan. De tre första artiklarna fokuserar alla på regeringsstabilitet men har olika substantiella intressen. I den första artikeln är fokus på Östeuropa. Startpunkten för artikeln är det faktum att i många tidigare studier av regeringsstabilitet i Östeuropa har fynden tenderat att variera. Det vill säga, många av de faktorer som visade sig vara viktiga för stabiliteten i en studie var det inte i de andra. Huvudmålsättningen i den här artikeln är därför att undersöka om det går att identifiera några nyckelfaktorer som har en mer konsekvent inverkan på regeringsstabiliteten i Östeuropa och som tenderar att ha en effekt oavsett hur regeringens övriga situation ser ut.

Resultaten visar att det finns en kärna av faktorer som konsekvent påverkar regeringsstabiliteten i Östeuropa. Dessa faktorer handlar framförallt om hur stark regering är (alltså vilket stöd den har i parlamentet), hur många partier som tävlar om makten, och hur det går för ekonomin (framförallt utvecklingen i arbetslöshet). Regeringar som sitter i en svag parlamentarisk ställning, har många konkurrenter om makten och där ekonomin går tungt verkar mer benägna att avgå i förväg oavsett vilka andra omständigheter regeringen befinner sig i. Detta står i bjärt kontrast till hur det är i Västeuropa där även politisk ideologi och politiska institutioner verkar spela en stark roll för hur stabil regeringen är.

Artikel nummer två, samförfattad med professor Torbjörn Bergman, drar inspiration från ett annat håll. Här ligger fokus på hur statsapparaten påverkar regeringens varaktighet. Frågan om hur den statliga byråkratin och korruption påverkar regeringsstabilitet har tidigare inte undersökts i fältet men det finns teoretiska skäl att förvänta sig att om regeringen inte har en välfungerande statlig sektor att falla tillbaka på borde det vara svårare att bedriva framgångsrik regeringspolitik. Detta kan i sin tur påverka risken för att avgå i förtid. Resultaten visar att kvaliteten på statsapparaten har både direkta och indirekta effekter. Sannolikheten för tidig avgång påverkas direkt genom att hög korruption och en ineffektiv statsapparat hindrar regeringspartierna från att uppnå sina mål. Men det finns också en indirekt effekt, eftersom statsapparaten kapacitet påverkar hur väl regeringen kan hantera socio-ekonomiska problem som uppstår under mandatperioden. Det är svårare att framgångsrikt introducera ny policy som motverkar ekonomiska problem om man inte kan förlita sig på tjänstemännen i den offentliga byråkratin. Det betyder att regeringar i länder med låg styrningskvalitet också blir mer känsliga för externa kriser.

Den tredje studien, den här gången samförfattad med docent Johan Hellström, tar avstamp i etablerade fynd inom fältet som visar att ogynnsam ekonomisk utveckling, såsom stigande inflation och arbetslöshet, har en negativ effekt på regeringsstabiliteten. Vårt huvudargument här är att ekonomisk utveckling inte borde ha en statisk effekt som är samma för alla typer av regeringar. Istället borde regeringens förmåga att hantera situationen bero på hur den är sammansatt och hur lätt det är för den att introducera ny policy för att ta sig an problemen. Våra resultat här visar att så mycket riktigt är fallet. Enpartiregeringar visar sig inte påverkas av negativ ekonomisk utveckling i vår analys, medan koalitionsregeringar blir mindre stabila i takt med att ekonomin försämras. Detta eftersom det är svårare för koalitionsregeringar att komma överens om ny policy för att hantera de ekonomiska problemen. De här resultaten tyder på att det generella fyndet i fältet att ekonomi spelar roll för stabiliteten i själva verket behöver kvalificeras något, eftersom resultaten helt drivs av koalitionsregeringar.

Slutligen, artikel 4 och 5 fokuserar båda på opinionsmätningar. Målsättningen, och fokus för att artikel 5, är att granska hur sannolikheten för tidig avgång påverkas av regeringens opinionssiffror, men för att göra det behövde artikel 4 göra en del förarbete genom att testa olika tekniker för att mäta opinionsstöd i komplexa flerpartisystem. Opinionsforskningen har gått starkt framåt de senaste decennierna, men de flesta studier fokuserar på majoritetsvals-system där antalet partier ofta är lågt. I parlamentariska system med proportionell tilldelning av platser finns det ofta många fler aktiva partier vilket ställer högre krav på mätningarna. Artikel 4 diskuterar och testar därför olika etablerade tekniker för att göra valprognoser i flerpartisystem. Två huvudslutsatser dras. Den första är att det fungerar bra att göra tidsserieanalys på opinionsdata också i flerpartisystem och att prognoser med bra träffsäkerhet (mindre än 0.8 procentenheter fel per parti) kan göras även där. Den andra slutsatsen är att existerande tekniker kan förbättras genom att börja mäta cykliska trender i tidsserierna eftersom stödet tenderar att utvecklas i förutsägbara vågor. Inkorporerar man sådana trender i modellerna kan man förutsäga svängningar innan de faktiskt inträffar och därmed göra bättre prediktioner på ett tidigt stadium.

Dessa fynd låg sedan till grund för artikel 5 där jag återanvände modellen för opinionsstöd men nu för att undersöka hur regeringsduration påverkas av svängningar i opinionsstöd. Tidigare inom fältet har många lagt fram det teoretiska argumentet att opinionsstöd borde spela roll för regeringens benägenhet att avgå, men hittills har detta inte testats empiriskt. Resultaten i den här studien visar att popularitetssiffror mycket riktigt påverkar stabilitet och att effekten går i två riktningar. När regeringen är mer populär än genomsnittliga regeringar ökar sannolikheten för att regeringen utlyser ett nyval. Detta kan ses som en sorts opportunism där man försöker utnyttja goda tider medan de varar. Å andra sidan, när stödet istället minskar, ser vi en ökad sannolikhet för att regeringen ombildas eller ersätts men utan att ett val hålls. Opinionsstöd verkar därför viktigt för regeringar eftersom sannolikheten att avgå påverkas oavsett vilken riktning stödet utvecklas i så länge avvikelserna från genomsnittliga regeringar är stark nog. Båda dessa fynd gäller framförallt för enpartiregeringar där möjligheten att agera baserat på opinionssiffrorna är större.

Sammanfattningsvis kan man därför säga att de olika artiklarna som ingår i den här avhandlingen gör sina egna unika bidrag men att de också lämnar ett gemensamt bidrag till fältet som helhet. I flera av artiklarna försöker jag förbättra mått och metoder från tidigare forskning, mer direkt mäta de kausala processerna eller introducera nya förklaringsfaktorer som inte tidigare testats. Det är min förhoppning och övertygelse att vi nu vet lite mer om när regeringar faktiskt upplever det som sitt bästa alternativ att avgå i förtid.

1. Introduction and overview

I'm not afraid of death, I just don't want to be there when it happens.

Woody Allen

It is not just beings of flesh and blood that one day cease to exist, the same fate befalls governments. The main objective of this thesis is to explain why this often happens sooner than expected.

Both the life and the death of governments are of crucial concern in parliamentary democracies. While governments are in office, they have the ability to make essential decisions about the political future of the country. They try to implement their preferred policies and they make decisions that will influence the country in years to come. Governments are also ultimately responsible for the workings of the civil service and for ensuring that the public sector operates smoothly (Andersen, Møller, and Skaaning 2014).

Sooner or later, though, the government's time in office will come to an end, and the voters will hold the party or parties that made up the government accountable for their actions. This opportunity for voters to give their verdict on the government's job performance is a defining feature of democratic political systems (Manin, Przeworski, and Stokes 1999). By rewarding government parties that they like and punishing those that have failed to live up to expectations, voters can signal which political direction they want for their country (Gallagher, Laver, and Mair 2011). This assessment of the performance of government parties while in office helps ensure that an accountability mechanism is built into the system of parliamentary democracy.

In many cases, though, the ability of voters to hold incumbents to account in elections is circumvented because the government does not survive until the next regular election. When the government is partly or completely replaced before voters get the opportunity to communicate their views about its performance, i.e. when a replacement government is formed, the chain of responsibility becomes far less clear (Laver and Schofield 1990: Ch 6). It is difficult for the electorate to deduce whether a particular party is responsible for certain political developments if the parties in charge of that policy area have changed between elections. This makes the choice of whom to reward and whom to punish far more complicated. For this reason, government instability goes to the heart of the accountability mechanism that defines democratic political systems.

High government turnover also has more immediate socio-economic effects. In particular, a wide range of studies have found that political instability can hamper economic development by introducing uncertainty into the system (Alesina, Özler, Roubini, and Swagel 1996; Bernal-Verdugo,

Furceri, and Guillaume 2013; Roe and Siegel 2011). Frequent changes in political leadership can lead to a lack of predictability in the system, which tends to put a strain on economic activity. Uncertainty about the future rules of the game due to political instability has been linked to lower investment, reduced economic growth and inflation (Alesina, Özler, Roubini, and Swagel 1996; Cukierman, Edwards, and Tabellini 1989). This suggests that government stability not only has normative implications for the quality of democracy, but also more material economic effects.

The case of Italy can serve as an example of this. Italian governments are the least stable in Europe, and since 1945 more than 80% of governments have left office prematurely (Andersson, Bergman, and Ersson 2014). The frequent change in governing parties in the middle of a constitutional term of office has often led to difficulties in establishing and implementing a coherent, long-term policy agenda (Alesina, Özler, Roubini, and Swagel 1996; Allum 1974). In other countries, such as Sweden and Luxembourg, early terminations are rare occurrences, which has helped make government policy significantly more predictable.

Early government termination is thus socio-politically relevant, but it is also a theoretical puzzle. Holding office and influencing policy are the stated goals of most political parties. To relinquish the opportunity to achieve both before one is constitutionally mandated to do so might seem like a strange choice. Winning elections and being invited to be part of a government are rare events for most parties, so why would they leave office before they are required to? Odd though it may seem, that is, in fact, the most common way for governments in Europe to end. Since the Second World War, close to 60% of democratically elected governments have ended before the next election (Andersson, Bergman, and Ersson 2014). Analysing why, in many cases, premature dissolution is the preferred option of the parties in government is the main topic of this thesis.

The second key theme pursued here is the question of political polling and government popularity. How well can we predict the eventual election results for government parties through the use of polls, and how does their standing in the polls influence their willingness to remain in office? We know from previous studies that the results of political polls have tangible effects on the behaviour of both government and opposition parties. Parties rely on the information conveyed in polls to figure out what voters are interested in and to make adjustments that can improve their results in the next election (Van der Eijk and Franklin 2009, 146)¹. In polls, voters can say what they think of both the various cabinet parties as well as opposition parties, thereby

¹ The question of how retrospective (i.e. looking at past results) and prospective (i.e. looking at what the parties can do in the future if elected) voters are is complicated, and the exact trade-off changes depending on the party and the context (Wlezién, 1995; Van der Eijk and Franklin, 2009: 172)

expressing how they rate the actions of the parties so far and what they want for the future. Based on what people express through polls, parties might develop new policies, discard old ones, or even, as Paper 5 explores, leave or dissolve a government. Thus, government popularity and the propensity to terminate prematurely are related. A second ambition in this thesis is therefore to develop new and improved techniques for analysing polling data in multiparty systems. This will enable us to improve our understanding of how popularity affects cabinet stability.

This thesis is comprised of five distinct papers that deal with these topics in different ways. Their contributions will be discussed at length in Section 2, but some of the *main contributions and findings* can be briefly foreshadowed here. First, using a relatively new statistical technique, it is found that the durability of governments in Central and Eastern Europe is affected by slightly different factors than those that impact on governments in Western Europe. In particular, ideological factors, e.g. how closely aligned to each other the parties in government are, appear to be of less significance in Eastern Europe. On the other hand, factors related to formal power, such as whether the government controls a majority of seats in the legislature and the number of competing parties, matter more.

Second, the state is brought into government stability research. In a joint paper with Professor Torbjörn Bergman, it is shown that countries with lower quality of governance and a less efficient public sector have less stable governments. This is partly because government parties struggle to achieve their goals when they can not rely on a well functioning public sector and partly because they find it more difficult to deal with new policy challenges, such as economic troubles, when the state apparatus is inefficient and corrupt.

Paper 3, co-written with Associate Professor Johan Hellström, looks at how different types of governments (single party governments and coalition governments – both majority and minority) respond to economic challenges. In particular, this paper demonstrates that the same changes in economic circumstances (e.g. increases in unemployment or inflation) have different effects on cabinet stability depending on which type of government is in charge. Single party governments are better equipped to deal with economic changes, because they are better positioned to devise new policy responses without having to compromise with other parties. This suggests that although economic deterioration is often found to negatively impact on government duration (Saalfeld 2008, 2013; Warwick 1994), these results tend to be driven by coalition governments, while single party governments appear unfazed.

Finally, over the course of two papers I first explore new techniques for analysing polling data and then use them to empirically test whether governments sometimes choose termination as a way to cope with bad poll numbers. Most of the existing techniques for pooling polls and forecasting elections were explicitly designed with two party systems in mind. In Paper 4

I test some of these techniques to determine their usefulness in complex, multiparty systems, and I develop some improvements that enable us to take advantage of more of the information in the data. In the final paper, I combine the two themes by looking at how changes in government popularity affect the likelihood of premature dissolution. I find that governments, particularly single party governments, do, in fact, use terminations as a strategic response to changes in their popularity among the public. When support is high, governments tend to opportunistically call an early election whereas they tend to abandon or reshuffle the government when support is low.

The main questions that have shaped this thesis, such as government stability, popular support and party behaviour, are a subset of the larger field of study known as coalition politics (Müller and Strøm 2003). Although the name suggests that only governments, or even just coalition governments, are the main topic of investigation, the field is actually concerned with most of the formal features of the parliamentary democratic system. This includes questions such as: Which parties get into government? How do parties distribute cabinet portfolios? How important is ideology for understanding government policy? How much do institutions and constitutional provisions matter? What is the impact of external events and developments?

Figure 1.1. The government life cycle

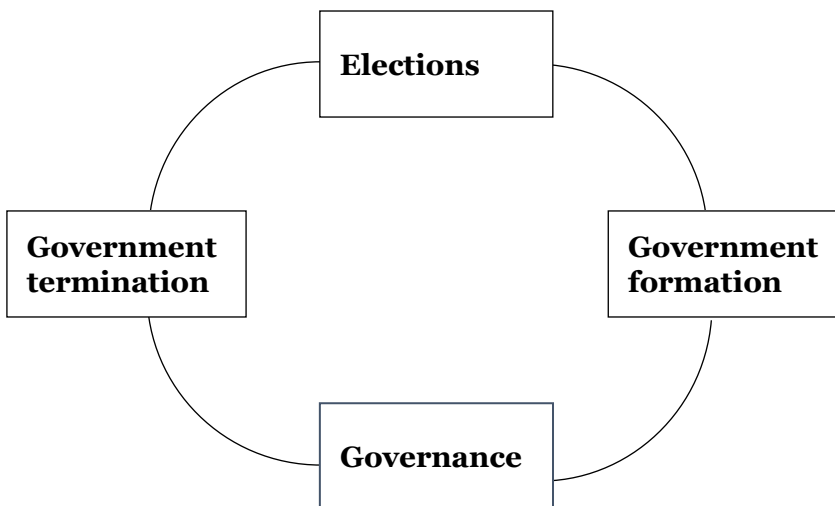


Figure adapted from (Strøm, Müller, and Bergman 2008)

In one of their seminal books on the topic, (Strøm, Müller, and Bergman 2008) presented a simplified but useful overview of the field's main research areas in what they termed the coalition life cycle. This is reproduced in Figure

1.1. The life cycle consists of four steps: parliamentary elections, formation, governance and termination. However, the steps are not necessarily sequential and not every government goes through all of them. For example, a government that comes to office after the previous one has terminated prematurely can go straight to formation without an intermediary election.

These 4 steps are distinct in the sense that they have different formal and informal features and the actors studied here – governments and parties – have different ways of pursuing their interests in the different steps. At the same time, though, the different stages of the government life cycle are closely connected. This connection runs in two directions. First, each preceding step influences the next. The government that is formed depends on the outcome of the election, and the ability of the government to pursue its policy priorities is tied to the kind of government that was formed. Simultaneously, however, the shadow of the future always plays a role. Fearing early termination, parties may choose not to take part in a government they believe will fail. And the knowledge that elections await them invariably informs and constrains the actions governments undertake while in power.

So while the steps can and should be analytically separated, their interconnectedness means that irrespective of what we focus on in a particular study, the full life cycle needs to be taken seriously. Therefore, the key questions of concern in this thesis – termination, elections and developments while in office – will be firmly grounded in the larger context established by the complete life cycle.

1.1 What is early cabinet termination and how common is it?

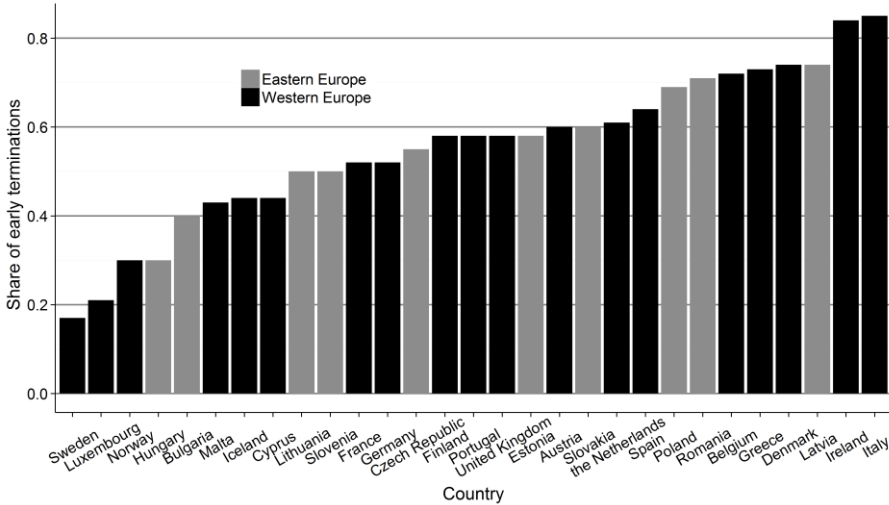
If you are from a country with high political stability and low government turnover, the frequency of early termination might surprise you. More than half of all democratically elected governments in Europe, 59% to be precise, have ended prematurely in the last 70 years (Andersson, Bergman, and Ersson 2014). However, there are large differences among the countries. Some, such as Sweden, Luxembourg and Norway, only rarely see a government terminate in any other way than through regular elections. For others, such as Italy, Latvia and Ireland, premature dissolution is the rule rather than the exception. In Italy, more than 80% of cabinets leave office before the next regular election, which suggests that a newly formed government faces dire odds irrespective of its other characteristics. Contrast this with Sweden, where only exceptional circumstances seem to precipitate a downfall. In fact, as the concept is defined here, the country has at the time of writing not seen a single discretionary early termination for 35 years. A graphical overview of the

frequency of early termination in different European countries can be seen in Figure 1.2.²

When faced with such pronounced differences, a common mistake is to assume that governments and countries have such distinct characteristics or are faced with such specific events that we cannot draw comparative conclusions. In fact, numerous researchers have claimed that the general rules of coalition politics do not apply in their particular country of interest (see Laver and Schofield (1990) for a discussion of this). Due to the country’s unique history, cultural context or particular party leaders, we are told to look to country-based and case-specific reasons for terminations rather than to use insights from comparative studies.

We know, however, that general, cross-country factors explain a large share of the variance across countries (Strøm, Browne, Frensdreis, and Glieber 1988). Italian governments are less stable than those of other countries mainly because the country also has a number of features that we know consistently affect cabinet durability negatively. Italy has more minority and surplus majority governments than the European average, a bicameral system, a large number of parties competing for power, high electoral volatility and large ideological divisions. Once we have adjusted for such factors, much of the

Figure 1.2 Share of early terminations in Europe since WWII



Data mainly from Andersson et al. (2014), but it was updated and extended for this thesis

² Countries such as Denmark and the UK have a fairly high level of government turnover, but it is often strategic opportunism (calling an early election when times are good) (Schleiter and Tavits 2016) rather than conflict and instability that brings down the government

propensity for Italian governments to terminate early can be explained, albeit not all of it.

A related problem is that the media often depicts cabinet terminations as mono-causal and as stemming from the last major event that preceded the government downfall. If high unemployment leads to public protests and government resignation, public disaffection with the state of the economy is said to have caused the downfall. But from comparative studies we know that a government with other characteristics (say, majority instead of minority) in a different political environment (e.g. with fewer other parties competing for power) and with other institutional provisions (e.g. a unicameral system or a demand for an investiture vote) might have been able to weather the storm. The public protests might be the straw that breaks the camel's back, but failing to pay sufficient attention to the underlying factors is like focusing only on the straws while ignoring the camel.

So while there undoubtedly are unique events and specific country characteristics that are hard to capture through quantitative comparative studies, many of the factors that explain the consistent cross-national differences are known and can be empirically modelled. Or expressed differently, we can never predict the future *duration* of one particular government, but we can model its *durability* (Laver 2003; Robertson 1983). Duration in this case refers to the absolute length of time that the government survives, for example 650 days. In that sense, duration is an empirical measure of how long the government managed to stay in office. Durability, in contrast, is the probable duration of the government – the general ability that it has to survive. There are many factors that influence the general durability of the government and these can be used to estimate the *average duration* of a government with particular characteristics. Since the characteristics that make up durability differ systematically among European countries, it is possible to account for some of the differences in survival patterns.

Even a government that has high durability can, of course, be brought down at any time by stochastic events such as a corruption scandal or sharp economic downturn. But the explanatory factors identified in the field can help us predict how likely it is that a government will be able to withstand such eventualities. Since many of the factors that bolster stability are already known when the government takes office, its durability, and thus the likelihood that it will terminate prematurely, can be estimated even if we do not know what its actual duration will be. In a stochastic world, knowledge of general tendencies and average odds are the best we can hope for.

Of course, in order to measure the durability of different types of governments in different contextual conditions, we must first define what constitutes a termination. This is actually more difficult and subjective than one might assume (Browne, Frendreis, and Gleiber 1986; see Frendreis, Gleiber, and Browne 1986; Grofman and Van Roozendaal 1997 for

discussions). A popular definition is the one provided by Browne, Frendreis & Gleiber (1984), which says that a government has terminated when one of the following four events occur:

1. An election
2. The prime minister is replaced
3. The partisan composition of the government changes
4. The government resigns

Thus, whatever party or coalition of parties that comes to power after the occurrence of at least one of these events is defined as a new government³. Reasonable though these criteria might seem, it is not always clear that the implications are suitable for the study of government stability. For example, the PM has a unique position with a decision-making mandate that is crucial to how the government operates. But if the PM dies and is replaced by another senior member from the same party, it is not obvious that this is an example of the type of government instability that we are interested in here. We cannot turn to explanatory variables to help us predict government termination due to the premature death of the PM, and we cannot draw any general lessons that would be applicable to other governments.

Similarly, it is not immediately obvious that an election should always constitute a termination. If the same parties return to power to form a new government and they retain the same PM, one could argue that the government is similar enough to be considered the same. On the other hand, a counter argument to this is that even if the composition of the cabinet parties remains the same, the seat distribution in parliament has likely changed in ways that affect both the government and the opposition. This will influence the political bargaining environment, which suggests that it might be analytically relevant to treat the government as a new one.

These types of problems can be partly overcome by separating so called ‘discretionary’ from ‘technical’ terminations (Laver, 2003). Retaining the criteria for what constitutes a government termination as outlined by Frendreis et al., but further subdividing the cases into those that can be prevented by the government (such as a voluntary early election) and those that cannot (such as a constitutionally mandated regular election), allows us to develop models to predict discretionary, deliberate terminations. The statistical techniques for doing this and a more detailed explanation of how I have coded the different types of terminations will be discussed in chapter 4.

³ In the ERDDA dataset (Andersson et al., 2014) used here, only the first three criteria are used to define a new cabinet. The reason for this is that a resignation is sometimes not accepted or simply leads to the same government coming back into power. Thus, in this thesis, if none of the other three factors changes simultaneously, a resignation does not count as a new government.

However, it should be noted that some of the distinctions are bound to be inherently subjective, which makes transparency in the process imperative.

1.2 A brief history of the field

Thus, durable governments can be defined as those that are better able to avoid early discretionary termination. Given this definition, what established findings are there in the field of cabinet duration studies? What are the factors and attributes that serve to render governments more or less stable, and how have the studies in this thesis been informed by existing research?

The first empirical studies of comparative cabinet stability date back to the 1970s (De Swaan and Rapoport 1973; Dodd 1976; Sanders and Herman 1977; Taylor and Herman 1971; Warwick 1979). Researchers also studied democratic stability more generally, for example focusing on larger questions such as the stability of the whole public sector or even the country as such (Hurwitz 1972; Lijphart 1968). Some work looked in more depth at individual countries (Damgaard 2008; Eckstein 1966). Nonetheless, a distinct subset of studies focused more directly on government stability from a comparative perspective. These were inspired by theoretical literature on coalition politics and party behaviour, but found their own niche by taking a more empirical approach. In practice this often meant doing statistical analyses of how democratic governments in Europe tended to behave, using hypotheses and theoretical expectations based on insights from rational-choice inspired models of how government parties should act.

Taylor & Herman (1971) broke new ground when they produced the paper that set this new subfield in motion. Their main focus was on how party system characteristics and government attributes correlate with average government longevity. They used a set of 196 mainly European cabinets to explore how factors such as number of parties in government, number of parties in opposition, fractionalisation (Rae 1971) and majority or minority status correlate with time spent in office. Many of the explanatory variables invoked by Taylor and Herman have stood the test of time and are still regularly included in empirical models today.

But even if some of their insights remain solid, Taylor & Herman's general approach is unlikely to win the approval of modern journal reviewers. One problem is that they measure cabinet duration simply as the number of days the government is in office. This does not separate technical from discretionary terminations, nor does it take into account institutional regulations, e.g. the maximum length of the term of office. A newly elected British government in the 1970s could sit for five years, but a Swedish government only for three. This leads to pronounced differences in the number of days that governments can potentially remain in power, but not

Figure 1.3. Relationship between early termination and number of parties

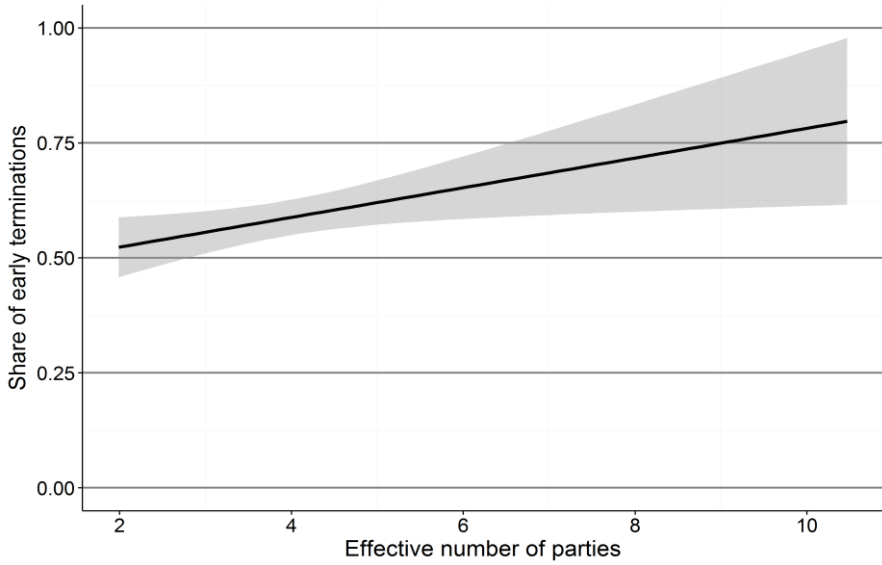
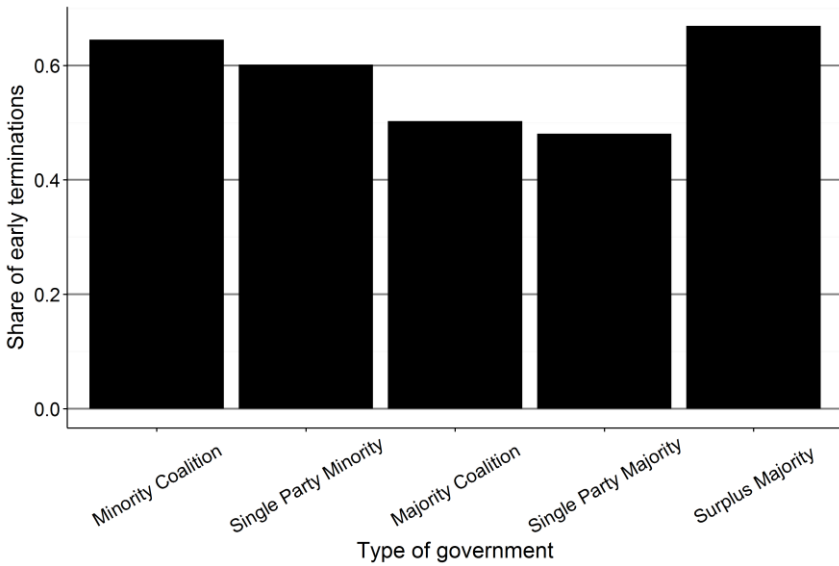


Figure 1.4. Share of early terminations for different government types



Graphs made from data in Andersson et al. (2014)

necessarily to differences in the actual durability of the governments. Another problem was that they relied exclusively on bivariate correlations to empirically test their hypotheses. Each potential explanatory factor was run in a standard linear (Pearson) correlation together with the number of days that government was in office. The correlation coefficient (r) and squared correlation coefficient (r^2) were calculated, which enabled them to measure how much of the variation in government duration was picked up by each factor. Today, simple correlations are generally considered insufficient because they do not control for other relevant factors; they require you to run repeated tests, and the risk of spurious results increases dramatically.

Despite these concerns, many of Taylor & Herman's findings remain descriptively true today. Evidence for two of their hypotheses can be seen in Figures 1.3 and 1.4. First, in Figure 1.3, we see the relationship between the effective number of parties in a system and the frequency of termination. Effective number of parties is defined as the number of parties represented in parliament weighted by their seat share⁴. Weighting for seat share also captures how disbursed power is, since the bargaining environment is more complex when there are many players of approximately equal size. The graph shows that when the effective number of parties increases from low to high, the likelihood of early termination increases from around 50% to over 75%. The most stable systems, on average, appear to be the two party ones (such as the UK, until recently), where the bargaining environment is highly predictable.

Similarly, in Figure 1.4, we see the share of discretionary early terminations for different types of governments. Here it is clear that the least stable governments, i.e. those with the greatest share of premature dissolution, are minority cabinets and cabinets composed of a coalition of parties. Single party minority governments are more stable than minority coalitions, and surplus majority coalitions (i.e. those that contain more parties than are technically needed for a majority) are far less stable than smaller or single party majority cabinets. Size and number of cabinet parties thus seem to matter for stability.

Sanders & Herman made progress on many of the issues in Taylor and Herman's study in their 1977 paper. They defined 'survival' as duration in days divided by total possible duration in order to get the relative rather than absolute longevity of the government. This means that a British government surviving for 900 days in 1977 would get a survival rate of 0.49 (900/1825), whereas a Swedish government would get a survival rate of 0.82 (900/1095). Measuring survival like this solves the problem that the length of term of office varies across countries. Sanders and Herman also argued that a government

⁴ Mathematically this can be done using $ENP = \frac{1}{\sum_{i=1}^N p_i^2}$, where N is the number of parties and P is each party's parliamentary representation squared (Laakso and Taagepera, 1979).

ends whenever a regular election is held, even if the same parties and PM assume office after the election, because party system fractionalisation and bargaining power will change. Warwick (1979) developed this further, and introduced more precise measures of ideological divisions within coalitions. Since governments that are ideologically heterogeneous have more difficulty agreeing on policy, and also find it harder to adopt new policies in light of changed circumstances, they should also be more likely to end prematurely. Warwick used linear regression (OLS) models in an important attempt to move beyond the bivariate correlations that can so easily lead to spurious findings. Thus, by the late 1970s, important cornerstones of the field as we now know it had already been laid.

1.2.1 *Browne et al, Strøm and the stochastic nature of terminations*

The 1980s were dominated by what Warwick memorably termed the ‘survival debates’ (Warwick, 1994). The debates were initialized by Browne, Frendreis, and Gleiber in a series of papers (Browne, Frendreis, and Gleiber 1984, 1986; Frendreis, Gleiber, and Browne 1986) (see also Browne and Dreijmanis 1982). The Browne et al. team made a significant contribution to the field in general, and they had one core argument that was aimed at its very foundation.

They argued that government terminations are largely stochastic and driven by external events. That is, all governments run a certain risk of ending prematurely, and whether this fate befalls a particular government is a result of chance events such as wars, economic crises, corruption scandals or personal conflicts.

This approach was wildly different from how the field had progressed in the 1970s. In that tradition, government stability was modelled as a direct function of the attributes of the government, parliament and general party system. This meant that we could look at the characteristics of a certain government and its political environment at the time that it took office and make a prediction about how it would fare in terms of stability compared to the average government in the sample. For example, according to the early research findings, if a particular government enjoyed a majority position, it would, on average, be more durable than if it had only minority support in parliament. Browne et al. argued that such an approach was likely to be misguided. If governments are not brought down by their inherent attributes, but by chance events whose occurrence cannot be predicted ahead of time, then directly modelling their probable duration is conceptually problematic. Instead, the question is how many potentially terminal events and crises a government will experience and how well it will be able to weather them.

One implication of this theoretical argument is that the timing of terminations should be randomly distributed throughout the governmental life cycle. If terminations are caused by external crises, and crises are equally

likely to happen at any time during the cabinet's term of office, then a termination at day 25 should be as likely as a termination at day 950 (Browne et al., 1986, 636). Browne et al. used a comparison with the Poisson probability distribution to model this. The Poisson distribution is often used to model the probability that a certain phenomenon happens within a given time span (e.g. how many cars pass a certain bridge per hour and how long do I have to wait, on average, for the next one?). This distribution provides information about the likely number of events within the specified time span, but there is no way of knowing when the events will happen within that time span.

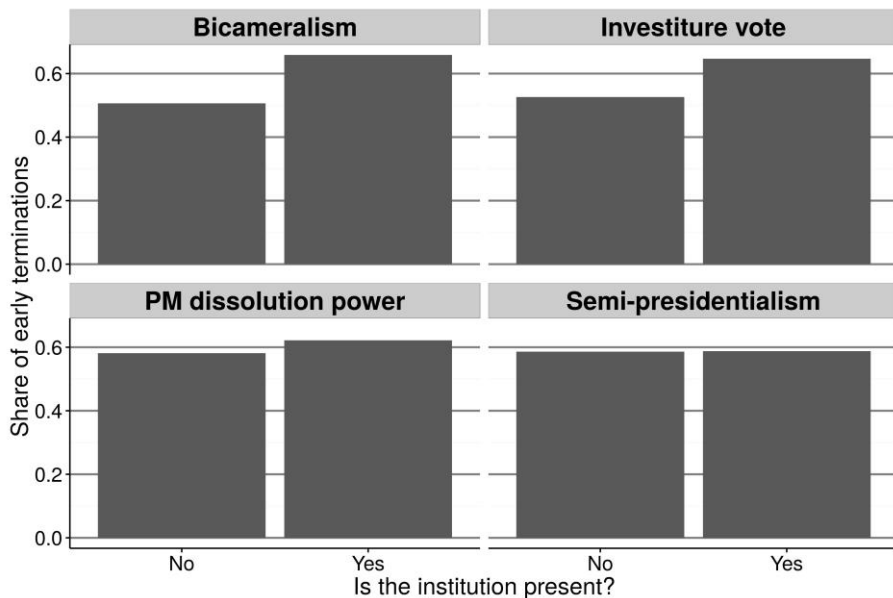
Browne et al.'s key argument is that the critical events that bring down governments follow such a pattern. We know that certain countries or types of government are more prone to termination (i.e. have a higher mean rate), but we generally do not know when, if at all, during the government's term of office it will fail. The precise timing is therefore driven by other, external events. Browne et al. then argued that in some countries this stream of critical events seems to follow the essentially stochastic process suggested by the Poisson distribution.

This controversial standpoint immediately received far-reaching criticism, most notably from Strøm (see, e.g., Strøm et al. (1988)). He argued that 8 of the 12 countries investigated by Browne et al. had a dissolution pattern that differed from the one predicted by the Poisson distribution. In addition, Strøm argued that Brown et al. did not properly define 'critical events', nor did they collect data on all possible critical events in order to determine when they took place. Rather, they inferred the distribution from the timing of all terminations. This is problematic because it only measures 'successful' critical events, not all events. Finally, Strøm noted that Browne et al. looked at all terminations, including those stemming from the death of the PM or other constitutionally mandated reasons, rather than only discretionary terminations. Strøm (1988) summarized his conclusions bluntly, arguing that:

The stochastic models reviewed are both ill-conceived and unilluminating. There is no convincing micropolitical connection between critical events and cabinet dissolutions. No theoretical definition of the dependent variable is provided, and the operationalization is ill suited to the purported theoretical concerns. The independent variable is neither defined nor observed.

There were other developments in field that took place outside of the 'survival debates'. In one important paper, Strøm (1985) began focusing on the role of institutions. For example, he argued that governments that form in countries that have an investiture vote are more likely to survive. This is because potentially weak governments will not form in the first place, since they fail to

Figure 1.5. Share of early terminations for different institutional contexts



Graph made from data in Andersson et al. (2014)

pass the investiture vote. On the other hand, the opposite argument can also be made – i.e. that to pass the investiture vote, it might be necessary to include other, more dissimilar parties, which could make the government less coherent and thus more unstable. Regardless of the direction of the relationship, the point seemed clear – i.e. political institutions can affect cabinet stability and should therefore be included in the models. Later research has also shown that additional institutional factors such as bicameralism, semi-presidentialism and the constitutionally mandated powers of the PM can impact on cabinet stability (Damgaard 2008; Schleiter and Morgan-Jones 2009). A descriptive overview of the average number of terminations under different institutional contexts is provided in Figure 1.5. It confirms that some types of institutions – e.g. bicameralism and a mandatory vote of investiture – do seem to be associated with a higher likelihood of early termination.

1.2.2 ***‘The unified approach’ and a formalization of critical events***

Although both Strøm and others (such as Laver (1986)) rejected some aspects of the stochastic approach advocated by Browne et al., it had a few notable features that had a strong impact on developments in the 1990s. The early

modellers in the tradition of Taylor and Herman had had some success in explaining why some governments have a longer lifespan than others, but they could still only account for 30-40% of the variation in outcomes. This meant that there was still significant room for progress.

One insight to build on was that the general risk of ending prematurely is something that all governments face, and it should be more formally included in models. That is, since all types of governments in all countries run a certain risk of failing, cabinet life expectancy cannot be modelled only as a direct function of some explanatory variables. While they might be able to explain why government stability varies, they cannot measure the underlying risk shared by all. A second promising insight was that governments are not brought down by their general characteristics, but by events occurring sometime during their term of office. All governments can be said to be stable at the outset of their period in office, since the government that was formed was preferred to all others and no party had an incentive to defect at that stage (Diermeier and Stevenson 2000). If this changes and the government dissolves, it must be because the initial equilibrium has been disturbed. Measuring only stable characteristics such as type of government and effective number of parties (both of which are static throughout the term office) effectively renders us incapable of pinpointing the precise causal triggers.

Gary King et al. published a landmark paper in 1990 that addressed the first of these concerns (King, Alt, Burns, and Laver 1990). Using a novel statistical technique developed by Cox (1972) and known as the Cox proportional hazards model (discussed in Section 4.1), the basic risk affecting all government could be directly included in the model. At the same time, their model included a range of covariates capturing government, parliamentary and institutional attributes. This meant that the durability (expected duration) of any particular government could be modelled as a result of both a standard set of explanatory variables and an underlying risk component. Moreover, there is a feature of the Cox model known as censoring, that makes it possible to separate observations where the event of interest occurs from those where it does not. This allowed King and co-authors to separate discretionary from technical terminations.

The question of how to conceptualize and measure critical events was not directly dealt with by King et al. Instead, one of the co-authors of the paper, Michael Laver, wrote a paper with Kenneth Shepsle, in which they tackled the problem head on (Laver and Shepsle 1998). Their basic argument was that all governments are exposed to a continuous stream of potentially destabilizing events. Durable governments are those that are more likely to survive in the face of such events. To make the term 'critical events' more precise they divided the concept into four distinct categories:

- Policy shocks
- Agenda shocks
- Decision rule shocks
- Public opinion shocks

Policy shocks are external events that force governments to re-evaluate their policies. For example, a sudden increase in immigration might force the government to reconsider its migration policy. Developing new policies that were not on the table when the government assumed office can create problems about how to proceed. An agenda shock, in contrast, is a situation in which policies remain the same, but latent ideological differences between coalition members suddenly become salient. For example, if an issue not previously on the agenda, e.g. abortion, suddenly becomes politicized and the government must take a stance, ideological differences that were previously unimportant for the functioning of the government can suddenly become unbridgeable.

Decision rule shocks are about more formal or institutional changes in how decisions are made or how the government and the opposition interact. For example, during a vote on the annual budget in Sweden in 2014, the populist right-wing party the Sweden Democrats decided to reject decades of political praxis by not abstaining from future votes after their own budget proposal had failed. Instead, they supported a joint budget put forward by the other opposition parties, thus defeating the incumbent government's proposal. This nearly led to the downfall of the red-green government, which shows that changes to established decision rules can have substantial consequences. The fourth and final shock is perhaps the most intuitive and also most easily measurable, namely, changes in public opinion.

Enjoying public support and doing well in future elections are key concerns of any political party, so when the polls show that voters are shifting away from the government, we might imagine that the preferences and strategies of the cabinet parties will also change. Public opinion shocks are the main topic in two of the papers in this thesis.

A useful element of Laver & Shepsle's approach to critical events is that it enables researchers to incorporate them into the traditional approach that focuses on government and party system attributes. For example, if we are interested in policy shocks, we can ask ourselves when they are likely to be particularly detrimental? Probably when the government consists of a coalition of parties, and especially when those parties are not ideologically connected. Single party or ideologically unified governments (i.e. where the parties are adjacent on the left-right spectrum) should be far better equipped to deal with sudden policy crises. Ideologically dissimilar governments, in contrast, will find it far more difficult to successfully negotiate a new political

approach following a policy shock, because they will have different ideas about how to respond. This means that for such governments, a relatively small policy shock could prove destabilizing. The events approach thus enables us to develop better theoretical accounts of why certain empirical results tend to occur in our comparative studies and provides us with testable propositions about how the causal mechanisms that lead to cabinet termination should operate.

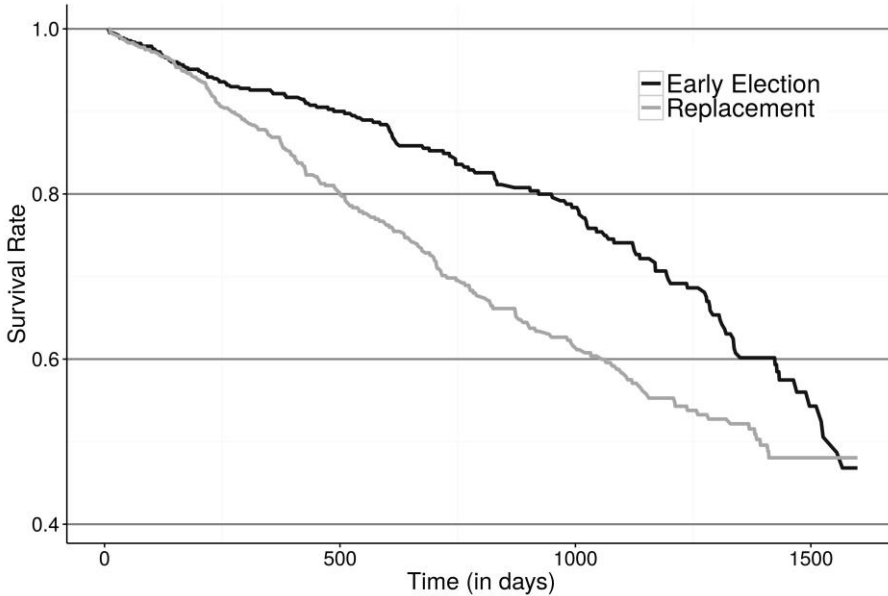
1.2.3 *Separating different termination types*

The paper by King et al. (1990) showed how technical and discretionary terminations could be empirically and analytically separated to help us predict how long different governments are likely to survive before terminating through preventable reasons. Lupia and Strøm (1995) then published an influential paper that argued that the category of ‘discretionary terminations’ should be broken up into two subcategories, because some of the terminations are logically and theoretically distinct. In particular, they argued that analyses should separate replacement terminations and early elections, because party incentives for the two termination types are different and because they occur at different times during the term of office. Replacement terminations take place when the composition of the government changes, e.g., when individual parties leave or join the incumbent government or when the entire government is replaced by opposition parties. In contrast, early elections are situations in which the incumbent government steps down and a new national election is called before the date when the regular, constitutionally mandated election was set to take place.

Empirical support for the differences between these termination types is presented in Figure 1.6. The figure is a Kaplan-Meier graph that shows how the likelihood of surviving beyond a given time point changes over time (Box-Steffensmeier and Jones 2004). The graph shows that the likelihood of termination through a replacement government develops in a fairly linear fashion over the course of the government’s lifetime. After only a couple of months in office, the risk of this type of termination starts to increase, and after 500 days only 80% of governments in the sample have not ended through a replacement. After 1000 days, 40% of governments (that have not exited the sample for some other reason) have ended through a replacement. In contrast, the risk of termination due to early elections increases only gradually for the first two and a half years. The risk then starts to increase exponentially, and after 1300 days, 40% of governments that have not already left the sample for other reasons (e.g. because of replacement) have experienced early election.

This is in line with our theoretical intuition. When the government has just assumed office, it has a lot of time left to enjoy being in power. Holding an early election and jeopardizing everything would therefore be a risky endeavour at this point in time (Lupia and Strøm, 1995; Diermeier and

Figure 1.6. How the likelihood of different termination types develops over time



The graph shows the Kaplan-Meier curve of the two termination types over time. Data comes from an updated and extended version of Andersson et al. (2014)

Stevenson, 2000). In contrast, a more limited replacement, e.g. a single party leaving the coalition, means that the rest of the government can remain intact. Thus, in the early stages of the cabinet life cycle, a replacement should be the preferred option. On the other hand, when the next regular election is drawing near, a replacement should be less likely. If the government is unstable and needs to change when there are only 6 months left in the term, it seems more reasonable to hold a premature election rather than go through the arduous process of reforming the government on the eve of a mandatory election. Obviously, losing the remainder of your term in office by calling an early election is not as big a sacrifice when there is less time left.

Based on this line of thinking, Lupia and Strøm (1995) argued that the two types of terminations might have different causes and that the effect of those causes might change over time. An external shock of a certain magnitude in the first half of the term of office might not be sufficient to trigger an early election at that point in time, but it could do so when the next regular election is approaching. Since the two termination types are both theoretically and empirically distinct, most studies now distinguish between them and run separate models for each outcome.

1.2.4 *Turning to the east*

A final significant development in the field that has had a direct bearing on the work presented in this thesis is the broadening of the sample of countries to also include the new democracies in Central and Eastern Europe. It has been a matter of some concern that a significant share of the research has focused on just 20-odd countries in the region that can broadly be described as ‘Western Europe’. A major reason for this concern is that there was an unhealthy interplay between theory development and empirical testing. That is, the set of countries that was used to develop theories about government behaviour was also used to test the theories (Laver, 2003). Although some of the more general theories about party behaviour were based on abstract game-theoretic calculations (Downs 1957; Riker 1962), most of the specific hypotheses came from observations about how governments in Western Europe behave. Since this process of both developing theories and testing them in Western Europe went on for decades, a natural question was how generalizable the results were to other contexts.

One useful opportunity for investigating whether the findings are supported in other settings came when democracy was gradually adopted in the former communist countries that were part of the Soviet sphere of influence until 1989-1990. There were now a number of new countries in Central and Eastern Europe where existing models could be tested. Researchers waited more than 15 years, until a sufficient number of cabinets had come and gone, before carrying out the first large-N comparative studies. Somer-Topcu and Williams (2008) published the first such study in 2008, and since then a handful of other papers have been published (Bergman, Ersson, and Hellström 2015; Grotz and Weber 2012; Savage 2013; Tzelgov 2011).

In general, the findings of these studies (which are discussed at length in Paper 1 in this thesis) provide a mixed verdict on the question of how far the results from the studies on Western Europe can be generalized. Some of the findings about the causes of government instability in Western Europe have been confirmed to apply in Central and Eastern Europe too. In particular, type of government seems to matter, as does the number of parties in the system, the distribution of bargaining power and economic developments. On the other hand, some factors seem to play less of a role. Ideological divisions and polarization are salient factors in Western Europe, but have little consistent impact in Central and Eastern Europe. The effect of institutions also seems less clear-cut.

In addition to such general differences in the effects of particular explanatory factors, differences in political behaviour between the two regions were also identified. For example, in Western Europe, surplus majority governments (i.e. those governments that have more parties than are technically needed to form a majority government) are far less stable than the average government. In contrast, this type of government is substantially

more stable than average in the East (Andersson, Bergman, and Ersson 2014). Also, voters seem noticeably less loyal in CEE. Electoral volatility, i.e. the number of voters who change party from one election to the next, is far higher in CEE, and it is usually government parties who experience the largest drop in electoral support (Roberts 2008). This means that being in power is costlier in CEE, and the initial government equilibrium is therefore more easily disturbed. Such differences between the two regions have made it possible to investigate how popular explanations of cabinet instability fare in quite different socio-political environments.

All in all, the developments outlined in the preceding sections have helped to define the field of cabinet stability studies as we know it today. Government stability is understood as a result of the joint impact of an underlying risk affecting all governments and the unique risk a particular government faces. Duration models, e.g. the Cox proportional hazards model, are used to capture both of these dimensions. Actual terminations are triggered by different types of critical events that occur during a government's the term of office; however, a government's sensitivity to such causal triggers is determined by pre-existing factors. Early elections and replacements are two empirically and theoretically distinct types of termination and should therefore be analysed separately. Some of the field's main findings appear to be generalizable to other contexts, but the variation between Eastern and Western Europe suggests that more work is needed in order to specify when and how individual factors matter. Taken together, these overarching features have informed the field's research agenda for the past two decades, and have thus ensured that we have a minimum set of agreed-upon principles necessary to enable the construction of a cumulative knowledge base.

2 Scope and contribution

This thesis is comprised of five distinct articles, and my goal is to make contributions in four main areas by:

- Identifying more reliable explanations for the causes of government instability in Central and Eastern Europe (Paper 1)
- Taking quality of governance and the state more seriously as explanations for government instability (Paper 2)
- Improving our understanding of the causal process that links economic changes to government termination by making the importance of type of government more explicit (Paper 3)
- Measuring public opinion trends and testing how they affect government duration (Papers 4 and 5)

Although these contributions might seem disparate, there is a joint mission. Together the five articles constitute an attempt to improve our ability to explain actual government duration by taking new explanatory factors seriously or by better modelling existing ones. In different ways, the research presented in the five papers is aimed at improving our understanding of when, how and why early termination happens in both Eastern and Western Europe, so that we can make more accurate predictions about how future cabinets will behave. By firmly situating all of the studies in the theoretical framework and methodological outlook of previous work, the intention is to enable the papers to make distinct but mutually compatible contributions to the overarching field of government duration studies.

The lion's share of attention in the thesis is focused on cabinet stability and political polling; therefore, some of the popular topics in coalition studies will not be dealt with directly. In particular, government formation, cabinet ministers and coalition agreements are all prominent parts of the overarching field of coalition studies, but they receive little attention here. They will be mentioned in the articles because they are part of the government life-cycle (and thus affect duration); however, they will not be the subject of any independent or original research. Moreover, the main focus is on causal mechanisms and methodological improvements and novelties. Less attention is paid to underlying theoretical premises and formal models of actor behaviour.

A more detailed overview of the contributions of this thesis is best given by looking at each paper in turn, so let us now provide a brief account of the focal points and main findings of each study. A more extensive summary of the papers can be found in Chapter 5. First, in existing studies of cabinet stability in Eastern Europe, a key challenge is that the findings from one study tend not

to be replicated in other ones (Sommer-Topcu and Williams, 2008; Tzelgov, 2011; Grotz and Weber, 2012; Savage, 2013; Bergman et al., 2015). Many explanatory factors that are shown to have a significant impact in one study are found to be irrelevant in others, or even to have an opposite effect. This problem can partly be explained by a lack of data and by the statistical problems that arise when there are many variables to test but a limited number of observations. That is, when we have too many explanatory variables compared to the number of cases, the result is low statistical power (Finlay and Agresti 2009). This, in turn, makes it difficult to uncover robust findings and makes the models highly sensitive to small changes in specifications.

Given this, one key goal of the first paper is to think carefully about both which factors should matter most in the newly established democracies in Central and Eastern Europe and how this can be tested in a way that avoids complications stemming from low statistical power. By combining theoretical insights from duration research with results from more general studies of party competition in new democracies (Kitschelt 1995; Teorell 2010), I put forward a simplified and streamlined set of hypotheses about which factors should be of particular relevance in CEE countries. Then, by using a statistical technique that has not previously been applied to duration studies – Sensitivity Analysis (Leamer 1985; Sala-i-Martin 1997) – it is here shown that the set of factors that systematically influences cabinet stability in CEE countries is far smaller than the number of factors that matter in Western Europe. In particular, this study shows that ideology and political institutions are less important in the East and that this is likely to continue to be true until the patterns of party competition have stabilized. Instead, the main determinants of government duration in CEE seem to be: level of support the government has in the legislature, number of competing parties and economic performance. This paper thus makes a contribution to the field by helping us pinpoint which factors are of primary importance when we seek to model government stability in the region.

A different type of contribution is made in the second paper, which deals with government stability and the state. The main premise here is that public sector management and state capacity (Dahlström, Lindvall, and Rothstein 2013) have important, but previously unacknowledged, explanatory capacity. Thus, while the paper on Eastern Europe contributes by pinpointing robust survival trends in the region (uncovering the signal in the noise, as it were (Silver 2012)), the goal in this paper is to broaden the scope of what is generally considered relevant in duration research. As discussed above, most of the explanatory variables have focused on attributes of the government, party system, institutions or socio-economic development. The public sector itself, however, has largely escaped scientific attention. This might be surprising given that the government and public bureaucracy are often

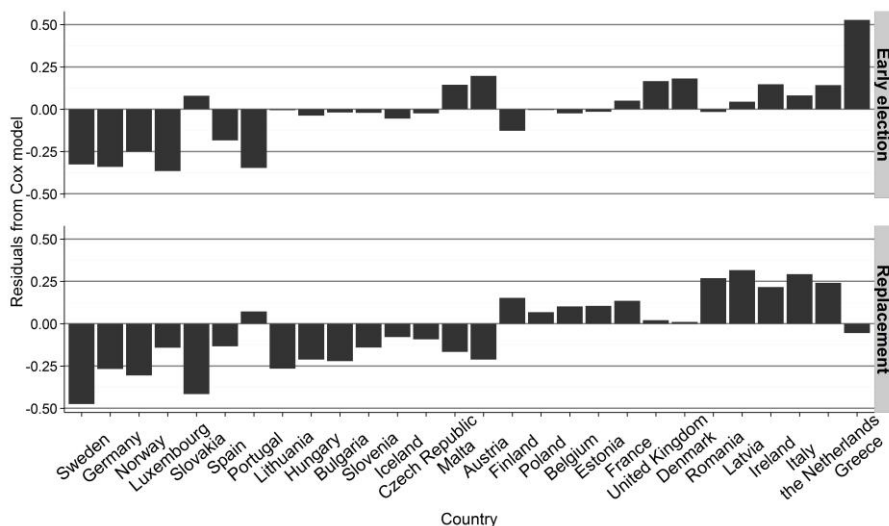
conflated in the eyes of voters and that their satisfaction with government performance is partly informed by how the general affairs of the state are handled. Corruption and inefficiency might lead to public disaffection with how the bureaucracy operates (Anderson and Tverdova 2003). In addition, if the government cannot rely on a competent and impartial civil service, it creates uncertainty in decision-making. If policy decisions are not reliably implemented, then it is difficult for the parties in power to successfully introduce the new policies that they and their voters favour. All of this might make it less valuable to be in power, which raises the risk of early termination.

Taking state capacity seriously might therefore be one way to extend existing models and capture an additional slice of the daily political reality that informs government parties' strategic decision-making. One important advantage of this approach is that it can help us account for regional and national differences in stability that previous models have missed. A simple empirical overview shows that governments in both Southern Europe and Eastern Europe are noticeably less stable than their counterparts in the northwest. That is, a government with the same characteristics in a similar institutional environment is more likely to fail in the southeast than in the northwest. To deal with this, statistical models routinely use fixed effects on the country level in statistical models in order to pick up this residual variation. But it would, of course, be better if we could directly model why governments in different countries appear to behave in different ways.

Figure 2.1 presents an overview of how the actual duration of governments in the countries in the dataset used in this thesis (Andersson, Bergman, and Ersson 2014) differs from the expectations of a statistical model that controls for some common explanatory factors in the field. Thus, if we run a typical model and use the residuals to calculate how the observed duration of governments in the different countries deviates from our expectations, we see that there are predictable mistakes. Governments in Sweden, Germany and Norway tend to be a lot more stable than the attributes of the governments would lead us to believe, whereas countries like Greece, the Netherlands and Italy are less stable. Such systematic problems in explaining actual survival patterns in particular countries suggest that important variables are missing. Paper 2 makes the theoretical case that state capacity appears to be one such previously neglected factor that has important explanatory power, and then demonstrates this empirically. Thus, state capacity can help us account for some of the inter-country differences we observe, and is therefore a useful addition to existing models.

The third paper is concerned with economic developments, and in particular with how different types of governments respond to similar changes in economic circumstances. Most studies on government stability include some measures of economic developments, e.g. GDP growth, unemployment or inflation (see Warwick (1994) for a longer discussion of the importance of

Figure 2.1. Difference between model estimations and actual government duration



The graph plots the residuals from an attempt to predict the actual duration of governments using the estimated coefficients from the Cox model. Negative residuals (i.e. downwards pointing bars) mean that the governments in the country were more stable than the model predicted. The upper half of the graph shows the results for early elections while the lower shows replacements.

economic affairs). The relevance of economic factors – unemployment and inflation in particular – in predicting premature government dissolution has been empirically confirmed repeatedly and is also theoretically intuitive. Economic problems can lead to a decline in popular support for the government and might also contribute to policy challenges if new measures must be devised to deal with the economic situation.

Article three presents evidence that the effect of economic deterioration does, in fact, vary depending on government type. That is, economic changes do not have the same effect on all types of governments. Some governments become significantly more likely to end prematurely when the economy takes a turn for the worse, while others are unfazed. In particular, single party governments remain robust in the face of economic deterioration, while coalition governments (regardless of whether they enjoy minority or majority support in parliament) become more likely to end prematurely. This can be explained by the greater policy discretion of single party governments. Coalition governments often have explicit or implicit policy agreements that determine what they can do during their term of office. Significantly changing such agreements to deal with economic problems can be difficult and might contribute to tension in the coalition. In contrast, single party governments

have far more discretion and are therefore better positioned to alter their economic agenda or replace ministers associated with failing policies (Dewan and Dowding 2005).

The fourth and final contribution, which required two separate papers, is concerned with how government stability is affected by changes in public opinion. The general idea behind this link is both intuitive and a direct implication of Laver & Shepsle's (1998) work on critical events. The intuitive part is that government parties want to win elections and if their public support (as measured by opinion polls) changes radically, this might prompt them to rethink their current government membership. In particular, we can imagine that government stability is affected both by bad poll numbers (when cabinet parties might have an incentive to step back in order to minimize vote loss) and by good poll numbers (when parties might have an incentive to opportunistically seize the moment and call an early election while times are good (Kayser 2005; Schleiter and Tavits 2016)).

Thus, it seems reasonable to assume that the future electoral prospects of the cabinet parties would play an important role in government survival. In fact, whether and, if so, how popularity plays a role in explaining stability has not been previously incorporated into comparative duration models. One reason for this is a practical matter: reliable opinion data has been scarce, especially for earlier decades. Another reason is that it is not obvious how to aggregate the polling data into a coherent time series that allows us to accurately model developments over time.

The latter concern led me to conduct a separate study about how developments in party support can be modelled. Unlike many other policy areas for which polling data is collected, polling on party support can actually be tested for accuracy because we get one perfectly unbiased measure, namely, the general election. Thus, by comparing model estimations of the levels of party support with the actual election results we can get an indication of how trustworthy our calculations are. Moreover, if the estimations turn out to be reasonably accurate on election day, we can also assume that they provide reliable information at other times during the government's term of office. Most of the variables included in our duration models, e.g. government type, institutions and even economic developments, can be directly and objectively measured. Party support must be indirectly inferred from polls, and for this reason it was necessary to first conduct a separate test about how to do this as accurately as possible, before including the estimations in our duration models.

Paper 4 contributes to the steadily growing literature on electoral forecasting (Fisher, Ford, Jennings, Pickup, et al. 2011; Jackman 2005; Linzer 2013) by refining and testing some popular Bayesian techniques for aggregating polls in order to see what works best in the complex multiparty systems in our data. In particular, I first demonstrate that election results can

be forecasted with good accuracy even in PR systems with many active parties. I then suggest an improvement to existing models that captures seasonal fluctuations in the data so that we can predict seasonal shifts before they are actually present in the data.

Finally, the fifth paper is concerned with the empirical task of actually testing the practical impact of changes in public support on cabinet stability. In many ways, this can be seen as the main contribution of the thesis, because government popularity has been an implicit component of a wide range of theoretical and statistical models of government stability, but it has never actually been tested. For example, increases in unemployment have been found to lead to a higher risk of early dissolution, but the causal hypothesis is that this happens because high unemployment impacts negatively on the government's future electoral prospects (Robertson 1983; Somer-Topcu and Williams 2008). Similarly, governments comprised of ideologically diverse parties are less stable partly because they find it more difficult to pursue the types of policies their voters are interested in (Saalfeld, 2008), which might lead them to support another party in the next election.

Thus, changes in public support are theoretically assumed to be part of the underlying causal mechanism behind many other factors that have been found to matter in cabinet duration studies. But if public opinion is a fundamental component in the causal story, it should be tested directly. For this reason, and because public opinion shocks have such a central place in the Laver-Shepsle model, being able to test it empirically is an important next step in duration modelling. Directly operationalizing such a central component should make our models more accurate and more directly reflective of actual political praxis.

The fifth paper draws two main conclusions. First, the popularity of the government does have an impact on the likelihood that it ends prematurely. The government's standing in the polls influences its likelihood of ending both through early elections and replacements. In particular, good polling increases the likelihood of opportunistic early elections, while bad polling increases the likelihood of a replacement government, because individual party members or the entire government are more inclined to step back and rebuild their public support before the next election. The second conclusion of the paper is that this effect is strongest in single party governments. In coalition governments, popular support can shift in different directions for the different party members, which makes it more difficult to agree about how to respond to polls. All parties might have ideal individual responses, but it is difficult to find a joint response when the parties' poll numbers go in different directions. For that reason, popularity mainly affects single party governments, because they have greater discretion about how to react.

In sum, together, the five papers that make up this thesis are intended to make slightly different contributions to the overarching field of cabinet

duration studies. But they are all firmly grounded in the, by now, canonical understanding of how government duration should be theoretically understood and empirically investigated. The overarching goal that has guided the research is to improve on the state of the field by teasing out the reliable trends in the government survival patterns in Central and Eastern Europe (Paper 1), extending our previous understanding of what influences stability by including previously omitted explanatory factors such as state capacity (Paper 2), elucidating the causal path linking economic changes to termination by taking the importance of government type seriously (Paper 3), and measuring and incorporating government popularity in novel ways to empirically test popular ideas in the field about how public support and stability are connected (Papers 4 and 5).

3 Theory

Good tests kill flawed theories; we remain alive to guess again.

Karl Popper⁵

In the next two chapters, I return to the broader theoretical and methodological discussions that have inspired and informed the field of coalition studies in general and the research on government duration in particular. The goal here is both to position the studies in the thesis within this larger context and to provide a template for thinking about future research.

Starting with theory, the research carried out in comparative coalition studies relies on theoretical frameworks with different levels of abstraction. On the most fundamental level, many underlying premises in the field are inspired by insights from rational choice and formal game-theoretic models (Axelrod 1970; Downs 1957; Neumann and Morgenstern 1944; Riker 1962). These theoretical starting points have subsequently been refined and concretized within the field of coalition studies to generate explanations that are more directly applicable to modern party behaviour. In some cases, the modern refinements of the early models are themselves formal game-theoretic exercises (such as the model by Lupia and Strøm (1995) discussed above). In many cases though, researchers apply a kind of ‘soft’, pragmatic rational choice approach that draws inspiration from the formal models but base their specific hypotheses on looser and less formalized thinking about how parties and governments are likely to behave. In the next two sections I briefly outline why I think rational choice is a suitable theory for the study of coalition politics and how it is used in this thesis.

3.1 A brief background: why rational choice is popular

In comparative political research, it seems fair to say that rational choice (RC) theory enjoys an overwhelming dominance. For a long time, RC was the leading approach in political science more generally (particularly in the US), and in the early 1990s about 40% of the articles in the prestigious *American Political Science Review* used RC as a starting point (Hindmoor 2006, 9). Although its popularity has declined since then, the theory has managed to retain its standing as the leading approach in the field of comparative government. Two key reasons can be put forward as to why RC is still dominant here. First, despite the rather late emergence of coalition research as a field of its own (1960s-70s), it already has a rather impressive bank of accumulated knowledge. This means that new researchers coming into the field do not need to start from scratch, but can proceed from the progress that

⁵ Quoted in Vary (2011)

has already been made and thus venture further than would otherwise be possible. This commitment to cumulative knowledge is an important strength inherent in RC and has served the field of coalition politics well.

Secondly, the assumptions of rational choice, such as strategic behaviour and utility-maximization, fit the sphere of formal political competition particularly well, and if the theory has any merit this is one of the areas in which it should be useful (Shepsle 2010). Indeed, the goal in the study of coalition politics is not to investigate how normal people make everyday decisions, but instead to understand how a few key individuals and organizations make crucial decisions. This includes whether or not to join a government coalition, whether a party leader should follow the wishes of the party base or the median voter, and whether to vote for or against a certain policy (e.g. EU membership) (Müller and Strøm 1999, 9–11). It seems reasonable to assume that people involved in making such decisions seek out more information, have a better understanding of the issues and other players involved, rely on a more conscious and self-aware deliberative process, and pay more attention to how long-term goals can be achieved effectively. They are also aided by other actors and organizational structures that have been set up to improve decision-making (Simon 1991; Tversky and Kahneman 1981). This means that even if perfect rationality can be criticized if it is assumed to apply to all humans everywhere, it is more likely to apply to the preference-ordering and strategy development of high-level political actors.

There is also a methodological reason why rational choice fits coalition politics well, namely, that the field is characterized by large-N statistical studies. This characteristic ensures that individual quirks and idiosyncrasies, as well as unusual norms or cultural practices in particular countries, play less of a role since the overarching trends that we are interested in should still be apparent. We do not need to predict the future actions of one specific government, but can instead aim to identify trends that apply on the group level. This means that average net effects are in focus. What is the mean effect of a certain variable of interest across our group of countries, once we have controlled for other factors? We may, for example, conclude that given certain scores on the vector of independent variables, a government may have a 70% risk of terminating prematurely. This forecast only applies on the group level, however, and thus means that given enough observations, 7 out of 10 governments with those characteristics should fail. The task of the rational choice-based explanations is only to provide an account of why this group level trend exists. Other, more specific explanations must be sought for the deviant cases.

3.2 Applying rational choice to party behaviour

Although it is important to keep the underlying theoretical concepts in mind, in the daily work of explaining the actions of real life coalitions, it is rarely

necessary to derive hypotheses and explanations from first premises. Instead, it is generally accepted as given that we are interested in the strategic behaviour of political parties, and we try to explain and predict what incentives and interests they are likely to have in different situations. How parties think about cabinet dissolution is then a subtype of their more general interests.

So what exactly are the strategic interests that parties compete over? A useful, if idealized, summary has been provided by Müller and Strøm (1999). They make the claim that it is possible to distinguish three main traditions in the field. In the early literature in the field, it was assumed that the main interest of parties (and of the individuals comprising them) was to gain office in order to enjoy the material benefits and status associated with holding public office (Downs, 1957; Riker, 1962). Thus, politicians desire office, and in order to achieve it they create policies that will convince voters to support them (Van der Eijk & Franklin, 2009, 15). The specific policies that parties choose to adopt are simply a means to an end rather than desirable goals in their own right. This assumption generated a number of specific hypotheses, some of which were confirmed, but many of which failed to live up to empirical testing, and the models generally did not explain as many of the observed outcomes as researchers had hoped. Numerous hypotheses stemming from office-seeking theory were rejected, including that parties will join a cabinet with any other party in order to gain office (ideology turned out to matter), that minority governments will be brought down by a parliamentary majority of parties wanting to gain office (they are not usually), and that cabinets will form with the parties that can create the smallest possible majority so that the spoils of office do not need to be shared with more people than necessary (surplus majority governments are common).

A second hypothesis was therefore put forward: that politicians are policy maximizers (De Swaan and Rapoport 1973). This assumes that politicians are ideologues who want to achieve certain policy outcomes. In many situations, this generates the same predictions as the office-seeking approach, since parties are better able to influence policy outcomes if they hold the reins of government. In the policy-seeking approach, however, office is a means to an end (i.e. changing society in a particular direction) rather than valued for its own sake (Van der Eijk and Franklin 2009, 2–3). And in some cases, the two approaches generate completely different hypotheses. Policy-seeking politicians will not form a government with just anyone. They will be more likely to rule harmoniously with parties of similar ideological preference, and they will sometimes refrain from changing their policies even if a different one would make it more likely that they end up in office. The policy-seeking approach was generally more successful in explaining observed outcomes and was therefore gradually introduced in the models as a complement to office-seeking.

A third main goal that parties might pursue is vote-maximization. In this case, office or policy goals are not pursued directly. Instead, future electoral gains are in focus because votes can be traded later for both office and policy (Müller and Strøm 1999, 200). This approach is the most difficult to separate from the others, because it is a means rather than an end. Votes are not pursued for their own sake, but are instead treated as a tool to achieve something else, i.e. either office or policy. The vote-seeking approach is therefore sometimes compatible with one or both of the other assumptions. A party can reject an offer to participate in a government because it fears the effects that doing so would have on its long-term electoral success. This is because being in office is likely to have a negative impact on future electoral performance if the party is forced to sacrifice its policy goals. The voters of a party are unlikely to continue supporting it if it does not deliver the type of policy it has promised. In this case, vote-seeking and policy-seeking explanations appear to coincide. Conversely, a party leader might need to ignore demands of ideological purity put forth by the party base because of the potential loss of votes that this might entail, which means that an excessive focus on policy would hamper the quest for office.

In addition to this, (Sjöblom 1968, 74) added a fourth goal – party cohesion. In addition to achieving their strategic interests today, parties also need to survive and remain relevant as coherent actors in the future. Therefore, all parties have an interest in avoiding factionalisation and infighting by developing sound internal decision-making structures and keeping the interests of the party base in mind. In the studies undertaken in this thesis, parties are generally treated as unitary actors, so this fourth goal will not be directly utilized.⁶

A general lesson to draw from these different perspectives is that parties have multiple and sometimes incompatible goals, and they are often required to prioritize one of them. For example, parties might focus more on holding office in the beginning of a term in office (when there is more time left to enjoy the advantages it offers) (Lupia and Strøm 1995), while focusing more on policy or votes as the next election approaches. It is often difficult to combine differing assumptions of preferences within a rational choice framework, since they generate incompatible predictions (Geddes 2003, 177). However, that does not mean that it is necessary to permanently adopt one or the other, only that it is necessary to specify which assumption is being utilized in a particular study. In the studies undertaken here, the basic approach has been to take into account the main causal mechanisms and contextual factors that are relevant for the particular study in focus, and then to think critically about which goal is likely to take precedence under those circumstances.

⁶ See Druckman (1996) and Laver and Schofield (1990, 15) for discussions of the unitary actor assumption.

For example, in Paper 2 (which deals with how state capacity influences government stability) it is argued that parties are more likely to desert the government when state capacity is low. State capacity should not have any direct impact on the ability of the parties to hold office (office-seeking), but with a poorly functioning state bureaucracy, it is difficult to implement new policies (policy-seeking), which, in turn, contributes to voter dissatisfaction due to a lack of tangible progress (vote-seeking). This means that when the next election is approaching and the importance of popular support rises, parties can find it in their interest to leave the government prematurely in order to minimize the vote loss. How parties make trade-offs between the different goals and when particular goals become dominant are thus crucial features of any causal theory of government instability.

3.3 Later theoretical developments in the study of government stability

The theoretical premise that to understand government behaviour is to understand how their constituent party members strategically pursue their different interests is also a cornerstone of the subfield concerned with government stability.

However, as discussed in the summary of the field presented in Chapter 1, in the 1990s there were several important theoretical developments that had more precise implications for government duration. The general inspiration behind many of these theoretical refinements was a goal to produce more detailed expectations and predictions by explaining the causal process (George and Bennett 2005) that leads to a particular type of termination. The work by Lupia and Strøm (1995) helped distinguish between two different termination types: replacements and early elections. It also helped us develop a clearer understanding of how the different interests that parties have influence their incentives for early termination in various stages of the electoral cycle. This was a useful concretization of the general theories about party behaviour, because it made it possible to articulate more precise expectations about when termination would be a rational option and what type of termination it would be.

Laver and Shepsle's (1998) work on 'shocks' and how much the external environment has to change during the term of office in order to make the strategic calculation of parties shift in favour of termination also helped render the general theories more specific. Their claim that the willingness to terminate prematurely has a micro-causal connection to particular developments in the political environment (such as changes in the policy agenda or the popularity of government parties) helped pinpoint when and how party interests change sufficiently to make termination rational.

An important theme in this and in later work (e.g. by Diermeier and Stevenson 1999, 2000; Strøm and Swindle 2002) is the creation of a more

detailed set of causal expectations specifically for duration studies. This gives us a shared set of theoretical starting points that are more precise and practically useful than the general rational choice work on party interests discussed above.

Another noteworthy, more recent development is the increased focus on conditional effects (Warwick, 1994; Saalfeld, 2008; Grotz & Weber, 2012). This is less a development in formal theory and more an advance in hypothesis generation and modelling. In any case, the new emphasis on how the effects of some explanatory factors are modified (reinforced or hindered) by other factors is highly useful for developing a better theoretical understanding of the operation of causal pathways. For example, the effect of increased unemployment is likely to depend on how well the government can respond to it. A majority government that consists of ideologically similar parties will find it easier to agree on and implement new policies to address the problem. In that sense, the same external development could have quite different effects depending on the government's other circumstances. This makes the effect of unemployment conditional rather than additive. I will further elaborate on the advantages of developing conditional theoretical predictions in the final chapter, when I discuss future projects.

Taken together, these and other theoretical contributions in the past two decades have helped make the theoretical framework more directly focused on the question of government duration and have helped clarify the causal process. This is the underlying theoretical framework that has inspired the papers included in this thesis.

4 Methodology and data

Essentially, all models are wrong, but some are useful.

Box and Draper (1987)

Comparative coalition research in general, and the field of duration studies in particular, relies mainly on large-N studies and quantitative models. Nonetheless, the dominance of statistical methods should not be taken as a given. The Popperian philosophy of science and the theoretical starting points discussed previously are equally compatible with studies that use a smaller sample size and qualitative methods. Case studies, elite interviews and fuzzy set analysis are examples of techniques that are equally suitable for exploring the questions and testing the hypotheses developed in the field (King, Keohane, and Verba 1994).

In practice, however, the vast majority of studies are quantitative, and that tradition is followed in this thesis. A primary reason for the dominance of the quantitative approach is that the effects we are interested in and the variable relationships we are positing are noisy. In such circumstances, a large number of cases are needed to separate the signal from the noise (Silver, 2012) in order to adequately test the hypotheses. In addition, given cross-national variation in political traditions and in formal institutional settings, it is often useful to include governments from many countries to control for some of the heterogeneity. This makes large-scale data analysis a logical choice.

Table 4.1. Countries and government duration in ERDDA

Country	Cabinets	Early Elections	Replacements	Survival rate	First entry in dataset
Austria	25	12	3	0.38	1945
Belgium	40	11	18	0.28	1946
Bulgaria	10	1	3	0.60	1990
Cyprus	9	0	4	0.56	1988
Czech Republic	11	0	6	0.45	1992
Denmark	35	21	5	0.26	1945
Estonia	12	0	7	0.42	1992
Finland	50	4	25	0.42	1945
France	29	5	10	0.48	1959
Germany	29	3	12	0.46	1949

Country	Cabinets	Early Elections	Replacements	Survival rate	First entry in dataset
Greece	15	10	1	0.27	1977
Hungary	10	0	3	0.67	1990
Iceland	32	7	7	0.55	1944
Ireland	25	16	5	0.16	1944
Italy	55	9	38	0.15	1946
Latvia	19	1	13	0.26	1993
Lithuania	12	0	6	0.50	1992
Luxembourg	19	2	2	0.78	1945
Malta	7	2	1	0.57	1987
Norway	30	0	9	0.70	1945
Poland	16	0	11	0.31	1991
Portugal	19	5	6	0.42	1976
Romania	17	0	12	0.29	1990
Slovakia	10	1	5	0.40	1992
Slovenia	12	0	6	0.50	1990
Spain	11	6	1	0.36	1977
Sweden	29	1	4	0.82	1945
The Netherlands	28	6	11	0.39	1946
United Kingdom	24	12	2	0.39	1945
Total	640	135	236	0.44	

The comprehensive dataset known as ERDDA – the European Representative Democracy Data Archive (Andersson et al., 2014)⁷ – is used in this thesis to conduct such analyses. The dataset contains data on 640 individual governments in 29 European democracies. The countries that are included are the 27 countries that were members of the EU before Croatia joined in 2013, as well as Iceland and Norway. A full list of the countries is presented in Table 4.1. The table shows that the countries enter the dataset at different dates. This is because only democratically elected governments are of interest, so Spain, Portugal and Greece enter the dataset in the 1970s, while the Central and Eastern European countries that were under the influence of the Soviet Union

⁷ This dataset is freely available at www.erdda.se

during the Cold War do not join the dataset until after 1989. Most other countries are covered from the end of the Second World War to the present.

For each of the included governments, the dataset contains more than 100 variables that describe things like the government itself, formal political institutions, the bargaining environment and socio-economic context. For example, there is data on the type of government (minority or majority, single party or coalition), the ideological spread within the government and within parliament, the existence of institutions such as a second chamber or semi-presidentialism, socio-economic developments (e.g. inflation, unemployment and GDP growth), as well as duration-related items such as how long the government survived and whether termination was by replacement or election. The comprehensive nature of the dataset makes it possible to test a wide range of hypotheses while controlling for other factors that could potentially influence and bias the results.

In order to conduct better tests of some of the hypotheses explored in this thesis, I did additional data collection and coding for some of the duration-related variables. In particular, I wanted to distinguish between discretionary terminations (ones that the government could have prevented) and technical terminations (ones that occurred for constitutional or other reasons and thus do not indicate cabinet instability). Such detailed coding had previously been done for the Western European countries up to 1999, but the information was missing for all Western governments after that date, as well as all Eastern European countries. Expanding the sample to include these countries was an important next step, and I therefore spent a couple of weeks sifting through archives and reading historical accounts in order to ensure that all of the 640 governments in the ERDDA were also fully coded in those areas that are of particular importance for duration studies.

Deciding whether a government terminated for discretionary or technical reasons is often an easy matter. If one of the coalition parties left after a failure to agree on policy or if the government called an early election after losing a vote of no confidence, then it is clear that the government came to an end for preventable and political reasons. Similarly, if the government ended through a regular election or the prime minister died (which automatically brings about a new government), then it is coded as a technical termination. However, in some cases it was more difficult to determine whether the termination should be treated as discretionary or technical. There were two reasons for this. In some cases, the facts of the termination were unclear and the main sources that were used (Keesing's World News Archive, the Political Data Yearbook, the Inter-Parliamentary Union and Wikipedia) emphasised different key reasons for the termination. In such, very rare, cases I tried to find more detailed information (e.g. articles or book chapters about the termination written by political scientists) and discussed the matter with knowledgeable colleagues.

A second problem was that in some cases it was clear that the termination was preventable, but it was unclear whether it was an example of the kind of government instability that is of interest here. To make the data useful for duration studies, we need to exclude cases in which a termination occurs for stochastic or other non-political reasons, since such cases tell us little about the structural causes of government instability. This can best be explained by looking at a couple of real world examples.

First, the Oddsson IV government in Iceland was terminated in September 2004 when Prime Minister Oddsson stepped down and handed over power to foreign minister Asgrimsson. At the time, there was almost three years left until the next regular election. The decision was voluntary, and was thus clearly not prompted by technical or constitutional reasons. In addition, the decision followed a political defeat, when a bill that Oddsson had been championing was vetoed by the president, which seems to suggest that the termination was politically motivated and thus of interest here. On the other hand, Asgrimsson and Oddsson had made a deal already when the government assumed office in May 2003, that Oddsson would step down in September 2004 as part of the coalition agreement. This meant that the termination was pre-planned and not immediately caused by other external circumstances at the time. Because of this previously existing agreement, the termination was coded as technical, even though it was within the government's power to avoid it.

Another example, even more difficult to classify, is the Ahern III government in Ireland. This government came to an end in May 2008, when Prime Minister Ahern voluntarily stepped down. He had received cash gifts from businesses that he had cooperated with, and this was widely seen as problematic and a potential case of corruption. In public opinion polls, between 55 and 65% of Irish voters said that it was wrong of him to accept the payments, and he stepped down in May to protect his Fianna Fáil party from further ramifications of the scandal. This termination was clearly preventable, but it has nonetheless been coded as technical. The main reason for this decision is that the scandal involved only the Prime Minister and was therefore more of a personal event than an issue involving the party or government. The departure of a prime minister has sometimes been coded as discretionary, for example when it was prompted by intra-party fighting or a poor result in a regional election. However, a scandal involving only the prime minister and no other senior party figures can be seen as a stochastic, external shock in the tradition of Browne et al. (1984) rather than an event that says something about the general stability of the government.

These decisions can certainly be debated. However, out of the 640 governments in our sample, very few cases were truly problematic. Thus, even though future coding exercises could change some of the decisions, they are unlikely to alter the substantive results.

Once the dataset and the updated coding were complete, the next challenge was to select appropriate methodological frameworks for the different papers. Overall in this thesis, duration analysis was the most common statistical approach. In some papers, duration analysis was used in conjunction with other methods, such as sensitivity analysis (to measure how robust the explanatory factors are) and the Bayesian Dynamic Linear Model (for aggregation of polls). The most popular type of duration analysis relies on the Cox proportional hazards model, which I found offered enough benefits to use as the main method in this thesis. The necessarily constrained format of the articles did not provide the space needed to explore the underlying workings of this model in any detail, so I think it is worthwhile to devote some attention to that here, in particular to elucidate how it works and what advantages it offers given the goals of the thesis. In some papers the logit-based discrete-time model was used instead (Box-Steffensmeier and Jones 2004). But that model is asymptotically identical to the Cox model.

4.1 **How to analyse duration data and why the Cox model was chosen**

As mentioned briefly in Section 1.2.2, the introduction of the Cox proportional hazards model to the field of cabinet duration by King et al. (1990) enabled both theoretical and methodological improvements. It neatly captured the distinction between duration and durability and thus helped separate the general risk of termination affecting all governments from the specific risk that a particular government faces, which stems from its particular attributes. The model itself was initially developed by Cox (1972), who was interested in human mortality. These origins and the fact that the model first gained widespread acceptance in medical research meant that the technique as such became known as ‘survival analysis’. One popular early application was to model how long patients with different pre-existing conditions survived a particular type of disease. The statistical model can, however, be used to model all types of time-to-event data. For instance, economists have used it to model how long industrial plants survive in different sectors of the economy (Tveterås and Eide 2000) and to explore what helps people get jobs after a period of unemployment (Korpi 2001). Sociologists have used it to analyse how long it takes for people with various characteristics to get married (Trussell and Bloom 1983) and so on. Adopting it to measure how long governments manage to stay in office is also a fitting application of the technique.

The model has a range of statistical properties that make it particularly suitable for the problem at hand, and I will briefly mention two that I found especially useful. One important feature is what is known as *censoring* (Allison 1995; Box-Steffensmeier and Jones 2004, 16–19). Censoring is used when the event of interest (early discretionary termination in our case) does

not occur, which happens, for example, when the government survives all the way to the next election. In such cases, we know how long the government survived, but not how long it would have survived if it had not been forced to leave because of the constitutional term limit. This means that we do not know how durable the government actually is, because it has not been allowed to terminate through a discretionary cause.

The censoring technique in the Cox model addresses this issue by using the duration information from the censored cases and adding it to the general pool of survival data in the dataset. So, for example, if we have two governments with identical characteristics, and one survives for the full four-year term of office while the other fails after three years, then we would, on average, expect to have one failure per 7 years for such governments. In contrast, Ordinary Least Squares regression, which was originally used by researchers in this field (Warwick 1979), would treat both cases as fully observed. The mean survival rate would be estimated to be 3.5 years, and this would be the expected one for future governments with similar characteristics, even though this does not do justice to the actual information in the data. In our case, we censor all technical terminations since those governments came to an end for reasons that are not of substantive interest here.

Being able to handle censored cases is immensely helpful in the analyses, and it was one primary reason why the model was chosen. However, an equally important advantage is the fact that the Cox proportional hazards model filters out the baseline risk component (Kleinbaum and Klein 2005). The hazard rate common to all subjects (i.e. the general risk that Browne et al. (1984) talked about) is cancelled out in the equations that define the Cox model, which means that we do not need to assign it a probability distribution. This means that the risk of termination that all governments share disappears from the equations and we can instead focus on how a particular government's characteristics make it more or less stable than the average government. We can investigate the relative effects and the statistical significance of the covariates we enter into the model without having to specify the underlying hazard rate affecting all governments. Or, to phrase it differently, we can measure deviations from the baseline hazard caused by the independent variables without modelling the hazard that they are deviating from. The model is therefore known as semi-parametric (Box-Steffensmeier and Jones, 2004). This makes the analysis significantly more flexible than would otherwise be the case.

To see more clearly how this works, we can briefly examine the underlying mathematics. To set up the model we first need the duration of each subject and then an indicator variable (censor variable) that tells us whether or not the event of interest (discretionary early termination) has occurred. The basic formula for the Cox model is then (Cox 1972):

$$h(t, \mathbf{X}) = h_0 t \times e^{\beta_i X_i} \quad (4.1)$$

where $h(t, X)$ is the hazard rate (the risk of early termination) at time t given the vector of independent variables X . $h_0 t$ is the baseline hazard (shared by all governments) that is then multiplied by the exponentiated variables and coefficients. Maximum likelihood estimation is then used to calculate the most efficient estimators (Kleinbaum and Klein 2005, 95). Note also that when all the independent variables are 0, the hazard rate reduces to the baseline hazard since $e^{\beta_i \times 0} = e^0 = 1$, and $h_0 t \times e^{\beta_i X_i}$ then reduces to $h_0 t$.

The goal of the model is then to calculate a hazard rate for an individual government. To do this we need to calculate which variables cause a statistically significant increase in the risk of early government termination, which variables lower it, and which variables have no significant effect on the baseline hazard. Importantly, we do not need to know the baseline hazard (i.e. $h_0 t$) in order to calculate how the variables influence the risk. The reason for this is that the difference in the hazard rate when we change a certain independent variable can be calculated through (Box-Steffensmeier and Jones 2004; Kleinbaum and Klein 2005):

$$\widehat{HR} = \frac{h(t, X_{change})}{h(t, X)} \quad (4.2)$$

In line with equation 4.1 we can rewrite this as:

$$\widehat{HR} = \frac{h_0 t \times e^{\beta_i X_{change}}}{h_0 t \times e^{\beta_i X_i}} \quad (4.3)$$

But now we have the baseline hazard $h_0 t$ in both the numerator and the denominator, which means that it cancels out! We thus get

$$\widehat{HR} = \frac{e^{\beta_i X_{change}}}{e^{\beta_i X_i}} \quad (4.4)$$

So to estimate how the hazard rate of early termination for a certain government changes compared to the baseline risk, we only need to look at

how the coefficient changes for new values of X. This becomes clearer if we simplify the math further:

$$\widehat{HR} = \frac{e^{\beta_i X_{change}}}{e^{\beta_i X_i}} = e^{\beta_i (X_{change} - X_i)} \quad (4.5)$$

And finally

$$\widehat{HR} = e^{\sum_{i=1}^N \beta_i (X_{change} - X_i)} \quad (4.6)$$

In the final equation, we can see that the hazard rate is simply the exponentiated summed change in the coefficients for marginal changes in X. This is the ingenuity of the Cox proportional hazards model and the key to its operation. Regardless of the shape, level or change in the baseline hazard, we can still measure how the variables we are interested in (e.g. government type, number of coalition parties, institutional set-up, etc.) affect the hazard rate. And we get this as a proportion. We can say that, holding the other variables constant, the risk of early termination increases by 50% if we increase a certain variable by one. This model, coupled with other descriptive data and tests in survival analysis, e.g. Kaplan-Meier curves, the log rank test and Nelson-Allen residuals, allows us to model and test government stability far more accurately than was possible in the past (Kleinbaum and Klein 2005).

The main upside of the model, its semi-parametric mathematical basis, is sometimes also a shortcoming. The result we get from the model is how much a certain variable increases or lowers the risk of early termination compared to the baseline risk. But since the baseline is filtered out, it is not immediately obvious what the total risk is. An increase in the hazard rate of 20% for a certain government is only meaningful if we know what the hazard rate was before. But since the baseline hazard rate was filtered out, we cannot use it to calculate the total risk of early termination for the government. One solution to this is to calculate the baseline hazard outside of the model, and then conduct a separate calculation in which the baseline risk and the change in the hazard rate for a particular government are added together.

Another approach to solving this issue, which I used a number of times while working on the thesis, but did not include in the final version of the papers, is to specify a fully parametric version of the duration model. That is, instead of filtering out the baseline hazard, it is assigned a probability distribution (typically a Weibull or exponential distribution), and the fully specified model is then used to calculate the risk that a particular government has of failing (Box-Steffensmeier and Jones 2004). Fully parametric models have occasionally been used in duration studies (Sommer-Topcu and Williams

2008), and they are useful when there are theoretical reasons to expect that the underlying hazard distribution takes a certain form. If it is difficult to decide on theoretical grounds which distribution to expect, one can also use a more empirical approach – e.g. run log-likelihood tests to compare the relative fit of models with different assumed probability distributions, and then use statistical criteria to determine which distribution seems to be the most accurate.

A third option is to use a discrete-time model with a logistic link function as I did in a couple of the papers. This model is described in more depth in those papers (see also Box-Steffensmeier and Jones (2004: Ch 5)).

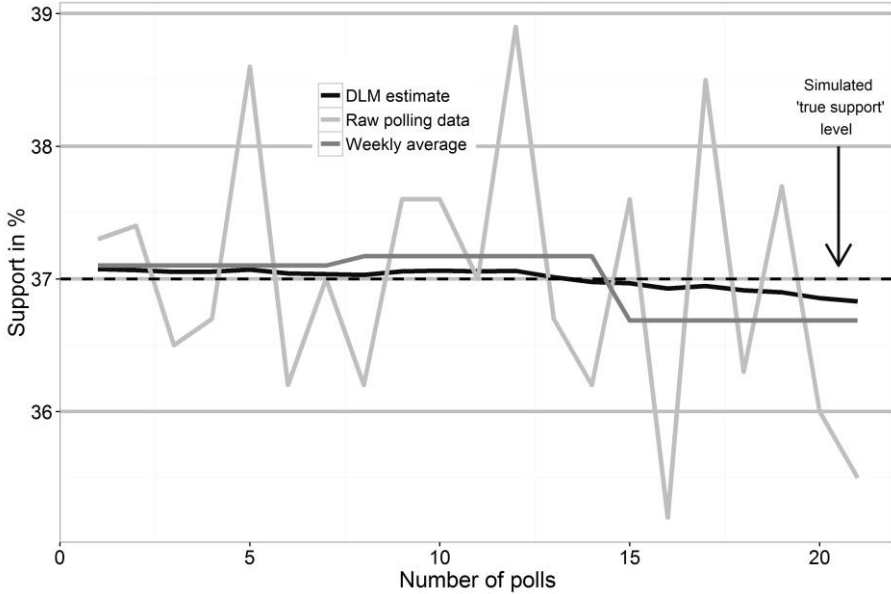
4.2 Polls and cabinet stability

In addition to the Cox model, the other main statistical technique used here is the Dynamic Linear Model, which was used to analyse polling data in Paper 4 and Paper 5. Both of these papers devote substantial attention to the workings of the model, and I will not repeat that discussion here. However, I will provide a brief discussion of the strengths and weaknesses of the model and how some of the limitations were dealt with.

Early in the process of working on this thesis I decided that although government duration would be the main topic, a prominent sub-theme should be polls and government popularity. The fact that previous research in the field had not directly engaged with polling data or incorporated it into the statistical model seemed like a noticeable shortcoming. Since public support is one of the main ‘exogenous shocks’ that Laver & Shepsle (1998) suggest can threaten the government’s position and thus contribute to an increased risk of termination, it seems logical that the empirical models used in the field can be improved by actually measuring it. There are also exciting new techniques available for analysing political polling data (Jackman, 2005; Silver, 2012). A practical concern, however, has been a lack of polling data from many countries. Either there was a complete absence of suitable data or it was sketchy and unreliable. To get reasonably accurate estimates of a government’s popularity, at least a handful of polls per month over the entire term of office are necessary. Fortunately, a large dataset of polling data that fulfils these requirements for many of countries in ERDDA was published recently by Jennings and Wlezien (2014). Their dataset contains more than 26 000 polls from 45 countries. From that group, I selected 12 000 polls that were relevant for the research in this thesis. I also updated the dataset with new data for Sweden.

To analyse this new wealth of data I chose to rely on the Dynamic Linear Model (DLM). In the version used here, the DLM is part of the Bayesian tradition of statistical science (Harrison and West 1997; Jackman 2005, 2009). The model makes it possible to aggregate polling data over time, while giving more weight to polls and polling houses that are judged to be more

Figure 4.1. Comparison between different ways of handling polling data



This comparison relies on simulated data. 21 observations (‘polls’) were simulated, each with mean 37 and a standard deviation of 1. So the closer the line is to 37, the more accurate the technique is.

reliable (Fisher, Ford, Jennings, Pickup, and Wlezien 2011; Linzer 2013; Pickup and Johnston 2007). We can thus create a time trend of estimated party support that is continuously updated and improved when new data becomes available. Figure 4.1 presents a comparison of raw polling data, weekly averages of the polls and the DLM. For illustrative purposes, in this case 21 observations of polling data (i.e. 1 measure per day for 3 weeks) have been *simulated*.⁸ The true support for this hypothetical political party was defined to be 37% with a standard deviation of 1. So, in this case, unlike in reality, we know the actual support for the party (namely 37%) and we can therefore compare how different approaches to polling capture it.

Figure 4.1 shows that even though actual support was held constant at 37%, the raw data (light grey line) jumps quite erratically between 34% and 39%. Random samples of limited size are likely to deviate from the true value, as we can see here. This means that if we only use raw time series, there will appear to be trends and shifts even when the actual popularity of the party is constant.

⁸ The R code for this as well as for all the other analyses, graphs and tables in this thesis is available on GitHub.

Taking weekly averages (dark grey line) gets rid of some of the extreme observations. Still, that approach provides static estimations for each week (even though actual support could change) and also treats each week as a completely new observation, rather than using the knowledge that we already have. The DLM (black line), in contrast, hovers nicely around 37%. The outliers are gone, the trend is continuous, and, most importantly, we are very close to the true level of support. Since we can test the model in these simulated circumstances⁹ and confirm its accuracy, we can also be confident in its application in real world circumstances.

Thus, when we have sufficient data, there seems to be little reason to doubt the reliability and usefulness of the DLM. Instead I had other concerns when combining this model (as developed in Paper 4) with duration analysis in order to measure how party popularity affects government stability (Paper 5). The first and main problem was data scarcity. Although 12 000 polls sound like a lot and is a clear improvement over previously existing polling datasets, the polls are not evenly distributed. Some countries, e.g. Germany, the UK and Sweden, are very well covered, whereas many of the countries in ERDDA are not covered at all. All of Eastern Europe is missing, and many countries in Southern Europe, e.g. France, Italy and Greece, are also insufficiently covered. This means that northwest Europe is clearly overrepresented, which can be problematic since we know that this region stands out as regards cabinet stability. Most of the governments included in the dataset were also in power fairly recently, since the availability of polling data has improved over time. This means that even though about 160 governments were included in the study, there are reasons to doubt the generalizability of the findings. In particular, one of the stated goals of the paper is to test popular theoretical arguments in the field and see how other explanatory variables are affected when we control for popularity. However, to fully engage with the existing literature, the sample must be extended to match the one used in the other studies. This should be a prime concern in future research in this area.

A second concern with this approach is more theoretical – how do decision makers take polling into account, and when does it directly affect their cabinet participation? Is it when they foresee that their support will decline in the future, when they have suffered a persistent slump (i.e. one that lasts for a long time) or when support shifts rapidly enough to shock them out of their equilibrium? Or perhaps it is enough to simply track the regular daily trends? Put differently, do parties look primarily at the past, future or the present when trying to decide how their fate in the polls should affect their cabinet membership? How one answers this question about the causal mechanism has a significant effect on statistical operationalisations. Before finally settling on

⁹ We can, of course, also test it on real-world elections, which was done with good result in Paper 4

the model used in Paper 5, I tried all of the above options as well as a few more. This process led to the finding that the standard daily estimate, i.e. the present situation, of the DLM model provides the best statistical fit. This is interesting, since it suggests that parties primarily look at their current situation, rather than retrospective appraisals or trying to forecast future developments.

The theoretical challenge about how parties react to changes in popular support was therefore dealt with using a statistical approach. Although this might seem overly inductive, in practice political science theory is often too imprecise to provide us with detailed guidance about how actors should behave in intricate situations. Even if we agree, as argued in section 3.2, that parties have certain predefined goals, it is often not immediately obvious how they will pursue them in complex and changing circumstances. In such cases, I often found it useful to use empirical trends to distinguish between contending and equally plausible theoretical options.

To round off this chapter on methodology, it should be stressed that although this thesis might come across as heavily focused on methods, the statistical techniques used and developed here have always been seen as a means to a theoretical end. The thesis has been driven by a substantive interest in explaining government decisions to remain in power or to pass the buck to someone else. However, to answer some of the questions that arose in that process, it was necessary to think carefully about the methodological approaches chosen and, in the case of polling data, to develop new techniques to get more information out of existing materials. This has meant that significant time in the past 4 years has been spent learning and working with statistical techniques. My hope is that the thesis has become theoretically more interesting as a result.

5 Summary of the articles

5.1 Paper 1: Cabinet survival in Central and Eastern Europe – what do we know after 25 years of democracy?

Under review

In this study, the main goal was to explore what we can say about how and when early government dissolution happens in Central and Eastern Europe (CEE). Most of the early studies in the field of government stability, as well as most of the theoretical claims, focused on Western Europe. With the sudden democratization of the countries in CEE, there was a new and geographically adjacent sample on which to test established findings.

Before this study was conducted, 5 articles had been published that focused explicitly on government survival in CEE (Bergman, Ersson, and Hellström 2015; Grotz and Weber 2012; Savage 2013; Somer-Topcu and Williams 2008; Tzelgov 2011). The studies incorporated both ‘conventional variables’ that had proved to be influential in Western Europe and new variables specifically tailored to the political dynamics in CEE. New variables included whether a communist successor party (i.e. a revamped version of the old ruling communist party) was in cabinet, how much governing experience the cabinet members had, as well as whether and, if so, when the country had joined the EU (Tzelgov, 2011; Grotz & Weber, 2012). Unfortunately, the papers came to radically different conclusions. Over 25 variables were tested, yet only 2 were found to matter in all studies: growth in GDP per capita and the presence of a communist successor party in the cabinet. Every other factor that was determined to be a statistically significant predictor in one study was disconfirmed in at least one of the others.

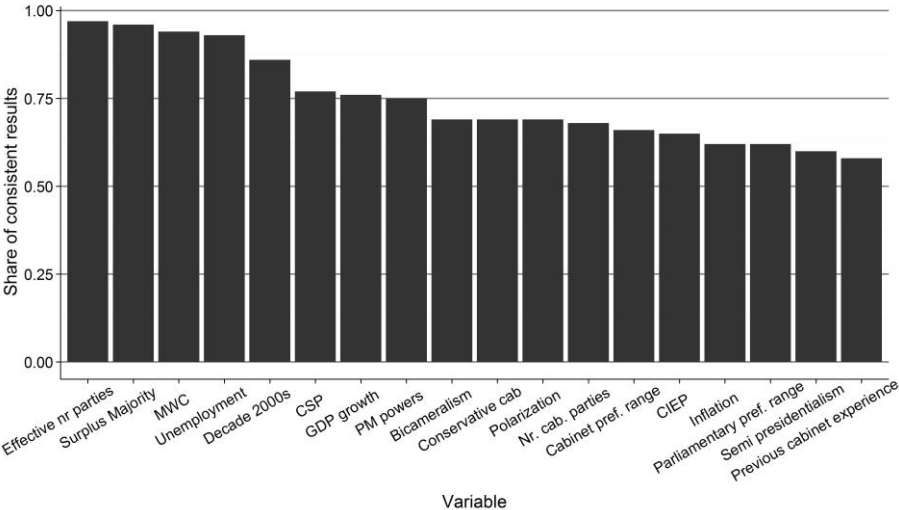
After doing an analysis of statistical power (Finlay and Agresti 2009, 169), I found that one of the main reasons for the inconsistency in results was that the sample of included countries was small and highly sensitive to changes in model specifications. Fewer than 100 governments were included in most studies, and when 15-20 variables are tested, the ratio between observations and variables is too small to get reliable findings. In particular, if the model is changed slightly by the removal or inclusion of a single variable, then the results can be noticeably different.

This led to the conclusion that new theoretical and empirical approaches were needed in order to generate findings that would be stable and reproducible in future studies. First, as far as theory is concerned, it seemed reasonable that, as a compliment to findings in the duration literature, we should also look at in-depth research on the CEE political environments (Bielasiak 2005; Kitschelt 1995; Roberts 2008) and at the more general field

of democratization studies (Lijphart 1968; Randall and Svåsand 2002; Teorell 2010). Based on insights from these different subfields, I could develop a scaled down set of predictions about what should be of particular importance for cabinet stability in CEE. In particular, the paper argues that in new democracies, voters are more concerned with effective management and tangible results and less with ideological considerations. High voter volatility also means that competition between parties is more prevalent. This suggests that there is reason to believe that variables that tap into formal power, party competition and economic performance are likely to be of particular relevance here, while those concerned with ideology or institutions should matter less. These conclusions are similar to the ones that Döring and Hellström (2013) reached as regards what influences government *formation* in CEE.

Rather than running a standard duration model to test those expectations, I used a technique that is more concerned with robustness of results. Specifically, Extreme Bounds Analysis (EBA), or Sensitivity Analysis, is a statistical technique that is designed to test how robust results are to changes in model specifications (Hlavac 2015; Leamer 1985; Sala-i-Martin 1997). Given a list of variables of interest, EBA is built on the idea of running many regression analyses, in which different combinations of those variables are included. For example, say that an analyst wants to test a set of 15 variables. Only a subset of the 15 are selected for inclusion in any particular regression, which means that every individual variable is included in a wide range of

Figure 5.1. Consistency of predictors of cabinet stability in CEE



The bars show how often a particular variable was significant in all the regressions run in the Extreme Bounds Analysis

different regression analyses with a changing set of other variables. A variable that has stable effects is one that continues to influence the likelihood of government termination in a similar way, regardless of which other variables are included in the model.

Some of the main findings from the EBA are reported in 5.1. Here we see that, in general, the variables that have consistent effects do, indeed, seem to tap into the main dimensions identified in the theoretical section. The effective number of parties and government type are the most reliable predictors, and they also seem to have a strong substantive effect. In all of the regressions in which it was included, effective number of parties has an average odds ratio of 1.22, meaning that for each extra party in the system, early government termination becomes 22% more likely (other variables held constant at their mean values). Effective number of parties is theoretically relevant since it captures the level of competition between parties in the system. Economic performance, measured here by the unemployment level, has a consistent effect in 93% of regressions, suggesting that it also impacts on the stability of most governments in CEE. This is in line with the expectation that tangible results and material improvements should matter. As predicted, factors related to ideology and institutions were less systematically relevant, since those issue dimensions do not become relevant until the party system has stabilized.

5.2 Paper 2: Government instability and the state

Under review

The starting point for this study, which was co-authored with Professor Torbjörn Bergman, was the idea that the state – its operation and competences – should be brought into duration research. In the field of cabinet duration studies, a wide range of factors that might influence government longevity have been tested, including indicators of government type, political institutions, socioeconomic development and the competition the government faces from other parties. However, the state as such has escaped attention. This can be seen as a theoretical oversight, insofar as the capacity and reliability of the public sector have a strong impact on other factors that we know influence government stability, e.g. the ability to achieve policy goals (Dahlström, Lindvall, and Rothstein 2013).

In addition to this theoretical starting point, the paper also draws inspiration from the empirical observation that there are consistent and predictable mistakes in the models that are commonly used in comparative duration research. In particular, our models consistently overestimate the duration of governments in some countries, and repeatedly underestimate the life of governments in others. That is, after we control for all the usual factors that we know influence cabinet stability, there are still predictable

discrepancies between the expectations from our models and real-life government duration. Moreover, these deviations are structured on a country level. This suggests that there are some factors that set countries apart and that make governments within them more or less prone to early dissolution irrespective of their other attributes. Figure 2.1 in Chapter 2 illustrates this point.

Paper 2 argues that state capacity and the effectiveness of public sector management might explain this. In particular, we argue that because it is more difficult to successfully introduce and implement new policy proposals when the bureaucracy is incompetent or corrupt, this negatively affects the stability of the government in two ways. First, the value of holding government power is lower when the government struggles to introduce policy that both it and its voters favour (Dodd 1976). Failing to implement the policy that it was elected to bring about can be electorally costly to the government, which goes against the main goal of political parties. Second, if policy implementation is difficult, then this hampers the government's ability to deal with sudden events that require decisive action (the external shocks that Laver and Shepsle (1998) talked about). This means that governments in countries with low state capacity find it more difficult to adequately respond to such events.

As an indicator of state capacity, we relied on a measure of quality of governance that measures impartiality in the conduct of public sector operations and efficiency in bureaucratic procedures (Rothstein 2011; Rothstein and Teorell 2008). The measure we used was initially developed by the International Country Risk Guide,¹⁰ and it has been extensively tested in the Quality of Governance (QoG) field. With its focus on competence and reliability, it nicely captures our theoretical intuition about how state capacity should affect government stability.

To test the relevance of state capacity, we included the indicator in our statistical models in two ways. It was entered on its own in one model, in order to test the direct, independent effect, and it was entered in a second model in an interactive relationship with unemployment. If our causal argument is correct – that governments in countries with low-state capacity are less well equipped to deal with external shocks – then these governments should find it more difficult to deal with an increase in unemployment. Unemployment, which we measured on a quarterly basis (i.e. four measurement points per year), thus acts as an example of an economic shock that leads to a demand for new policy.

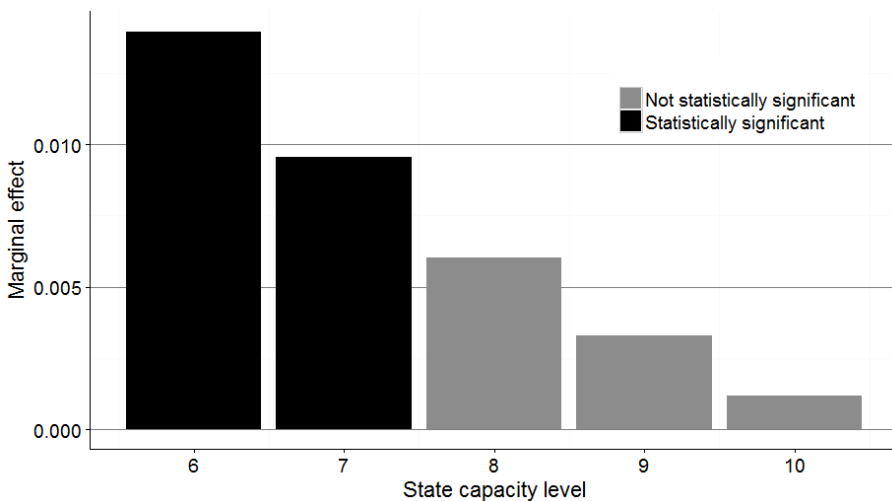
Using a discrete-time model, state capacity was found to matter for cabinet stability in our sample of 28 countries. State capacity had a strong independent effect on the likelihood of termination through a replacement,

¹⁰ We relied on the version of the measure that is provided by the Quality of Government Institute at <http://qog.pol.gu.se/>.

but seemingly no effect on the propensity to call an early election. When there was no interaction term, state capacity was found to have an odds ratio of 0.784 for replacements, which means that for each extra level of state capacity on our ten-point scale, the likelihood of ending through a replacement declines by 21.6%.

In the second model, where we added an interaction term composed of state capacity and unemployment, the results are a bit more complicated to interpret, because the effect of one variable is now conditionally dependent on the level of the other (Brambor, Clark, and Golder 2006). Therefore, in the paper we used an interaction plot, similar to Figure 5.2, to determine how the effect of unemployment varies depending on state capacity. The figure shows the effect of unemployment for different levels of state capacity. The results confirm our theoretical expectations – i.e. unemployment has a statistically significant effect on the likelihood of early dissolution (in the form of replacements) only when state capacity is low. To be precise, unemployment appears to matter when state capacity is lower than 8 on our 10-point scale, but has no statistically significant effect above that. This suggests that an increase in unemployment in a country with sufficient bureaucratic capacity to respond (e.g. the Nordic countries), does not have a meaningful effect on stability. On the other hand, the marginal effect when state capacity is six is about 0.014, which means that for each extra point that unemployment increases during a given quarter, the likelihood that the government ends through a replacement increases by 1.4%.

Figure 5.2. Marginal effect of unemployment for different values of state capacity



Bars in black are statistically significant at the 95% level.

The two main takeaways from this paper are, first, that the state in general, and the competence and efficiency of the bureaucracy in particular, should perhaps be taken more seriously in duration research than it has been in the past. Second, some explanatory factors that are often found to matter for duration (such as unemployment) might, in fact, have a conditional effect that depends on other qualities and competences, including the capacity to deal with them. To really capture the causal dynamics at work, it might be worthwhile to consider interaction effects and mediating variables in additional scenarios.

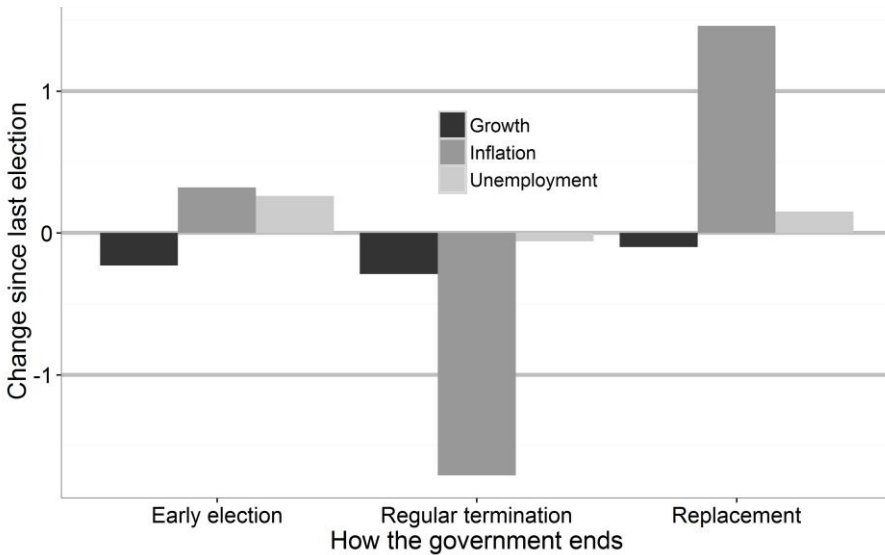
5.3 Paper 3: How is government termination affected by the state of economy? Payoff structures, type of government and economic changes

Accepted for publication in Government & Opposition, January 2017

Economic factors, e.g. GDP growth, inflation and unemployment, are frequently included in studies of cabinet stability (see Warwick (1994); Saalfeld (2013) for longer treatments). Such economic developments, particularly inflation and unemployment, are often found to have significant effects on the likelihood of early termination, both early elections and replacements. There are various theoretical explanations for this. Incumbent parties can use the economic situation to evaluate their future electoral prospects. Since voters care about the economy, a strong one might give the government an incentive to opportunistically schedule an early election (when the institutional setting allows for it) (Damgaard 2008; Kayser 2005). In contrast, a failing economy might contribute to infighting and blame shifting among coalition partners, which, in turn, might contribute to the dissolution of the cabinet (Lupia & Strøm, 1995).

Some descriptive support for the proposition that the economy matters can be found in Figure 5.3. Here we see the predicted pattern for both unemployment and inflation. That is, if we measure how economic circumstances changed between the time that a government assumed office and when it left (either by a premature dissolution or a regular, non-discretionary termination), we find that both unemployment and inflation tend to increase in cases of early termination. In contrast, when the government manages to survive to the next regular election, both unemployment and inflation tend to decrease – quite strongly as regards

Figure 5.3. Change in economic circumstances for governments that terminate prematurely and ones that do not



The bars show how things have changed since the government took office. So a positive bar e.g. for unemployment means that unemployment has increased whereas a bar pointing downwards means that unemployment has decreased since the election.

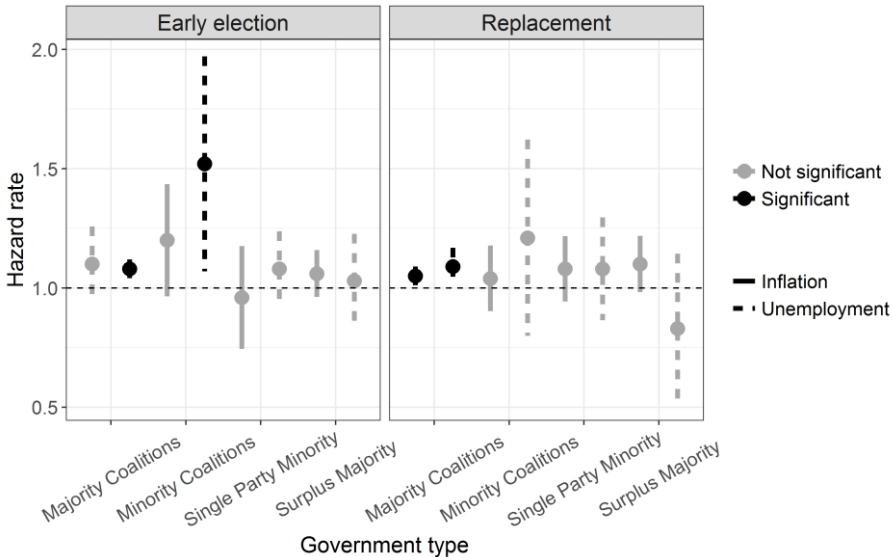
inflation. There is no clear pattern for growth. This pattern, together with the findings of previous research, led us to focus on unemployment and inflation as indicators of the state of the economy in this paper.

Although the relevance of economic factors constitutes a well-established finding, few attempts have been made to understand whether and, if so, how the type of government in power influences the impact of economic change on duration (for an exception see Robertson (1983)). In particular, we can hypothesize that if the causal mechanisms outlined in the previous section are correct, the precise effects of changes in the economy should depend on the number of parties in government. If one of the impediments to quickly and decisively responding to economic difficulties is that coalition partners disagree about the best way to address the problem, then this ought to be less problematic for single party governments. Coalition governments often have formal or informal agreements that, in broad strokes, define what types of policies will be pursued. However, if the external economic situation changes drastically, these pre-existing deals must be renegotiated. This can lead to disagreement about the best way to change existing policies, and crafting a response is also likely to be a more time-consuming process for a coalition government than for a single party government.

The paper develops a 2-by-2 typology of government types (minority coalition, single party minority, majority coalition, single party majority) and tests whether these different types respond in different ways to the same underlying changes in the economy. We first run a standard baseline model with all types pooled into one sample. This confirms previous findings, i.e. that both unemployment and inflation have statistically and substantively significant effects on the likelihood of both early elections and replacements. We then use a split sample approach and run separate Cox proportional hazards models for the four different types.

The empirical findings are summarized in Figure 5.4, and they confirm our main expectations. Single party governments appear to be largely unfazed in the face of economic turbulence, while coalition governments, both minority and majority types, are more likely to terminate early when faced with an economy in decline. A second hypothesis in the paper was that minority government would be more strongly impacted in substantive terms, i.e. not just in terms of statistical significance but also in terms of the size of the effects. The analysis also supports this hypothesis, since the strongest substantive effects by far were for minority coalitions. For example, the baseline risk of early elections for minority coalitions increases by 52% for

Figure 5.4. Main results for how economic effects depend on government type



Effects in black are significant at the 95% level. Dashed lines measure unemployment, solid lines inflation.

each extra percentage point that unemployment increases during the government's term of office.

The main goal of this paper was to qualify our understanding of the relevance of economic changes and to pinpoint what drives the aggregate results. As will be discussed in the final chapter, this is one of the directions in which the field of cabinet studies is changing. That is, it is moving away from general, aggregate results and focusing more attention on detailed findings that explain how the main effect of a variable depends on and is modified by different contextual factors.

5.4 Paper 4: Picking the winner(s): Forecasting elections in multiparty systems

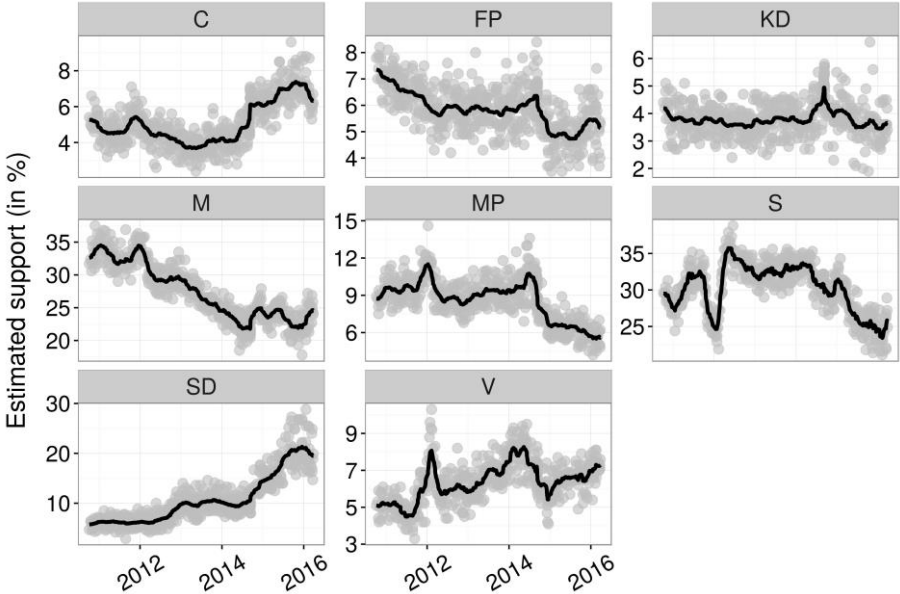
Published in Electoral Studies, December 2015

This paper is an extended attempt to answer the question of whether it is possible to predict with reasonable accuracy the outcome of elections in complex systems where many different parties are competing for power. Election forecasting is an interesting business, because it is one of very few situations where we eventually get a perfectly unbiased measure, i.e. the general election. By developing different techniques to aggregate and analyse polling data, we can compare different predictions with actual election results in order to determine how well they perform. Such tests are difficult to do for other types of polling or survey research, where data from the whole population, rather than just a sample, is hard to come by.

In the field of election forecasting, focus has traditionally been on stable two-party systems such as the US and the UK. One popular forecasting technique relies on OLS regression and uses economic and political indicators to try to predict election results (Abramowitz 2008; Bartels and Zaller 2001; Foucault and Nadeau 2012). For example, incumbent governments are generally punished slightly for being in power, but if the country is doing well economically and the president or PM is well liked, then a government might still end up getting a net gain in votes in the subsequent election. Using data about exactly how much the economy has improved or how well liked the leader is, we can try to estimate the government's election result. Another popular approach has been to model government support as a time series built from polling data. Rather than looking at external political and economic developments, this approach uses polling results and tries to find accurate ways to aggregate data of varying quality from different polling houses.

Although regression techniques have worked well in two party systems, they are likely to be ineffective in more complex parliamentary settings where a wide range of parties are competing. If regression models help us predict the result for the government party, in a two-party system that is sufficient to give us the result of both parties, because the opposition party gets 1 minus the vote share of the government party. However, this does not work in multiparty systems, where there are many parties in opposition. Also, if the government consists of a coalition, it is difficult to know which – if any – of the parties will be punished or rewarded for socio-economic developments during the government’s term of office. Thus, in multiparty systems we must primarily rely on a time series approach. Based on previous studies (Fisher, Ford, Jennings, Pickup, and Wlezien 2011; Jackman 2005; Pickup and Johnston 2007), the DLM presented in section 4.2 seems like a suitable choice. The advantages of this model have already been discussed in some detail, and they will not be repeated here. However, a brief overview of the technique is shown in Figure 5.5. Here we can see how the DLM works to create a joint trend for each party based on polling data from a variety of polling houses. The DLM nicely synthesizes the information in the data, while avoiding the highs and lows associated with individual polls.

Figure 5.5. How the DLM estimates the development of party support in Sweden between 2010 and 2016



The 8 parties are the ones that have parliamentary representation in the 2014-2018 Swedish parliament. The grey dots are individual polls.

This paper makes two key contributions. The first is to demonstrate how well the model can perform when applied in two complex, multiparty systems – i.e. Germany’s and Sweden’s. Predictions were made for three elections per country, based on all available polls up to the day before the elections. In both countries, the model worked well. In Germany, the average prediction error in the three elections is 0.69 percentage points, and in Sweden it is 0.78. These results are as good or better than what many previous studies have achieved in the more easily predicted two party systems. However, there are some differences in the accuracy of the model for some elections and parties. Support for some parties, such as the Green Party in Sweden, tends to be overestimated by polls. Support for others, e.g. the populist right Sweden Democrats, are underestimated. Nonetheless, on balance, the forecasts must be seen as reasonably successful.

The downside of the method is that the election forecast was made at a very late stage, that is, the day before the election. In practice, lead time is very important, and, ideally, we want to be able to predict something meaningful about the results of an election a lot earlier than that. For this reason, the standard DLM model was extended with a *seasonal component*. A seasonal component is often used in economic time series analysis to account for re-occurring trends. For example, consumption tends to increase just before Christmas, so if we want to know how the base level of consumption is changing we must control for the general surge associated with Christmas. And during summer, we have a predictable slump in economic activity. There are techniques for analysing time series that directly model such seasonal shifts, which makes it possible to uncover an underlying temporal trend.

Such techniques have not previously been used in political polling. However, political parties are also subject to seasonal trends. Government parties tend to lose support in the middle of their term of office, only to regain some or all of it before the next election. In addition, support for some small parties tends to be just below the parliamentary threshold up to the final weeks preceding the election, at which point some voters decide to support them in order to ensure that they are not completely eliminated (Fredén 2014).

The DLM can be extended to model these developments. In Paper 4, ARIMA models were used to empirically infer the seasonal trend for each party in two-party systems. By incorporating this control for seasonality, the prediction error one month before the election was lowered by 17% in the Swedish case and 8% in Germany, as compared to just using the standard DLM. Thus, this paper shows that we can use polling data to get reasonable estimates of political parties’ levels of support. This is good news for cabinet duration analysis, since we now have a reliable tool for taking on the long-standing question of how popularity affects a government’s chances of survival.

5.5 **Paper 5: The verdict in the polls – how government stability is affected by popular support**

Under review

In the fifth and final paper, the main ambition was to tie together the techniques for estimating popular support that were developed in Paper 4 with cabinet duration analysis. Previous studies have not directly modelled government popularity, which seems like a lacuna that needs to be addressed for both empirical and theoretical reasons. First, popularity is a core component in formal models like the ones developed by Lupia and Strøm (1995) and Laver and Shepsle (1998). In Laver and Shepsle's model, public opinions shocks are one of the main external developments capable of weakening cabinet equilibrium. Since governments are generally stable at the outset, some type of shock is needed to make parties re-evaluate their membership in them (Diermeier 2006). Changes in potential fortunes are key among these shocks, because, as Müller and Strøm (1999) have argued, the goals of being in power and influencing policy must sometimes be sacrificed in order to ensure long-term electoral survival.

In addition to such theoretical reasons for focusing on popularity, there are also many empirical ones. For example, a common explanation about why a high inflation rate negatively affects a government's chances of survival is that it threatens to hurt its performance in the next election. However, if inflation leads to a decline in popular support for the government, we should be able to see this directly in the polls rather than taking the unnecessary detour of inferring level of popular support based on the inflation rate. Similarly, a critical event such as a corruption scandal might propel the government's downfall, but if it does, it should be because the government has lost support and is afraid of a strong electoral backlash. Thus, for a variety of reasons, directly measuring the effects of changes in popularity is a logical next step in the development of the field.

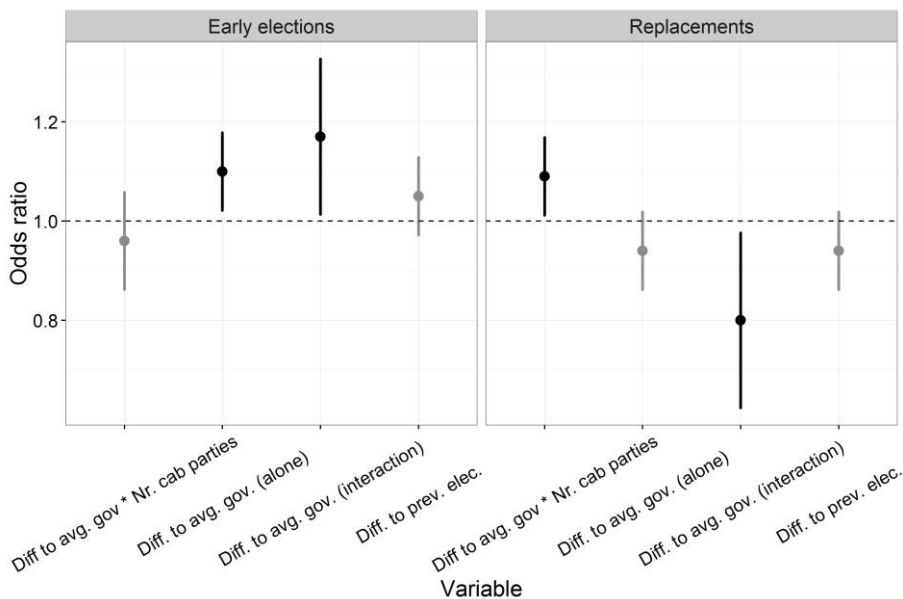
In order to estimate popular support, the DLM model used and extended in Paper 4 was applied to the Jennings and Wlezien (2014) dataset discussed previously. As noted, the subset of the dataset used here contains 12 000 unique polls on more than 160 cabinets, and, after running the models, the obtained estimations of party popularity were merged with data from ERDDA. Theoretically, the expectation was that popularity would matter for cabinet stability, but that there would be an interactive relationship. In particular, since popular support might develop differently for the individual members of a coalition government, a party that governs alone should find it easier to strategically time a response to changes in popular support. This hypothesis is similar to the one in Paper 3 where it was argued that responses to changes in economic circumstances are easier to devise and implement in a single party

government. Thus, the precise effect of a change in popularity should depend on the number of parties in power.

Another theoretical concern in the paper was how to operationalize support. If governments are influenced by popularity, are they influenced by absolute level of support (e.g. 48%) or by changes since the last election (e.g. a loss of 3 percentage points)? As it turns out, most governments lose support when they are in office. Halfway through its term of office, the average government polls at about 5 percentage points lower than the election result. However, part of this is recovered in the run-up to the next election. Since loss of public support compared to the last election fluctuates significantly as part of the political business cycle, this paper hypothesizes that polling matters when government support differs significantly from the general ups and downs that all governments tend to experience. When polling is significantly better, opportunistic early elections might be warranted, and when polling is worse, the government opts for a replacement (to give it time to recapture some support before the election is held).

As can be seen in Figure 5.6, total change in support compared to the last election appears not to matter. It is not significant in any of our models. However, the difference between a particular government and the average government does matter. This is calculated based on the month of the term of

Figure 5.6. Different ways of measuring popular support and how it impacts on government duration



Black lines are significant at 95% level.

office. For example, say the average government polls at -3,5 one year after the election. If a particular government has lost 5 percentage points at that stage, then it is at -1.5. Figure 5.6 shows that when this value is positive – i.e. when the government is polling at better than average – the likelihood of early elections increases. Conversely, when support is lower than it is for the average government (i.e. the value is negative), the risk of replacement rises, but only after we control for number of parties in the government.

In fact, further analysis of the interaction effects reveals that for both termination types, the findings are statistically significant only when there is just one party in government. That is, popularity seems to affect government stability only in the case of single party governments. This confirms our expectation that although all types of governments care about polling, it is easier to act on it strategically when there are no coalition partners to take into consideration.

The main goal of this paper was to test the often-invoked argument that popularity matters for cabinet stability, and this general intuition appears to hold. However, the paper is only a first attempt to formally include popularity in duration research, so the results must be seen as tentative. Only 160 cabinets were covered in the polling data (compared to the 640 included in the ERDDA), and there were no governments from Central and Eastern Europe. More data covering additional countries and a longer time frame might enable us generate more precise findings and lead to greater confidence in the results. We should also think more carefully about how popular support covaries with other commonly used explanatory variables, e.g. unemployment and inflation, and be more careful about how we model the interactive relationship between those explanatory factors.

6 Where do we go from here?

To round off this four-year effort to understand government instability, let me briefly outline what I see as some promising avenues for further progress in the field. Most of these suggestions deal with methods and data, but all of them are inspired by a desire to ensure that our empirical studies come closer to the theoretical intuitions that inspire them.

One general ambition should be to gradually move away from overarching statistical correlations and towards more precise and theoretically interesting causal identification (Angrist, Imbens, and Rubin 1996). Most of the variables used in the field's popular statistical models have a significant impact on cabinet stability under some circumstances, but those circumstances are not usually directly tested. For example, having a second parliamentary chamber is often found to be detrimental to cabinet stability (e.g. Saalfeld, 2008). Theoretically, we would expect this effect to occur primarily when the government is in a minority position, and lacks a reliable support party, in the second chamber. If the government controls a reliable majority in the second chamber, then the impact should be small. In general, however, this causal pathway is not what the models actually test. Rather, in most cases bicameralism is simply included as a static control variable with little connection to the actually posited causal mechanism.¹¹

Similarly, unemployment is one of the most reliably influential predictors of early termination, but, as shown in some of the papers here, its effect depends on other factors such as the capacity of the state to handle it and the type of government in power. In that sense, the estimated average effect of an increase in unemployment is not particularly illuminating, because it masks the different situations in which governments find themselves. For an ideologically unified single party government operating in a well-functioning country, the actual effect of increased unemployment on the likelihood of early termination is probably close to zero. For governments in particularly dire circumstances, the effects of the same change could be devastating. The price we pay for not carefully modelling the actual causal pathway is that we get a watered down average estimate that often says precious little about the specific circumstances of a particular government.

Thinking more carefully about conditional effects and contextual conditions and then actually testing them empirically is one way to improve our research. In the same vein, time and the temporal dependence of most effects should also be explicitly modelled. For example, being in a coalition government is likely to be less of a problem in the early days of the term of office, because coalition deals are quite fresh and the next election is far off (Diermeier and

¹¹ For an important exception, see Druckman & Thies (2002). See also Druckman et al (2005) and Eppner and Ganghof (2015) for similar discussions in the case of cabinet formation.

Stevenson 2000). However, as time passes and external shocks force cabinet members to re-evaluate their policy agenda (Laver and Shepsle 1998), and when the need to distinguish oneself from the crowd before the next election starts to grow, a coalition government should become prone to dissolution in ways that a single-party government does not.

A second, but related, suggestion about what to do next is the mandatory call for more and better data. In particular, everything that changes during the term of office should be measured as a time-varying covariate. This includes factors such as all economic indicators (unemployment, inflation, growth), government popularity, changes in the party system and bargaining environment (e.g. party ideology could change) and relevant external shocks (including general ones such as the 2008 financial crisis or the 2015 migrant crisis, as well as crises affecting only certain countries). This is a salient point because most of our formal models are built on the assumption that the government is in equilibrium when it assumes office (Diermeier, 2006). If we do not measure how the explanatory variables change after this initial equilibrium, we are hard pressed to determine exactly why the government was displaced from its original, stable position. When static variables are used, we can draw conclusions about which variables seem to be associated with a higher risk of a disturbed equilibrium, but not about when it happens or how much those variables need to change.

To make this point more explicit, one option would be to refrain from using governments as the unit of analysis in a dataset, and use government-year or government-half year instead.¹² Building time directly into the construction of the dataset would make it easier to update variables that change during the term of office. In addition, we could measure whether and how the impact of static variables (such as government type) changes depending on how much time is left in the government's term of office. This approach would also save us the trouble of having to include time as an independent variable (which is sometimes done when time-dependent effects are discovered). Fixed effects or shared frailty (Kleinbaum and Klein 2005) could be used to control for the fact that the same government is measured on multiple occasions.

Data and operationalisation strategies could also be improved by using more detailed measures of the different types of termination. For example, there have been some recent attempts to measure and test the idea of 'opportunistic election calling' (Riera 2015; Schleiter and Tavits 2016), where governments call early elections not out of necessity, but out of a desire to reap the benefits of a promising situation. As Paper 5 showed, there are good reasons to believe that when the electoral outlook for the government is promising and the institutional setting allows it, government parties might

¹² This is offered in one version of the Seki & Williams update of the Party Government Dataset (Seki and Williams 2014).

find it in their interest to opportunistically call early elections. Since these situations are likely to be quite different from ones in which governments hold early elections out of necessity, treating them all as the same type of termination is likely to give mixed results and obscure the actual causal pathway.

In a similar vein, what we now call 'replacement terminations' is a broad category that contains dissolutions of quite different scopes. Replacement terminations include all changes in the composition of the government, which means that one party joining the government is coded the same as the entire government resigning. A minor coalition member leaving a surplus majority government is coded the same as the resignation of the only party in a single party government. In fact, it is likely that these different ways to reconstitute a government are actually theoretically distinct. That is, they might have different immediate causes, and the parties might have different incentives and opportunity structures. Greater changes in the initial equilibrium are probably required to prompt a single party government to give up its mandate, as compared to changes sufficient to prompt a smaller party to give up the limited influence it has in a surplus majority government. If this is true, it might also be necessary – depending on the research question – to separate instances of replacement into different types in order to get a clearer understanding of which explanatory factors increase the risk of which particular type or types of replacement termination.

The overarching concern that arises from vague causal identification, time-invariant operationalisation and the aggregation of theoretically distinct termination types is that the results of our analyses are more muddled and less meaningful than they could be. There are some seemingly inescapable problems inherent in our field, e.g. too few data points compared to the number of interesting variables and the difficulty associated with operationalising abstract concepts like 'culture' and 'political tradition'. In contrast, the problems discussed above can actually be addressed, and we would likely get sharper and more interesting findings as a result.

Having said that, it is worth reiterating a point made by Laver (2003). He notes that despite all the improvements in data and statistical techniques that occurred during the field's first three decades, the explanatory factors that were found to reliably matter remained more or less the same. The signs of the coefficients and the asterisks indicating statistical significance shifted only marginally, which suggests that having the correct theoretical set-up matters more than the precise statistical techniques used. For that reason, it is possible that even if we implemented all the changes suggested above, we might still end up with more of a moderate refinement than a radical overhaul.

Nonetheless, many of the suggestions discussed here are more than simple changes in statistical modelling or operationalisation. The goal is to go beyond the ability to say whether and how much particular variables matter, and to

explain when and why as well – i.e. in conjunction with which other variables and under which circumstances particular variables matter. The hope is that this will make our models more realistic and better positioned to say something meaningful about the actual situation a particular government finds itself in.

Finally, in addition to the points raised above, another interesting research project would be to try to bridge the gap between governments, governance and political reforms. For example, investigating the relationship between government type and legislative output (Cheibub, Przeworski, and Saiegh 2004; Martin and Vanberg 2011, 2014) might reveal more about the impact of coalition dynamics on legislative decision-making. Under what circumstances are coalition governments less successful at getting bills passed than single party governments? How similar are the voting patterns of the different parties in government, and do government parties that display more disparate voting patterns find it more difficult to work together (and thus to survive until the next election)? Another interesting question is whether important legislative reform, e.g. fundamental changes in the way the welfare state works, tends to be undertaken by certain types of governments under particular circumstances. That is, are such ambitious reforms primarily driven by external needs (in which case the legislature considers it necessary to approve the changes regardless of which government is in power), or are governments that are composed of particular parties more likely to pursue policy changes that have far-reaching impacts? Combining data on governments, ministers, legislative output and welfare state reforms might thus yield pertinent insights into the interplay between government and governance. Some promising studies have taken initial steps in this direction by bringing together the insights of research into policy output, political institutions and coalition governments (see e.g. Bäck and Lindvall 2015; Häusermann, Picot, and Geering 2013; Knotz and Lindvall 2015).

This final point brings us back to where we started with Figure 1.1 and the coalition life cycle. Elections, formation, governance, and termination are inextricably linked and giving governance and policy output more attention in coalition research seems like a promising avenue for further research. In the process, it might also answer some outstanding questions about the messy workings of policy-making in real-world democracies.

7 Bibliography

- Abramowitz, Alan I. 2008. "Forecasting the 2008 presidential election with the time-for-change model." *PS: Political Science & Politics* 41(04): 691–695.
- Alesina, Alberto, Sule Özler, Nouriel Roubini, and Phillip Swagel. 1996. "Political instability and economic growth." *Journal of Economic Growth* 1(2): 189–211.
- Allison, Paul D. 1995. "Survival analysis using SAS: A practical guide. SAS Institute." *Inc., Cary, NC.*
- Allum, Peter Antony. 1974. "Italy—Republic without Government?" *The World Today* 30(11): 448–459.
- Andersen, David, Jørgen Møller, and Svend-Erik Skaaning. 2014. "The state-democracy nexus: conceptual distinctions, theoretical perspectives, and comparative approaches." *Democratization* 21(7): 1203–1220.
- Anderson, Christopher J, and Yuliya V Tverdova. 2003. "Corruption, political allegiances, and attitudes toward government in contemporary democracies." *American Journal of Political Science* 47(1): 91–109.
- Andersson, Staffan, Torbjörn Bergman, and Svante Ersson. 2014. *The European Representative Democracy Data Archive, Release 3. Main sponsor: Riksbankens Jubileumsfond (In2007-0149:1-E) [www.erdda.se].*
- Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American statistical Association* 91(434): 444–455.
- Axelrod, Robert M. 1970. *Conflict of interest: A theory of divergent goals with applications to politics.* Markham Publishing Company, Chicago.
- Bäck, Hanna, and Johannes Lindvall. 2015. "Commitment Problems in Coalitions: A New Look at the Fiscal Policies Of Multiparty Governments." *Political Science Research and Methods* 3(01): 53–72.
- Bartels, Larry M, and John Zaller. 2001. "Presidential vote models: A recount." *Political Science & Politics* 34(01): 9–20.
- Bergman, Torbjörn, Svante Ersson, and Johan Hellström. 2015. "Government formation and breakdown in Western and Central Eastern Europe." *Comparative European Politics* 13(3): 345–375.
- Bernal-Verdugo, Lorenzo E, Davide Furceri, and Dominique M Guillaume. 2013. "The Dynamic Effect of Social and Political Instability on Output: The Role of Reforms." *IMF Working Paper.*
- Bielasiak, Jack. 2005. "Party competition in emerging democracies: representation and effectiveness in post-communism and beyond." *Democratization* 12(3): 331–356.
- Box, George EP, and Norman R Draper. 1987. *Empirical model-building and response surfaces.* New York: John Wiley & Sons.
- Box-Steffensmeier, Janet M, and Bradford S Jones. 2004. *Event history modeling: A guide for social scientists.* New York: Cambridge University Press.

- Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding interaction models: Improving empirical analyses." *Political analysis* 14(1): 63–82.
- Browne, Eric C, and John Dreijmanis. 1982. *Government coalitions in Western democracies*. London: Longman Publishing Group.
- Browne, Eric C, John P Frenreis, and Dennis W Gleiber. 1984. "An events approach to the problem of cabinet stability." *Comparative Political Studies* 17(2): 167–197.
- Browne, Eric C, John P Frenreis, and Dennis W Gleiber. 1986. "The process of cabinet dissolution: An exponential model of duration and stability in western democracies." *American Journal of Political Science* 30(3): 628–650.
- Cheibub, Jose Antonio, Adam Przeworski, and Sebastian M Saiegh. 2004. "Government coalitions and legislative success under presidentialism and parliamentarism." *British Journal of Political Science* 34(04): 565–587.
- Cox, David R. 1972. "Regression models and life-tables." *Journal of the Royal Statistical Society. Series B (Methodological)* 34: 187–220.
- Cukierman, Alex, Sebastian Edwards, and Guido Tabellini. 1989. *Seigniorage and political instability*. National Bureau of Economic Research.
- Dahlström, Carl, Johannes Lindvall, and Bo Rothstein. 2013. "Corruption, bureaucratic failure and social policy priorities." *Political Studies* 61(3): 523–542.
- Damgaard, Eric. 2008. "Cabinet Termination in Western Europe." In *Cabinets and Coalition Bargaining: the democratic life cycle in Western Europe*, eds. Kaare Strøm, Wolfgang C. Müller, and Torbjörn Bergman. Oxford: Oxford University Press.
- De Swaan, Abram, and Anatol Rapoport. 1973. *Coalition theories and cabinet formations: A study of formal theories of coalition formation applied to nine European parliaments after 1918*. Amsterdam: Elsevier.
- Dewan, Torun, and Keith Dowding. 2005. "The corrective effect of ministerial resignations on government popularity." *American Journal of Political Science* 49(1): 46–56.
- Diermeier, Daniel. 2006. "Coalition government." In *The Oxford Handbook of Political Economy*, eds. Barry R Weingast and Donald Wittman. Oxford: Oxford University Press, p. 162–179.
- Diermeier, Daniel, and Randolph T Stevenson. 2000. "Cabinet terminations and critical events." *American Political Science Review* 94(3): 627–640.
- Diermeier, Daniel, and R.T. Stevenson. 1999. "Cabinet survival and competing risks." *American Journal of Political Science* 43(3): 1051–1068.
- Dodd, Lawrence C. 1976. *Coalitions in parliamentary government*. Princeton: Princeton University Press.
- Döring, Holger, and Johan Hellström. 2013. "Who gets into government? Coalition formation in European democracies." *West European Politics* 36(4): 683–703.

- Downs, A. 1957. *An economic theory of democracy*. Harper & Row, London.
- Druckman, James N. 1996. "Party factionalism and cabinet durability." *Party Politics* 2(3): 397–407.
- Druckman, James N, Lanny W Martin, and Michael F Thies. 2005. "Influence without confidence: Upper chambers and government formation." *Legislative Studies Quarterly* 30(4): 529–548.
- Druckman, James N., and Michael F. Thies. 2002. "The Importance of Concurrence: The Impact of Bicameralism on Government Formation and Duration." *American Journal of Political Science* 46(4): 760–771.
- Eckstein, Harry. 1966. *Division and cohesion in democracy: a study of Norway*. Princeton: Princeton University Press.
- Eppner, Sebastian, and Steffen Ganghof. 2015. "Do (weak) upper houses matter for cabinet formation? A replication and correction." *Research & Politics* 2(1): 1–5.
- Finlay, B, and A Agresti. 2009. *Statistical methods for the social sciences*. Upper Saddle River, NJ: Pearson/Prentice Hall.
- Fisher, Stephen D, Robert Ford, Will Jennings, Mark Pickup, and Christopher Wlezien. 2011. "From polls to votes to seats: Forecasting the 2010 British general election." *Electoral Studies* 30(2): 250–257.
- Foucault, Martial, and Richard Nadeau. 2012. "Forecasting the 2012 French presidential election." *PS: Political Science & Politics* 45(2): 218–222.
- Fredén, Annika. 2014. "Threshold Insurance Voting in PR Systems: A Study of Voters' Strategic Behavior in the 2010 Swedish General Election." *Journal of Elections, Public Opinion & Parties* 24(4): 473–492.
- Frendreis, John P, Dennis W Gleiber, and Eric C Browne. 1986. "The study of cabinet dissolutions in parliamentary democracies." *Legislative Studies Quarterly* 11(4): 619–628.
- Gallagher, Michael, Michael Laver, and Peter Mair. 2011. *Representative government in modern Europe*. New York: McGraw-Hill.
- Geddes, B. 2003. *Paradigms and sand castles: Theory building and research design in comparative politics*. Michigan: University of Michigan Press.
- George, Alexander L, and Andrew Bennett. 2005. *Case studies and theory development in the social sciences*. Cambridge, MA: Mit Press.
- Grofman, Bernard, and Peter Van Roozendaal. 1997. "Modelling cabinet durability and termination." *British Journal of Political Science* 27(3): 419–451.
- Grotz, F., and T. Weber. 2012. "Party Systems and Government Stability in Central and Eastern Europe." *World Politics* 64(4): 699–740.
- Harrison, Jeff, and Mike West. 1997. *Bayesian Forecasting and Dynamic Models*. New York: Springer.
- Häusermann, Silja, Georg Picot, and Dominik Geering. 2013. "Review article: Rethinking party politics and the welfare state—recent advances in the literature." *British Journal of Political Science* 43(1): 221–240.
- Hindmoor, A. 2006. *Rational choice*. New York: Palgrave Macmillan.

- Hlavac, Marek. 2015. "ExtremeBounds: Extreme Bounds Analysis in R." *R package version 0.1.5.1*. <http://CRAN.R-project.org/package=ExtremeBounds>.
- Hurwitz, Leon. 1972. "Democratic Political Stability: Some Traditional Hypotheses Reexamined." *Comparative Political Studies* 4(4): 476.
- Jackman, Simon. 2009. *846 Bayesian analysis for the social sciences*. Chippenham, Wiltshire: John Wiley & Sons.
- Jackman, Simon. 2005. "Pooling the polls over an election campaign." *Australian Journal of Political Science* 40(4): 499–517.
- Jennings, Will, and Christopher Wlezien. 2014. "The Timeline of Elections: A Comparative Perspective." In *APSA 2014 Annual Meeting Paper*.
- Kayser, Mark Andreas. 2005. "Who surfs, who manipulates? The determinants of opportunistic election timing and electorally motivated economic intervention." *American Political Science Review* 99(1): 17–27.
- King, Gary, James E. Alt, Nancy Elizabeth Burns, and Michael Laver. 1990. "A Unified Model of Cabinet Dissolution in Parliamentary Democracies." *American Journal of Political Science* 34(3): 846–871.
- King, Gary, Robert O Keohane, and Sidney Verba. 1994. *Designing social inquiry: Scientific inference in qualitative research*. Princeton: Princeton University Press.
- Kitschelt, Herbert. 1995. "Formation of Party Cleavages in Post-Communist Democracies Theoretical Propositions." *Party Politics* 1(4): 447–472.
- Kleinbaum, DG, and M Klein. 2005. *Survival analysis: A self-learning approach*. New York: Springer.
- Knotz, Carlo, and Johannes Lindvall. 2015. "Coalitions and Compensation The Case of Unemployment Benefit Duration." *Comparative Political Studies* 48(5): 586–615.
- Korpi, Tomas. 2001. "Good friends in bad times? Social networks and job search among the unemployed in Sweden." *Acta sociologica* 44(2): 157–170.
- Laver, Michael. 1986. "Between theoretical elegance and political reality: Deductive models and cabinet coalitions in Europe." In *Coalition behaviour in theory and practice*, ed. Geoffrey Pridham. Oxford: Oxford University Press.
- Laver, Michael. 2003. "Government termination." *Annual Review of Political Science* 6(1): 23–40.
- Laver, Michael, and N. Schofield. 1990. *Multiparty governments*. Oxford: Oxford University Press.
- Laver, Michael, and Kenneth A Shepsle. 1998. "Events, equilibria, and government survival." *American Journal of Political Science* 42(1): 28–54.
- Leamer, Edward E. 1985. "Sensitivity analyses would help." *American Economic Review* 75(3): 308–13.
- Lijphart, Abend. 1968. "Typologies of democratic systems." *Comparative political studies* 1(1): 3–44.

- Linzer, Drew A. 2013. "Dynamic Bayesian Forecasting of Presidential Elections in the States." *Journal of the American Statistical Association* 108(501): 124–134.
- Lupia, Arthur, and Kaare Strøm. 1995. "Coalition termination and the strategic timing of parliamentary elections." *American Political Science Review* 89(3): 648–665.
- Manin, Bernard, Adam Przeworski, and Susan Stokes. 1999. *Elections and representation*. Cambridge: Cambridge University Press.
- Martin, Lanny W, and Georg Vanberg. 2011. *Parliaments and coalitions: The role of legislative institutions in multiparty governance*. Oxford: Oxford University Press.
- Martin, Lanny W, and Georg Vanberg. 2014. "Parties and policymaking in multiparty governments: The legislative median, ministerial autonomy, and the coalition compromise." *American Journal of Political Science* 58(4): 979–996.
- Müller, Wolfgang C., and Kaare Strøm. 2003. *Coalition governments in Western Europe*. Oxford: Oxford University Press.
- Müller, Wolfgang C., and Kaare Strøm. 1999. *Policy, office, or votes?: how political parties in Western Europe make hard decisions*. Cambridge: Cambridge University Press.
- Neumann, J von, and Oskar Morgenstern. 1944. *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Pickup, Mark, and Richard Johnston. 2007. "Campaign trial heats as electoral information: Evidence from the 2004 and 2006 Canadian federal elections." *Electoral Studies* 26(2): 460–476.
- Rae, Douglas W. 1971. *The political consequences of electoral laws*. New Haven, Co: Yale University Press.
- Randall, Vicky, and Lars Svåsand. 2002. "Party institutionalization in new democracies." *Party politics* 8(1): 5–29.
- Riera, Pedro. 2015. "Economy, Type of Government, and Strategic Timing of Elections: Calling Opportunistic Early Elections in OECD Democracies." *West European Politics* 38(6): 1129–1151.
- Riker, W.H. 1962. *The theory of political coalitions*. New Haven, Co: Yale University Press.
- Roberts, Andrew. 2008. "Hyperaccountability: economic voting in Central and Eastern Europe." *Electoral Studies* 27(3): 533–546.
- Robertson, John D. 1983. "The political economy and the durability of European coalition cabinets: New variations on a game-theoretic perspective." *The Journal of Politics* 45(04): 932–957.
- Roe, Mark J, and Jordan I Siegel. 2011. "Political instability: Effects on financial development, roots in the severity of economic inequality." *Journal of Comparative Economics* 39(3): 279–309.
- Rothstein, Bo. 2011. *The quality of government: Corruption, social trust, and inequality in international perspective*. Chicago: University of Chicago Press.

- Rothstein, Bo, and Jan Teorell. 2008. "What is quality of government? A theory of impartial government institutions." *Governance* 21(2): 165–190.
- Saalfeld, Thomas. 2013. "Economic Performance, Political Institutions and Cabinet Durability in 28 European Parliamentary Democracies, 1945–2011." In *Party Governance and Party Democracy.*, eds. Wolfgang C. Müller and Hanne-Marthe Narud. New York: Springer.
- Saalfeld, Thomas. 2008. "Institutions, Chance and Choices: The Dynamics of Cabinet Survival in the Parliamentary Democracies of Western Europe (1945–1999)." In *Cabinets and Coalition Bargaining: the democratic life cycle in Western Europe*, eds. Kaare Strøm, Wolfgang C. Müller, and Torbjörn Bergman. Oxford: Oxford University Press.
- Sala-i-Martin, Xavier X. 1997. "I just ran two million regressions." *The American Economic Review* 87(2): 178–183.
- Sanders, David, and Valentine Herman. 1977. "The stability and survival of governments in Western democracies." *Acta Politica* 12(3): 346–377.
- Savage, Lee. 2013. "Party System Polarisation and Government Duration in Central and Eastern Europe." *West European Politics* 36(5): 1029–1051.
- Schleiter, Petra, and Edward Morgan-Jones. 2009. "Constitutional power and competing risks: monarchs, presidents, prime ministers, and the termination of East and West European cabinets." *American Political Science Review* 103(03): 496–512.
- Schleiter, Petra, and Margit Tavits. 2016. "The Electoral Benefits of Opportunistic Election Timing." *The Journal of Politics* 78(3): 836–850.
- Seki, Katsunori, and Laron K Williams. 2014. "Updating the Party Government data set." *Electoral Studies* 34: 270–279.
- Shepsle, Kenneth A. 2010. *Analyzing politics: rationality, behaviour and institutions*. New York: W. W. Norton & Company.
- Silver, Nate. 2012. *The signal and the noise: Why so many predictions fail-but some don't*. New York: Penguin.
- Simon, Herbert A. 1991. "Bounded Rationality and Organizational Learning." *Organization Science* 2(1): 125–134.
- Sjöblom, Gunnar. 1968. *Party strategies in a multiparty system*. Lund: Studentlitteratur.
- Somer-Topcu, Zeynep, and Laron K Williams. 2008. "Survival of the fittest? Cabinet duration in postcommunist Europe." *Comparative Politics* 40(3): 313–329.
- Strøm, Kaare, Eric C Browne, John P Frensdreis, and Dennis W Gliieber. 1988. "Contending models of cabinet stability." *The American Political Science Review* 82(3): 923–941.
- Strøm, Kaare. 1985. "Party goals and government performance in parliamentary democracies." *The American Political Science Review* 79(3): 738–754.

- Strøm, Kaare, Wolfgang C. Müller, and Torbjörn Bergman. 2008. *Cabinets and coalition bargaining: the democratic life cycle in Western Europe*. Oxford: Oxford University Press.
- Strøm, Kaare, and Stephen M Swindle. 2002. "Strategic parliamentary dissolution." *American Political Science Review* 96(3): 575–591.
- Taylor, Michael, and Valentine M Herman. 1971. "Party systems and government stability." *The American Political Science Review* 65(1): 28–37.
- Teorell, Jan. 2010. *Determinants of democratization: Explaining regime change in the world, 1972–2006*. Cambridge: Cambridge University Press.
- Trussell, James, and David E Bloom. 1983. "Estimating the co-variables of age at marriage and first birth." *Population Studies* 37(3): 403–416.
- Tversky, Amos, and D Kahneman. 1981. "The framing of decisions." *Science* 211(4481): 453–458.
- Tveterås, Ragnar, and Geir Egil Eide. 2000. "Survival of New Plants in Different Industry Environments in Norwegian Manufacturing: AS Semi-Proportional Cox Model Approach." *Small Business Economics* 14(1): 65–82.
- Tzelgov, Eitan. 2011. "Communist successor parties and government survival in Central Eastern Europe." *European Journal of Political Research* 50(4): 530–558.
- Van der Eijk, Cees, and Mark Franklin. 2009. *Elections and voters*. Basingstoke: Palgrave Macmillan.
- Warwick, Paul. 1994. *Government survival in parliamentary democracies*. Cambridge: Cambridge University Press.
- Warwick, Paul. 1979. "The durability of coalition governments in parliamentary democracies." *Comparative Political Studies* 11(4): 465–498.
- Wlezien, Christopher. 1995. "The public as thermostat: Dynamics of preferences for spending." *American Journal of Political Science* 39(4): 981–1000.

Department of Political Science

Umeå University, SE-901 87 Umeå, Sweden

www.pol.umu.se



UMEÅ UNIVERSITY

ISBN: 978-91-7601-699-2

ISSN: 0349-0831

Research Report 2017:2