

# **How Governments Strategically Time Welfare State Reforms: Empirical evidence from welfare state legislation in five European countries**

## **Abstract**

Building on studies on the political business cycle, the literature on welfare state retrenchment has argued that governments which cut the welfare state try to avoid blame by implementing painful measures in the beginning of the mandate and expanding benefits as elections approach. In contrast to this linear relationship, we argue that governments often feel pressured to fulfill (mostly expansionary) campaign promises during the first months in office. Consequently, cutting right away is not what should be expected. Instead, a more nuanced, u-shaped timing trajectory is probable with a period in the beginning characterized by both cuts and fulfillment of expansionary pledges, followed by a period of cutbacks, and finally an expansive phase toward the end of a mandate. We test this argument on a new original dataset of legislative changes in five European countries - Britain, Denmark, Finland, France and Germany - during the last four decades.

## **Keywords**

Electoral pledges – electioneering – political business cycle – welfare state

Version Accepted at *West European Politics*, September 2019

<https://doi.org/10.1080/01402382.2019.1668245>

## 1 Introduction

Following Paul Pierson's landmark study on the British and American welfare state reforms in the 1980s (Pierson 1994), research on welfare state retrenchment has argued that governments will resort to blame avoidance strategies in order to avoid electoral punishment if they cut back on social policies (for a recent overview, see Vis 2016). Since then, a plethora of studies has found a wide variety of blame avoidance strategies both in relation to how retrenchment is organized and how it is communicated (e.g. Hering 2008; Hinterleitner and Sager 2015; König 2016; Wenzelburger 2011). Following the notion of political business cycles (Nordhaus 1975; cf. de Haan and Klomp 2013 for a review), one of the most prominent organizational blame avoidance strategies identified by this literature is related to the strategic timing of welfare state cutbacks (Fernández 2012; Hübscher 2016; König and Wenzelburger 2017; Tepe and Vanhuyse 2010; Zohlnhöfer 2007). The underlying logic of this strategy is straightforward: Because the welfare state is very popular among the electorate (Brooks and Manza 2007), governments engaging in a cutting exercise have to do so right in the beginning of their mandate and not towards the end if they want to keep office. The reason is that voters are expected to have a recency bias and decaying memory, which is exploited by re-election motivated incumbents as they implement popular policies and fiscal stimulus at the end of the term and unpopular ones in the beginning (Healy and Lenz 2014; MacKuen et al. 1992; Shi and Svensson 2006).

However, while the empirical evidence for strategic timing of welfare state cutbacks has been found in both qualitative (König and Wenzelburger 2017; Wenzelburger 2011; Zohlnhöfer 2007) and quantitative studies (Fernandez 2012; Hübscher 2016; Hübscher and Sattler 2017)<sup>1</sup>, the current literature suffers from two weaknesses. First, the theoretical argument

---

<sup>1</sup> The only study with more mitigated findings is the paper by Tepe and Vanhuyse, who find different effect sizes depending on their model specification and the operationalization of the timing variable (Tepe and Vanhuyse 2010: 1228-1229). This is partly due to the fact that the authors cannot pinpoint whether a cutback

neglects the fact that governments do not start their mandate from a scratch, but that they have promised certain policies to their voters. Hence, they are not free to time legislation, but have to weigh these considerations against the need to fulfill electoral pledges. Second, the measurement of the dependent variable is not entirely convincing, as most studies rely on spending data (which is known to be far from a good proxy for welfare state change, e.g., Clasen and Siegel 2007, Kühner 2007) or on legislative cutbacks only (Hübscher 2016, Hallerberg and Scartascini 2017, Strobl et al. 2019). thereby overlooking possible expansions.

On the first point, the finding that governments use the first months in office to curb the welfare state runs against the literature on electoral pledges which have found 1) that pledges are mostly fulfilled (Thomson et al. 2017b), 2) that pledge fulfillment is most probable in the first months in office (Duval and Pétry 2019) and 3) that pledges are mostly expansionary in nature – at least in the realm of the welfare state (Horn and Jensen 2017, Thomson and Costello 2016). These findings are also in line with mandate theory according to which parties to a large extent do fulfill what they have promised in the election campaign (although to varying degrees) (Budge et al. 2001; Klingemann 2006; Klingemann et al. 1994). If we take these three findings seriously, governments should often face a trade-off in the first year of their new mandate because they have to choose between fulfilling expansionary welfare pledges or strategically cutting the welfare state. The reason simply is that voters' memory of pledges from the election campaign is still fresh in the first months of a government's mandate and fulfilling these promises will therefore lead to increased credibility for the government and improves the image of the parties – something which is less likely four years down the electoral term.

---

occurred before or after the election date in an election year (which always takes the value of 0, see footnote 11 in their paper). Our approach using monthly data does not suffer from this problem.

Bringing together the literature on welfare pledges and strategic timing, the present article investigates the resulting timing pattern empirically. Rather than testing (once more) the simple linear model which expects governments to cut much in the beginning and less as time goes by, our approach is open to the possibility that governments are pressured to do both things at the same time. In this scenario, governments still adopt cutbacks more in the beginning than in the end of the election period, but only after having introduced legislation in line with their party manifesto right in the first months of their mandate. Hence, for the case of expansionary pledges, our theoretical expectation leads to a mixed picture in the very beginning of a mandate, followed by a cutting period and a more expansionary stance toward the end of a government's term, i.e. a U-shaped relationship (H1). If a government instead pledges to shrink the welfare state, we expect the well-known linear pattern (H2).

On the second point, this paper uses a new and more appropriate measure of welfare state change. We test our argument on a new data-set which is based on our original coding of a total of 1582 legislative changes in pension and unemployment policies between the early 1970s and 2014 in five industrialized countries – Denmark, Finland, France, Germany and the United Kingdom. This “Welfare State Reform Dataset” (WSRD) enables us to overcome two weaknesses of existing studies. First, we can account for instances when expansionary measures and cutbacks are implemented at the same time, whereas most existing studies on strategic timing look at cutbacks or fiscal consolidation only.<sup>2</sup> Accounting for changes in both directions is important because we know from case studies that welfare state reforms are often characterized by a mix of different measures that involve cutbacks *and* expansions (partly in order to “throw good money after bad” (Weaver 1986: 385)). Second, the new dataset on legislation enables us to look beyond aggregate spending data as a proxy for

---

<sup>2</sup> The works of Armingeon and Giger (2008), Giger and Nelson (2011) and Schumacher et al. (2013), for instance, code cutbacks as “1” and status quo and expansions as “0”. When both status quo and expansion are coded together as “0”, years (or election cycles) without any reforms are treated as identical to years (or election cycles) with expansions, making it impossible to assess the independent effect of expansions.

welfare state or fiscal policy changes – which is the measure usually used in political business cycle (PBC) studies. This is an advantage in respect to existing work because it is conceivable that legislation is much easier to be “timed” by governments than spending.<sup>3</sup>

Evidence from the analysis of our original dataset on the welfare legislations in the five countries supports our theoretical claims. Our findings suggest that, indeed, governments that promise to expand the welfare state do not start to cut the welfare state right in the beginning of their mandate – a behavior that results in the expected U-shaped trajectory over time. In contrast, the linear relationship can be found in cases of negative welfare pledges.

In the following section, we review the existing literature on the timing of welfare state cutbacks, and present our theoretical argument and testable hypotheses. After having discussed data and methods in section 3, the results of the empirical analysis are presented in section 4. A final section summarizes the results and concludes.

## **2 Theory**

### ***2.1 Timing welfare state change: The basic logic of “strategic timing”***

The idea that governments time welfare state changes strategically in order to maximize votes at the next election has been developed as a response to Paul Pierson’s claim that welfare state cutbacks carry electoral risks and are therefore unlikely unless governments use blame avoidance strategies to prevent electoral punishment (Pierson 1996: 156). In the years after Pierson’s landmark study, welfare state scholars have found a huge variety of such blame avoidance strategies (König and Wenzelburger 2014; Vis 2016). These strategies can be

---

<sup>3</sup> Admittedly, legislation also has to pass the legislative process, which may take some time; but politicians who care about their re-election might know these delays and will therefore already include this in their calculation of when the proposal has to be adopted. In any case, government spending is much more difficult to time because it is not only affected by the delays of legislative processes, but also by many other factors, such as demand for benefits (e.g. because of rising unemployment or population ageing), economic growth, or the possibility to actually spend the money that is allocated in a budget.

classified in two types: Strategic organization, which means that governments organize cuts in a way that does not endanger re-election, and strategic communication which relates to how cuts are framed in the governmental rhetoric (Wenzelburger 2011). Strategic timing clearly falls into the category of strategic organization and follows the basic logic that governments will try to generate a political business cycle, i.e. implement the painful cutbacks of benefits as early as possible in their mandate and expand the welfare state toward the end of the mandate.

The reason for why governments frontload cutbacks and become more generous the closer it gets to the next election is the voters' 'recency bias' when it comes to rewarding or punishing incumbents for their past policy record (Nordhaus 1975). This recency bias comes in several versions, all of which underpin the incentive for governments to time their policies strategically. MacKuen et al. (1992) find, for instance, that the most recent economic events have the greatest impact on evaluating the incumbent, because recent events serve as a proxy for the incumbent's future actions and the electorate's welfare. In a similar fashion, Healy and Lenz (2014) argue that voters have cognitive and time constraints and therefore rely on information that is easier to recall and gather. Moreover, they also discount past policy records. Voters thus substitute the 'the end for the whole' (Healy and Lenz (2014: 33) and rely disproportionately on the last year in office (respectively take the last year as heuristic) when evaluating incumbents' policy records. This might lead to biased decisions and distorted electoral accountability. Accordingly, Shi and Svensson (2006) argue that politicians may behave opportunistically and exploit that some voters remain uninformed and rely on heuristics. They further argue that the greater the number of voters who fail to identify election-motivated fiscal policy manipulations, the higher the incentive for incumbents to increase expenditures before an election. Based on this literature, we therefore expect that re-election seeking governments should exploit the recency bias and time

cutbacks in the early phase of their mandate, while they become more generous towards the end of the term where voters really approve what has been done for them.

Applying these arguments to budget consolidation, Zohlnhöfer's analysis of fiscal adjustment illustrates how different governments in the UK, the Netherlands and Germany have used this tool (Zohlnhöfer 2007). Several other case studies have also delivered empirical evidence of such patterns (König and Wenzelburger 2017; Wenzelburger 2011). On the quantitative side, too, studies by Hübscher (Hübscher 2016; Hübscher and Sattler 2017), Tepe and Vanhuysse (2010) and Fernández (2012) point to similar timing patterns. In terms of data, most of these studies are based on regression models and either public spending or measures of austerity programs as collected by the IMF (Devries et al. 2011), whereas expansions as well as the balance between both expansions and cutbacks is mostly ignored.

At the same time, however, the literature on blame avoidance strategies has come under some attack recently for not testing its basic precondition that welfare state cutbacks are indeed electorally risky. Several studies have shown that welfare state retrenchment is not always suicidal for governments and that the degree of punishment varies largely upon, for instance, on the overall situation of the economy (Giger and Nelson 2013) or the partisan composition of the government (Schumacher et al. 2013).

Should governments therefore abstain from strategic timing? In fact, evidence from the case study literature suggests that blame avoidance is still highly relevant, even though cutting the welfare state may actually be less risky than previously thought. From elite interviews, it is obvious that the perceived risk of being punished for welfare state retrenchment is widespread among politicians (Klitgaard and Elmelund-Præstekær 2014; Wenzelburger 2014). Similarly, research on other areas of unpopular policies, such as foreign military interventions involving casualties, point to a massive misperception of public opinion by the

political elite (Gelpi et al. 2009). Therefore, it is highly probable that the question of how to strategically time welfare state cutbacks is still an issue when a government plans to retrench social policies. What follows from these two bodies of literature – the studies on the timing of welfare cutbacks and the analyses of political business cycles (PBC) pointing to recency bias – is therefore a very clear expectation for the strategic timing of cutbacks and expansions to the welfare states that assumes a linear trajectory: *Governments will adopt cutbacks early in the electoral term and expansions toward the end of their mandate (strategic timing expectation).*

## **2.2 Complicating the issue: Fulfilling electoral pledges on the welfare state**

However, the political game is more complicated than the basic logic of early cuts and late expansions suggests. Several points complicate the issue. First of all, policy changes only seldom consist of cutbacks or expansions *only*. Given the highly sophisticated nature of modern welfare states, it is probable that reforms change many instruments at the same time resulting in expansions and reductions of generosity at the same time. In fact, studies have even shown that complementing overall cutbacks with expansion for some groups is even a common pattern used by governments when retrenching social policies (Bonoli 2012; Häusermann 2010; Jensen et al. 2017) (“division” and “compensation” according to Pierson (1994: 19)). It seems therefore more reasonable to use the net change of the status quo, instead of using only one side of the picture (namely cutbacks *or* expansions only), to assess whether strategic timing occurs.<sup>4</sup>

Second, and more importantly on a theoretical level, a new government mandate only rarely starts from nowhere. Instead, it has to deal with the pledges from the election manifesto.

---

<sup>4</sup> By “net change”, we refer to the overall direction of changes in welfare legislations made by a government in a given period. For example, if there were more expansionary changes than cutbacks, the net change would be expansionary and this overall net position, rather than the sheer extent cutbacks *or* expansions, would inform the voters the general image of the government (on the matter of welfare policies).



Indeed, mandate theory has repeatedly shown that, on average, governments do what they announce to do in their election manifestos (Budge and Hofferbert 1990; Hofferbert and Budge 1992; Klingemann et al. 1994; Klingemann et al. 2006). A recent comparative study based on an extensive coding of electoral pledges and their fulfillment (fully, or partially) also clearly corroborates this expectation (Thomson et al. 2017b). In the case of welfare pledges, it has been shown that parties mostly keep their welfare promises (Horn and Jensen 2017).

Moreover, and thirdly, we know from the welfare state literature and studies on electoral pledges, that most governments tend to promise expansionary politics. Thomson and Costello (2016) show, based on a handcoded material from Ireland, that expansionary pledges are most common. Similarly, the pledge data collected by Thomson et al. (2017a) in 12 countries also supports this result: The authors coded 3,373 pledges to increase spending or policy programs and a bare 364 instances where a pledge referred to a cut in spending or programs. Hence, we should expect parties to pledge expansions much more than cutbacks.

Finally, fourthly, we know not only *that* pledges matter, but also *when* pledges matter most to governments, namely in the very first months of the mandate – commonly discussed as the “honeymoon period” or the “first hundred days” (Dominguez 2005). This seems to be the period in which newly elected governments try to adopt the promised policy changes in order to signal to the voters that they live up to the campaign pledges (Huber 1996). Several case studies have pointed out the importance of the first 100 days (e.g. on Obama in the US: Coglianese 2009; on Mitterrand in France: Smith 1990). And a recent study on Canada, based on manually coded data about fulfilled pledges, has also found that “if the government does not enact pledges within the first half of its mandate, the probability of these pledges ever being fulfilled drops drastically” (Duval and Pétry 2019).

Why should this be the case? Arguably, incoming governments do their best to keep their pledges because not doing so can rub off on the public's view of the government's trustworthiness and likeability (Schedler 1998). This logic builds on work by Lodge et al. (1989), McGraw et al. (1990), and Lodge et al. (1995), showing how voters' evaluating of candidates and parties is affected by an impression-based "running tally". Voters can often not recall all relevant information from memory (especially if it dates back in time), but instead draw on their running tally when deciding if they like a politician (or a government) and want to vote for her. Based on this, our assumption is that breaking a pledge creates negative feelings towards an incumbent, while simply implementing the same policy without first having to break a pledge does not. Making an explicit promise to expand the welfare state and then retrenching it, is worse than just retrenching it. Doing so implies that (welfare-supporting) voters not only consider the incumbent misaligned with their policy position, but also as untrustworthy – a characteristic that can be very damaging for future vote intention. As studies have shown, trust can be seen as a heuristic (Rudolph 2017, Hetherington 2005) that increases the probability of citizens to vote for the incumbent (Hooghe 2018).

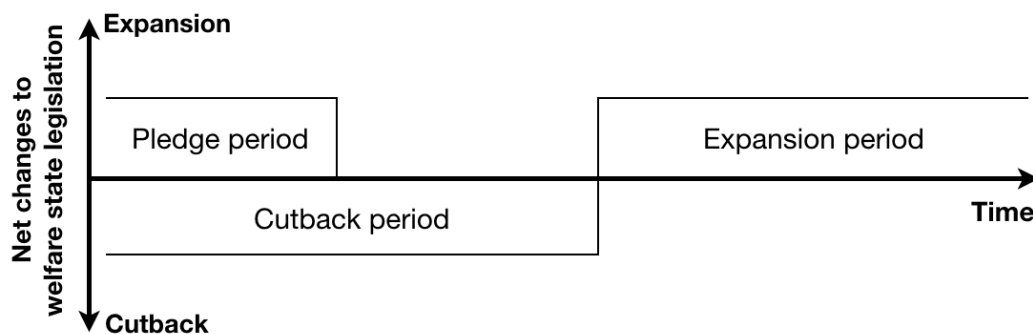
Hence, although most voters may not recall whether a specific pledge was implemented at election day, breaking pledges at the early stage of an election cycle will have negative downstream consequences because it alters voters' trustworthiness of the government. What is more, the negative impact on trust should be highest 1) when voters can still remember campaign pledges well and 2) when they expect that the government has most leeway to adopt pledges. In other words, when voters have a fresh memory about what government party has promised a couple of months before, fulfilling pledges will boost the credibility of the party and breaking them will lead to a drop in trustworthiness. This is much less the case toward the end of a government's mandate when voters may not recall what has actually been said five years before. Evidence on beneficial policies, that may resemble fulfilled pledges,

show that such positive evaluations can last longer than the voting literature usually acknowledges (Bechtel and Hainmueller 2011). These two considerations explain why the first 100 days are often used in the media to establish to what extent pledges had been fulfilled and also explain the finding by Duval and Pétry (2019) why the chance to fulfill pledges is highest during the first years of a government's mandate. From this follows a clear expectation on the timing of pledge fulfillment: *Governments fulfill welfare pledges during the first months of a mandate (welfare pledge expectation)*.

### **2.3 The pattern of welfare state legislation change over the time of a mandate: A U-shaped relationship**

If our theoretical expectations above are true, two different and partly overlaying dynamics should be visible during a government's mandate: A "pledge effect" on the one hand, which must be visible mainly in the first months; and a "strategic timing" effect on the other, which results in more cutbacks early on and expansions later. From these two propositions, it is evident that the resulting pattern of welfare state legislation over a government's mandate will depend on the initial welfare pledge before the election. If a government uses an expansionary welfare pledge, which is by large the most common situation (Horn and Jensen 2017: 386), a U-shaped trajectory of welfare legislation changes should result: a first phase right in the beginning when the expansionary pledges are fulfilled but cutbacks may already been implemented to take advantage of the recency bias of the voters, which is followed by the classic, but delayed, "strategic timing" trajectory of continuing cutbacks (but without the counterweighing effect of pledges) and expansions toward the end. Figure 1 illustrates this theoretical expectation. The classic expectation of strategic timing is embedded with the cutback period in earlier stage and expansion period in later stage, but the early cutback period is overlaid by a pledge period in the beginning of an electoral term in which governments fulfill electoral pledges.

**Figure 1: Expectation of timing trajectory**



Note: Net change to welfare state legislation on the y-axis denotes the overall direction of the welfare legislation at a certain point in time. The larger value (the upper direction on the y-axis) indicates that the government enacted more expansionary legislations than cutbacks, and the lower value denotes that the government enacted more cutbacks than expansionary legislations. Please note that during the first months or so, the pledge period and the cutback period overlap, so that the net change may be positive (if expansionary pledges dominate), around zero (if cutbacks and expansions cancel each other out) or negative (if cutbacks dominate).

Clearly, the exact nature of the pattern is very hard to predict – in terms of the length of the periods and the amplitude of the net changes. We can neither exactly say how long the first period (in which campaign pledges dominate welfare policies) would last as the length may differ across countries depending on the historical contexts, institutional settings or the parties in power. Nor can we predict from the outset how strongly governments pledge welfare state expansions, whether the pledges are all fulfilled in the beginning, and how extensively there will be the cutbacks due to strategic consideration. Hence, during the first months, the overall net change can be slightly positive (if pledges are very expansionary) or slightly negative (if strategic cutbacks overweigh). Similarly, the pledge period can last for 100 days or even for one year. Nevertheless, what we can say is that the shape of the basic trajectory should follow a steeper or flatter U-shape. This leads us to our first hypothesis. We expect that *the net change of welfare state legislation change is U-shaped over the electoral term if the initial welfare pledge of the government is expansionary (H1).*

However, as we know from the welfare state literature, electoral pledges are mostly, but not always expansionary (see above: in Thomson et al.'s (2017a) data, the factor is 1/10 (cutback/expansionary pledges)). In fact, especially conservative parties (such as Thatcher in 1979 or the German government in 1983) did pledge cutbacks in their manifestos and implemented them. In their study of welfare pledges, Horn and Jensen (2017) report four instances in which a government was elected in office based on a clearly negative welfare pledge.<sup>5</sup> Consequently, a U-shaped trajectory of welfare state change over a government's mandate is therefore not to be expected in these cases of a negative initial welfare pledge. Instead, governments that announced to cut the welfare state in the election manifesto should do so right from the start – resulting in a continuous positive slope of the relationship between net change and elapsed time. Therefore, our second expectation is that *the net change of welfare state legislation will increase linearly over a governments mandate if the initial welfare pledge of the government is not expansionary (H2).*

### **3 Research design**

#### ***3.1 The scope and the data on the welfare legislation changes***

We test the expectations by analyzing legislative changes concerning the social rights of citizens in the two main social transfer programmes – old-age pensions and unemployment protection – in Britain, Denmark, Finland, France and Germany between the early 1970s and 2014.<sup>6</sup> Our basic coding units are changes to individual policy instruments (e.g. duration of benefits, level of benefits, see Online Appendix, p.8), which are nested in welfare reforms (usually legislative bills). While some of these reform events contain the change of only one instrument, bigger reforms usually encompass several instrument changes. As we coded

---

<sup>5</sup> These are the Danish governments starting in 1973 and 1982, the French election in 1986, the German election in 1983 and the British election in 1983. Hence, electoral pledges seem to reflect the common notion of parties' positions on social policies.

<sup>6</sup> Due to data availability, the time series start in 1970 for UK, 1971 for Denmark and 1974 for Germany, Finland and France.

whether an instrument involved a cutback or an expansion of the welfare state, we could calculate an aggregate monthly “net change” score (see below).

The research design involves three decisions on case selection: on welfare program, on countries, and on time. The two programs were selected for two main reasons. First, they constitute a core part of what is conventionally understood as ‘the welfare state’ and, as a result, are prominently featured in a very long list of publications on the politics of the welfare state (see for instance: Clasen 2005; Flora 1986; Huber and Stephens 2001; Pierson 1996). And second, both programs are transfer schemes and are guided by rules to transfer cash benefits or benefits in kind to recipients, which is why we can code them using the same coding scheme building on these rules.

The five countries were selected because they provide great variation in terms of their respective welfare state models (Esping-Andersen, 1990; Castles and Obinger, 2008) and the popularity of the welfare programs (Brooks and Manza 2007): Britain is a classic example of a residual welfare state, where benefits are mostly means-tested and relatively low. Denmark, in comparison, has substantially more generous benefits and less frequently employs a (strict) means test. France and Germany are archetypal insurance-based systems, where benefits are tied to earnings history and membership of occupational schemes. Finally, Finland is an interesting case as it unites features of the Scandinavian universal welfare state and the insurance-based Continental European model (see Kangas 2006). Moreover, in all five countries, support for the welfare state varies, with lower values for Britain and more support in Germany or Denmark (Brooks and Manza 2007; Svallfors 2012: 14). Hence, selecting five very different cases in terms of institutional setup and public support of the welfare state allows for a test whether our theoretical expectations on timing patterns hold in different contexts.

Finally, this study covers the period after oil crisis in 1970s. The year 1974 marks the first full year after the first oil price crisis. It is conventional to view 1973 as the end year of the Golden Age of welfare state expansion as well as the beginning of the new era of permanent austerity (Pierson, 1994; 1996). From the mid-1970s onwards, governments faced a higher pressure to introduce austerity measures, which we would suspect to be timed strategically by governments to lower electoral costs.

Our dependent variable is the net change of welfare state legislation measured as the difference between the number of positive (expansive) and the number of negative (retrenching) changes of individual policy instruments (see Online Appendix p.8), which then are aggregated to indicate the net change of the legislative status quo in every month of a year. We believe that this is the most appropriate indicator to measure the strategic timing of welfare legislation because focusing on cutbacks or expansions only would give a biased picture of the actual legislative activity: Very frequently, cutbacks and expansions occur at the same time as part of a bigger reform package and we expect political actors that try to fulfill pledges or take advantage of the recency bias of the voters to care most about the *overall* direction of cutbacks and expansions – something which the net change variable takes up. Hence focusing on the timing trajectory of expansions or cutbacks only (or separately) does not make sense from a substantive point of view.

The measure of net change is based on 1582 individual instrument changes, which are nested in 703 reforms. These make up the “Welfare State Reform Dataset” (WSRD). Every instrument change coded as “expansion” or “cutback” in the five countries has been collapsed into monthly observations. To construct the measure of net change, we first identified the 703 welfare reforms. We then identified, what individual instruments had been changed based on a fine-grained coding scheme of 13 policy instruments (see Online Appendix, p.8). The basic promise is to code individual reform events (i.e. instrument changes) as “cutbacks” when it

curtailed the social rights and “expansions” when the reform event improved citizens’ social rights.<sup>7</sup> For instance, in December 1981 the German government implemented series of social policy saving measures in the “Arbeitsförderungskonsolidierungsgesetz”. Among other things, the maximum duration period of unemployment benefits was cut from 12 to 10 months and the access to the benefits tightened (one had to contribute for one year within a period of 3 year to be eligible instead of 6 months). In this case, we coded the two changes as individual instrument changes since they concerned two different instruments (benefit level and contribution period), and both were coded as cutbacks. The net change variable is calculated by subtracting the number of cutback reforms from the number of expansionary reforms occurred in a given month. This yields a monthly measure of cutbacks, expansions, and net change.<sup>8</sup>

To illustrate the data structure, Figure 2 displays the coded expansionary and retrenching changes in the five countries. From the data, even from the bird’s eye view, several well-known country specificities are clearly discernible. First, the major cutbacks of the British welfare state in the 1980s (especially 1980 and 1986) and early 1990s (especially 1995) implemented by the Conservative governments are nicely illustrated by the graph, and so are

---

<sup>7</sup> The individual events were identified by a team of trained research assistants and a senior researcher on the basis of as many secondary sources as possible and supplementing these with searches in legislative databases when necessary (see Online Appendix, p.9-23, for a list of sources). The coding of the instruments and the direction (cutback, no change, expansion) were done by the assistants and subsequently controlled by a senior researcher. In the event the senior researcher did not agree with the original coding, the relevant research assistant and senior researcher discussed the coding decision in detail to reach agreement; however, there were very few such instances. We have double-checked our selection of reforms for 4 countries by contrasting it with studies based on a full search of legislative databases (due to the lack of data, we had to exclude Finland). To do so, we have randomly chosen one year from our sample – the year 1992 – and compared, whether the 9 reforms identified by us (out of 703) were similar to those identified by other researchers in this year. For 1992, we can show that our selection is very close to what other researchers starting from the full selection of legislative databases have found: 6 of the 9 reforms were completely matched, 2 reforms identified by us were part of budget bills (in Germany) and therefore not listed in the respective search category and only 1 reform (in Denmark) was not identified by us, but in the database of the Comparative Agenda Project (Baumgartner/Jones 1993) (for more details, see Online Appendix, p. 24-26).

<sup>8</sup> In the 2554 coded months, we find 254 months with a positive value of net changes (i.e., more expansion than cutbacks) and 163 months with a negative value of net changes (i.e., more cutbacks than expansions). The data include 2134 months with a value of zero, which indicates either the case where an equal number of cutbacks and expansions (a total of 26 months) or the case where there was no relevant legislative action (a total of 2108 months).

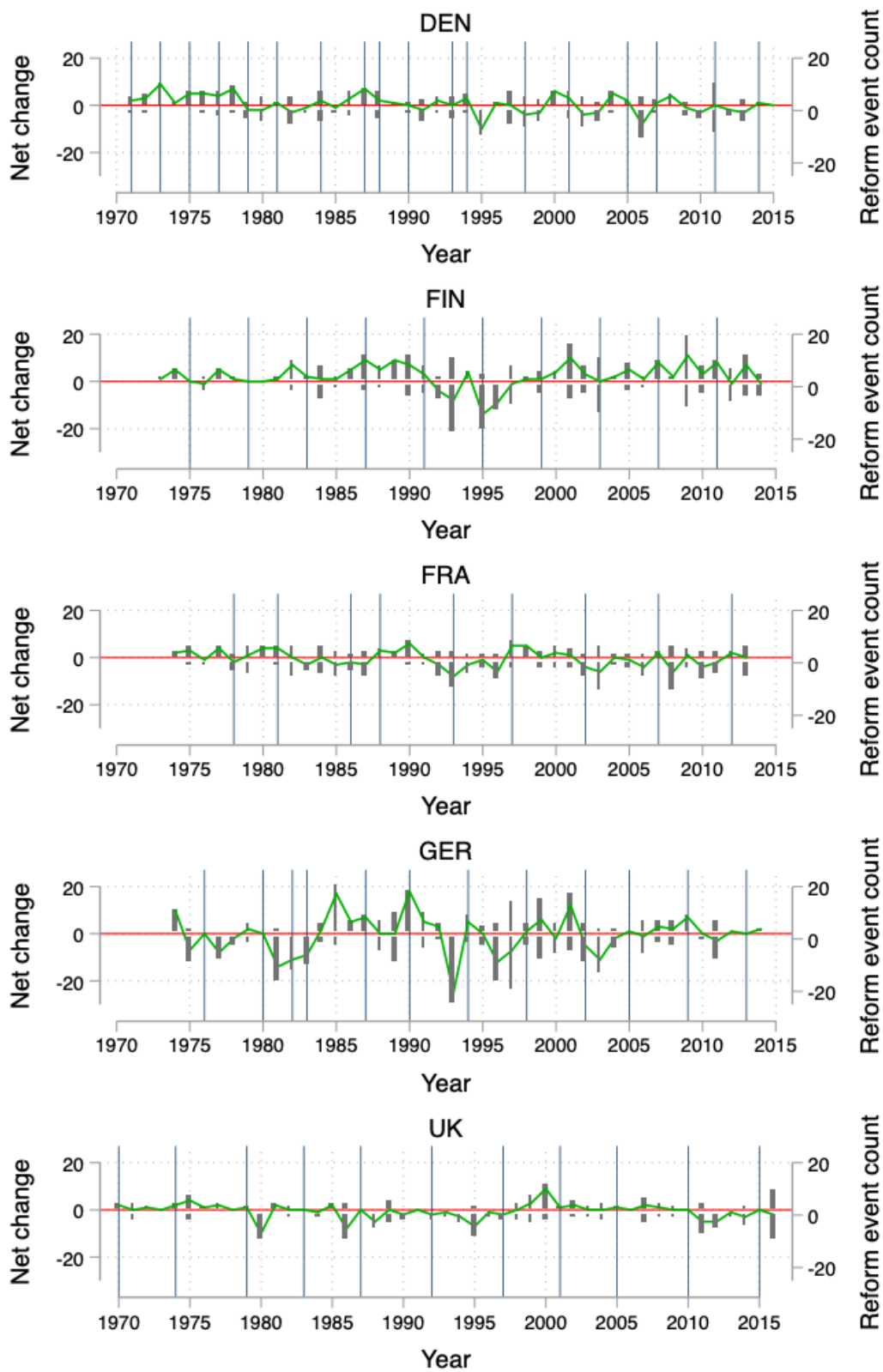


both the much more expansive stance followed by New Labour and the cutbacks in connection with the budget balancing efforts by the first Cameron administration in 2011. This pattern is in line with extant qualitative and case research, which often highlights the British experience as outstanding for the retrenchment it has witnessed (e.g., Allan and Scruggs, 2004; Clasen, 2005; Jensen, 2014) and shows the face validity of our data collection.

Likewise, the welfare cuts in Finland in the early 1990s are also clearly visible as well as the much more expansionist stance during the 1980s (cf. Timonen 2003). The German pattern shows both the cutbacks in the early 1980s under the new Kohl-government, the cuts in the early 1990s in the face of rising budgetary pressure and the retrenchment under Chancellor Schröder in the context of the Hartz-reforms. Finally, the very aggregate view on France and Denmark also fits the qualitative evidence, as we can for instance see several smaller retrenchment efforts in France in the beginning of the 1990s (governments Balladur and Juppé) as well as the Danish labor-market reforms in the mid-1990s.

Lastly, it is worth acknowledging that there are a large number of months without reforms. In fact our dependent variable suffers from high kurtosis, which is why we will use bootstrapping and estimate robust standard errors in our regression analyses.

Figure 2: Net changes in welfare legislations in five countries



Note: The x-axis denotes the year and the bars indicate the absolute number of expansions and cutbacks over time. The line plots the resulting net change. Vertical solid lines indicate years in which new governments began.

### 3.2 *Measurements and modeling strategies*

Our main explanatory variable is the timing over the course of the government's lifetime, measured by the percentage of time elapsed<sup>9</sup> in an electoral term.<sup>10</sup> As explained above, the combination of the classic political business cycle expectation with the 'pledge-keeping' expectation implies a U-shape relationship for expansionary pledges (H1) and a linear relationship for positive pledges (H2).

Welfare pledges are measured from party manifestos drawing on Horn and Jensen (2017). They use data from the Comparative Manifesto Project (Klingemann et al. 2006), and construct the pledge measure by subtracting the proportion of welfare-sceptic, anti-egalitarian and meritocratic positions in the party's manifesto from that of welfare state expansion, greater equality, state intervention, and market scepticism. In case of coalition governments, we weigh the respective values by the cabinet seat share of the individual parties in government (cabinet pledge). However, as legislation is often strongly influenced by the Ministry that develops and presents the bill (Alexiadou 2015), we account for this influence by calculating the average of the weighted cabinet pledge and the pledge as presented in the manifesto of the party holding the Ministry of Social affairs. The resulting measure varies from -33.6 to 51.7, with an average of around 11. Positive values denote that governments promised more expansionary welfare state. Hence, in our data, most of the times governments made a positive welfare pledge (81% of data points), which confirms the literature on the

---

<sup>9</sup> We have also tested the models using two other timing measures as independent variable: 1) time (months) until the next scheduled election and 2) time (months) since the last election. The substantive results do not change in both cases (see Online Appendix, p.6-7). We believe that the percentage of time elapsed since the last election is a better proxy of our theoretical argument as at least two of the countries (Denmark and the UK) often experience early elections.

<sup>10</sup> We always used the electoral term as basic unit of coding and did not code a change of Prime Ministers within an electoral term as a new start of the term, unless the change of Prime Minister did result in a major change of the political camp of the government that can be seen as a "new government". However, there were only two of such "changes of government" without election: Denmark in 1993 and Germany in 1982. Data source for elections and government changes is ParlGov (Döring and Manow 2018).

high probability that governments promise to expand the welfare state. The fact that parties are more inclined to talk about expansions than cutbacks in their manifestos is clearly in line with what we would expect given the high popularity of the welfare state.

To test our hypotheses, we split the sample on our pledge variable in two subsamples: one with cases in which pledges are expansionary, and one in which pledges are not expansionary (for descriptive statistics, see Online Appendix, p.3). Building on the coding in the comparative manifesto data (Volkens et al. 2017), this comes down to splitting the sample on the value of 0 – which means, substantively, that the share of positive and negative quasi-sentences related to welfare in a party program cancels each other out.<sup>11</sup> Because of the dominance of positive pledges, this approach yields two unequal subgroups with 2081 cases of positive pledges (mean 15.03) and 476 instances of negative pledges (mean -9.10). As a robustness check, we also use an alternative approach where we divide the sample in two groups on the median of our pledge variable. We also estimate a model on the full sample to see whether, overall, the U-shape or the linear relationship dominates.

To test the theorized relationships, we take two approaches. First, we run linear regression models for the positive and negative pledge cases to examine the linearity of the time trajectories in our hypotheses. In the OLS regression model  $Y_{\text{net changes}} = \beta_0 + \beta_1 * T + \beta_2 * X + \varepsilon$ , where T is the time elapsed since the last election and X denotes a set of covariates, we expect the  $\beta_1$  to be positive and significant for the sample of negative pledges (H2), whereas the linear relationship to be less clear with the sample of positive pledges (H1).

Second, we estimate a quadratic model to test the hypothesized U-shaped (H1) and the linear relationships (H2):  $Y_{\text{net changes}} = \beta_0 + \beta_1 * T + \beta_2 * T^2 + \beta_3 * X + \varepsilon$ , where T is the time elapsed since the last election and X denotes a set of covariates. To accurately test whether the U-

---

<sup>11</sup> Empirically, the pledge variable varies from -0.48 to 0.32, and there is no case in which the pledge variable takes the value of zero in the dataset.

shaped relationship exists, we use the test proposed by Lind and Mehlum (2010). This approach accounts for the composite nature of a U-shaped relationship and tests whether the relationship is decreasing for low values and increasing for high values of the independent variables. The test also indicates whether the extremum point falls within the analysed data range. We therefore report Lind and Mehlum's U-shapedness test statistics for quadratic models. In the cases of a negative initial pledge, if, as suggested in our theory, pledge-keeping effects in the first months are not discernible from the cutback period of the classic political business cycle and thus the relationship between elapsed time and net changes is more like a linear relationship (than a U-shaped), the test statistics will fail to reach the conventional significance. Conversely, for the cases of positive initial pledge, we expect to see the t-value of the test to be fairly high.

Besides our main variables of interest, we include several control variables that may affect both the timing pattern and the net change of welfare state legislation. To control for economic conditions, we include economic growth rates (quarterly, lagged) and unemployment rates (biannually, lagged). We also include Jahn's time-varying veto player variable (Jahn 2010) to control for veto player context because a high number of veto players may blur the clarity of responsibility for governmental action. However, if voters cannot clearly see who is to be held responsible for policies, governments have less need to time cutbacks and expansions strategically (Hobolt et al. 2013; Royed et al. 2000). Jahn's measure is based on Tsebelis (1995) conceptualization and uses policy distances to measure the distance between veto players to generate a time varying index. Lastly, country dummies are included to account for remaining unobserved cross-country heterogeneity.

We test our expectations by running an OLS regression on a pooled dataset (N=2563) which is based on monthly observations in all five countries as well as on the subsamples split according to the initial welfare pledge. Due to missing data on some covariates (e.g.

unemployment in the early 1970s in France, Jahn's veto-player index after 2012), most regressions are based on a slightly smaller sample. As discussed earlier, the distribution of our dependent variable is zero-inflated.<sup>12</sup> Therefore, estimations are bootstrapped and robust standard errors are used.<sup>13</sup> All regressions have been checked for autocorrelation given the time-seriesness of the data structure. Convenient tests indicate no violation of this assumption.<sup>14</sup> Heteroscedasticity is also accounted for via robust standard errors. In order to show the robustness of the results, we have run different models by using alternative measurements for our main explanatory timing variables (e.g. time until next scheduled elections) and including the continuous variable for initial welfare pledge as a covariate. All results for robustness checks are reported in the Online Appendix (p.6-7).

## **4 Empirical findings**

### ***4.1 Descriptive analysis: Timing trajectory of net changes in welfare legislations***

Pooling the data of all legislative changes to unemployment insurance and pensions in Britain, Denmark, Finland, France and Germany enables us to assess an overall pattern of strategic timing in these countries. Figure 3 presents how cutbacks and expansions as well as the resulting net change (our dependent variable) are distributed over time – measured in slices of 20% of the time elapsed in an electoral term. First, from the graphs in the left and center we can observe that cutbacks and expansions occur all the time and they counterbalance each other. Admittedly, the first and the last 20% of the mandate do see a bit

---

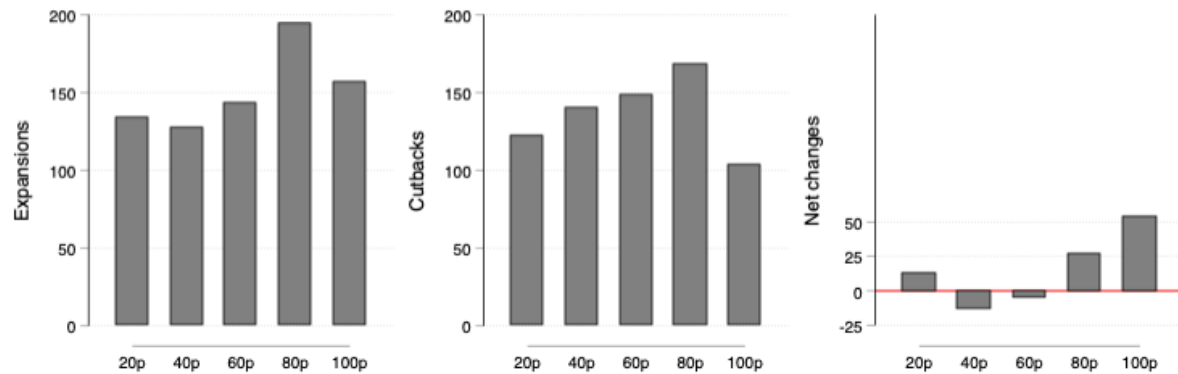
<sup>12</sup> The fact that the distribution of legislative events has a high peak around 0 is a classic finding of studies on legislative behaviour (Baumgartner and Jones 1993; Baumgartner et al. 2006).

<sup>13</sup> Hurdle models (or selection models) are not a useful option as we do not have strong theoretical assumptions about the underlying process that leads to legislative events as such. Using hurdle models would, however, imply that we have an argument about the first step, e.g. the probability of an event. Moreover, our dependent variable is not a pure count of events (i.e., the net change variable varies from negative to positive values as it is calculated by taking the difference between the two event counts) and all our theoretical arguments are linked to the overall direction of the legislative changes.

<sup>14</sup> A Wooldridge-test (xtserial-command in Stata) does not reject the null hypothesis of no serial correlation. The lagged residuals are not correlated with the residuals.

less events, but governments do adopt legislative changes to the welfare state already in the first months.

**Figure 3: Timing pattern of expansions, cutbacks, and net changes in five countries**



Note: The x-axis denotes the period of an electoral mandate, in percentage (e.g., if a government’s term was 5 years, each 20% points roughly indicate a time period of one year). The graphs on the left and center indicates the number of expansionary and cutback legislations, respectively, over the period of an electoral mandate from the pooled data. The y-axis of the right-most graph indicates net changes to welfare state legislation across the five countries, computed by subtracting the number of cutbacks from the number of expansions occurred in the given periods on the x-axis.

Second, more importantly, it is obvious from the right-most graph that the overall trend of net change follows the expected pattern described in our first hypothesis: The final 40 percent of the term duration can be featured as expansionary (and increasingly so), and the first 20 percent of the mandates are slightly expansionary followed by a period of cutbacks. These descriptive patterns also support our claim that the theoretical expectation about strategic timing (either based on the classic political business cycle or on our extension of this theory) from the perspective of voters and vote-seeking politicians has a much better fit with the net change – the consideration of expansions and cutbacks *simultaneously* – than the sole trend of expansions or cutbacks. This at the same time justifies our choice of dependent variable.

## 4.2 Results

Table 1 presents the results from the linear and quadratic regressions for the sample of governments with positive pledges (Models 1 to 3) and negative pledges (Models 4 to 6), and for the full sample (Models 7 and 8). For the three subsamples, we include results for 1) the linear regressions (Models 1, 4 and 7) and the quadratic regressions (Models 2, 5 and 8), as well as an alternative quadratic model where we split the samples on the median (Models 3 and 6) instead of the value of 0 (which can be, nevertheless, interpreted in a more meaningful way, see above). All models include the same set of covariates.

**Table 1: Results from linear and quadratic regression models**

	Positive pledge			Negative pledge			Pooled sample	
	Linear (1)	Quadratic (2)	Quadratic (median split) (3)	Linear (4)	Quadratic (5)	Quadratic (median split) (6)	Linear (7)	Quadratic (8)
<i>Timing variables</i>								
Percent of time elapsed	0.00089	-0.0070**	-0.0062*	0.0042***	-0.0032	-0.0069	0.0016*	-0.0063**
	(0.82)	(-2.00)	(-1.70)	(2.98)	(-0.51)	(-1.54)	(2.01)	(-2.54)
Percent of time elapsed <sup>2</sup>		0.000079**	0.000067*		0.000072	0.000100**		0.000079***
		(2.30)	(1.94)		(1.28)	(2.11)		(3.26)
<i>Time-varying or country-specific factors</i>								
Veto player index	-0.0031	-0.0034	-0.0022	-0.016	-0.018	-0.0088	-	-0.0040
	(-0.88)	(-1.16)	(-0.74)	(-0.76)	(-0.87)	(-1.15)	0.0037	(-1.24)
GDP growth (lagged, quarterly)	0.032*	0.032	0.019	0.036	0.035	0.043	0.031*	0.031*
	(1.66)	(1.44)	(0.99)	(0.89)	(1.09)	(1.33)	(1.73)	(1.83)
Unemp. rate (lagged, biannually)	-0.035***	-0.035***	-0.043**	0.0068	0.0074	-0.023	-	-0.032***
	(-2.77)	(-2.81)	(-2.44)	(0.27)	(0.29)	(-1.54)	0.032**	(-2.91)
Early election (dummy)	-0.14	-0.15	-0.18	-0.32*	-0.31	-0.18	-0.19	-0.20
	(-0.57)	(-0.68)	(-0.70)	(-1.84)	(-1.62)	(-1.56)	(-1.17)	(-1.01)
<i>Country fixed effects</i>								
Denmark	0.050	0.048	0.087	0.18	0.18	0.033	0.068	0.066
	(0.53)	(0.49)	(0.57)	(1.02)	(1.04)	(0.34)	(0.95)	(0.82)
Germany	-0.0037	-0.00041	-0.13	0.19	0.20	0.094	0.012	0.015
	(-0.04)	(-0.00)	(-0.78)	(0.46)	(0.44)	(0.69)	(0.13)	(0.15)
United Kingdom	-0.079	-0.081	-0.067	-0.26	-0.27	-0.22*	-0.13*	-0.13*
	(-0.90)	(-0.74)	(-0.65)	(-1.51)	(-1.43)	(-1.79)	(-1.65)	(-1.81)
Finland	0.22***	0.22**	0.17	--	--	0.48**	0.22**	0.22**
	(2.72)	(2.19)	(1.60)			(2.57)	(2.58)	(2.53)
France	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Constant	0.24*	0.37**	0.42**	-0.26	-0.13	0.25	0.18	0.31***
	(1.77)	(2.24)	(2.21)	(-1.17)	(-0.61)	(1.64)	(1.59)	(2.80)
<i>N</i>	1878	1878	1223	473	473	1128	2351	2351
<i>R</i> <sup>2</sup>	0.013	0.015	0.018	0.021	0.023	0.020	0.012	0.014



<i>R</i> <sup>2</sup> (adjusted)	0.0079	0.0097	0.0100	0.0041	0.0040	0.0110	0.0084	0.0100
<b><i>U-test results</i></b>								
<i>Utest (t-value)</i>		2.37***	1.46*		0.51	1.55*		2.00**
<i>Slope lower bound</i>		-0.0069694	-0.0062079		-0.0031992	-0.0068696		-0.0062613
<i>Slope Upper bound</i>		0.0088428	0.0071836		0.0112793	0.0131035		0.0094586

Note: In Models 4 and 5 Finland is dropped from the regression, because no government has had a negative pledge (a negative value on our pledge variable) during our period of observation; t-statistics in parentheses; \* p < .1, \*\* p < .05, \*\*\* p < .01.

For our main variables of interest, the models provide strong support of our two hypotheses.

First, Models 1 to 3 show that when the government's initial pledge was positive, the trajectory of welfare state change is U-shaped, rather than a linear trend, over the electoral term. Recall that these are the instances in which the government parties pledged to expand the welfare state before the election – and the great majority of our monthly observations belong to this case (1878 country-months in contrast to a bare 473 country-months for the negative pledge group). The results show that the coefficient for the timing variable in the linear model is not significant (Model 1) whereas the coefficients of the timing variables in the quadratic regression (Model 2) indicate a U-shaped relationship. More importantly, the u-test also yields a significant t-value with a negative slope for low values of the x-axis (i.e. first months of a mandate) and a positive slope for high values (i.e. the last months of a mandate). This pattern holds when we split the sample on the median instead of zero (Model 3), but is only marginally significant (at p<.10) with this reduced sample size for the positive pledges (since this classification generates much more balanced groups of around 1200 monthly observations).

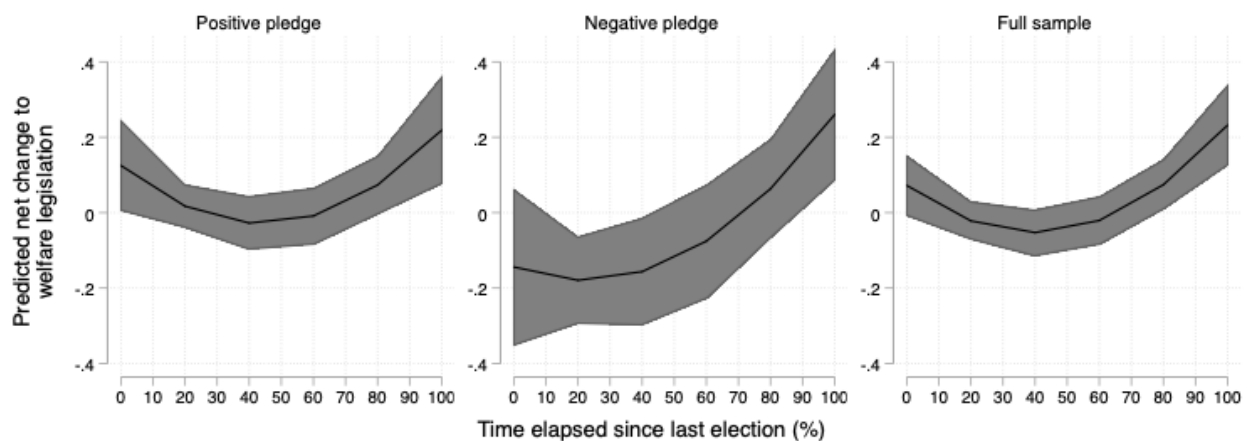
Second, for the case of negative pledges the regression results are quite different from the case of positive pledges: The coefficient of the timing variable is statistically significant in the linear model (Model 4); but the u-test for the quadratic model (Model 5) fails to indicate a U-shaped relationship within the boundary of the data. Overall, the linear and the quadratic

models demonstrate support for the U-shaped relationship for the case of positive pledges and the linear trajectory for the case of negative pledges.

As an additional test, Models 7 and 8 present the results from the sample of all cases – irrespective of the initial welfare pledge. The coefficients of the time variables are both significant at  $p < 0.05$  in the linear regression (Model 7) and the quadratic model (Model 8). However, the linear slope is much flatter than the one we saw from the negative pledge sample only (Model 4), and the quadric model results are very much similar to the quadratic model results from the positive pledge cases, demonstrating the U-shaped relationship attested by the u-test results and better model fit statistics for the quadratic model than the linear one. This is not surprising given the dominant proportion of positive pledge cases (ca. 80%) in the pooled sample.

To demonstrate the different timing trajectories more clearly, we present the predicted value of net changes over time for the positive pledge and negative pledge groups, as well as in the full sample based on the quadratic model results (Models 2, 5, and 8). As shown in Figure 4, the cases with positive pledges clearly show the expected U-shaped pattern (the left panel) and we can observe the similar pattern in the full sample (the right most panel). The U-shaped pattern is, however, less apparent in the cases with negative welfare pledges: As we see in the panel in the middle, after an almost flat line in the beginning, a linear relationship after the turning point dominates. This relationship resembles the expectations from the classic electioneering literature.

**Figure 4: Predicted timing trajectory from quadratic regressions**



Note: The panels are based on the quadratic models in Table 1: Model 2 for positive pledge (left panel), Model 5 for negative pledge (center panel), and Model 8 for the full sample (right panel). The y-axis indicates the predicted value of net changes in welfare state legislations, calculated using the estimates for time components in the respective models and holding all other covariates at their means and the grey area denotes the confidence intervals at 90%.

Besides our variables of main interest, our models in Table 1 include variables controlling for the economic context (unemployment and growth), a veto player index and a dummy indicating whether a mandate was terminated by an early election. The coefficients of these covariates point into the expected directions, although it is only for the economic variables that we find significant effects: Higher growth compared to the quarter before and low unemployment is related to welfare state expansion – at least in the full sample (and the sample of positive pledges).

The country dummies do perform as one could expect from the qualitative literature with Finland being the outlier exhibiting much more expansive welfare state legislation than France. This is due to the fact that the Finnish welfare state was still expanding in the 1990s when other countries already implemented cutbacks and very much in line with qualitative studies of the Finnish welfare state which emphasize that the Aho and Lipponen governments in the 1990s protected insurance-based systems from cutbacks (Timonen 2003: 61), but also

with the trends visible in replacement rate data (Scruggs et al. 2013). Hence, the findings based on our data actually match the different path of Finland quite well. Again, not surprisingly given the qualitative literature on welfare state change in the UK (Clasen 2005), Britain is the country which has the most negative stance.

## **5 Conclusion**

This paper has investigated whether and how governments strategically time welfare state reforms across the electoral cycle. We add to the literature in two ways. First, we analyze this relationship on the basis of a novel dataset on changes to welfare state legislation in five Western industrialized countries. This direct measurement focusing on changes in welfare legislation is a major contribution, filling the gap in the existing studies that have often used either public spending data, which is not the best proxy for legislative change, or cutback events only hence not taking account for moments when cutbacks and expansions occur at the same time. Second, we develop and test a novel theoretical claim according to which the timing of welfare state changes does not follow the linear logic mostly assumed by PBC theories. Instead, we propose that a U-shaped relationship is more probable, at least when parties have promised to expand the welfare state in the election campaign (which indeed is the majority of the cases – more than 80% in our sample). This U-shaped relationship emerges because the mostly expansionary pledges tend to be fulfilled in the beginning of a mandate. Hence, while we agree that the recency bias modeled in PBC studies will yield expansionary welfare legislation toward the end of a mandate, we argue that the first months of a government's mandate could be divergent from the pattern because governments need to deliver on the expansionary promises. This is because voters' memory of the campaign is still fresh and not fulfilling the promises would result in a loss of trustworthiness, which may have long-term consequences for the government's future. The only exceptions from this rational should be the cases when governments promise to cut the welfare state in their manifesto. On

these rare occasions, we would expect a linear relationship with governments cutting right in the beginning and expanding toward the end of their mandate. At any rate, in this first period, two dynamics overlay: pledge fulfillment and the necessity to adopt cutbacks early on.

In our empirical analysis, our basic expectation about this U-shaped relationship is clearly supported. Using quadratic regression, we show the U-shaped relationship when government parties have pledged to expand the welfare state. In line with our expectation, too, the pattern is different in the case of negative pledges as our linear model provides the best results indicating that the trend in the net change is rather monotonically increasing over time. This is exactly what the PBC studies suggest. Taken together, these results support our argument that governments pursue pledge-keeping expansions *and* cutbacks at the same time during the first months in government.

Another obvious alternative explanation for the reluctance of governments to implement cutbacks right away may simply be that governments need time to legislate cutbacks. It should be noticed, though, that our data shows that governments do indeed adopt welfare state legislation during the first months in office (see Figure 3), which may also be explained by the fact that we look at individual legislative events instead of policy outcomes, such as spending.

Overall, this study suggests that governments indeed have had the capacity to steer policies at their discretion – at least to a certain extent. While it is certainly true that politicians are often time-pressed, not informed adequately and constrained by past legacies and institutions, our finding indicates that they indeed have considerable room to maneuver, against all odds. At least when it comes to the survival at power, politicians seem to have the capacity to engineer policies in a way that maximize their reelection chances.

## References

- Allan, James P., and Lyle Scruggs (2004). 'Political Partisanship and Welfare State Reform in Advanced Industrial Societies', *American Journal of Political Science*, 48:3, 496-512.
- Armingeon, Klaus, and Nathalie Giger (2008). 'Conditional punishment: A comparative analysis of the electoral consequences of welfare state retrenchment in OECD nations, 1980-2003', *West European Politics*, 31:3, 558-580.
- Arndt, Christoph, and Kees van Kersbergen (2015). 'Social democracy after the Third Way: restoration or renewal?' *Policy & Politics*, 43:2, 203-220.
- Baumgartner, Frank R., Christoffer Green-Pedersen and Bryan D. Jones (2006). 'Comparative studies of policy agendas', *Journal of European Public Policy*, 13:7, 959-974.
- Baumgartner, Frank, and Bryan D. Jones (1993). *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.
- Bechtel, Michael M., and Jens Hainmueller (2011): 'How Lasting Is Voter Gratitude? An Analysis of the Short- and Long-Term Electoral Returns to Beneficial Policy', *American Journal of Political Science*, 55:4, 852-868.
- Bonoli, Guiliano (2012). 'Blame Avoidance and Credit Claiming Revisited', in Bonoli, Guiliano, and David Natali (eds.), *The Politics of the New Welfare State*. Oxford: Oxford University Press.
- Brooks, Clem, and Jeff Manza (2007). *Why Welfare States Persist: The Importance of Public Opinion in Democracies*. Chicago: University of Chicago Press.
- Budge, Ian, and Richard I. Hofferbert (1990). 'Mandates and Policy Outputs: U.S. Party Platforms and Federal Expenditures', *American Political Science Review*, 84:1, 111-131.
- Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara, and Eric Tanenbaum (2001). *Mapping Policy Preferences. Estimates for Parties, Electors and Governments 1945-1998*. Oxford: Oxford University Press.
- Castles, Frank, and Herbert Obinger (2008): 'Worlds, Families, Regimes: Country Clusters in European and OECD Area Public Policy', *West European Politics*, 31:1, 321-344.
- Clasen, Jochen (2005). *Reforming European Welfare States: Germany and the United Kingdom Compared*. Oxford: Oxford University Press.
- Clasen, J. and Siegel, N.A. (eds.) (2007) *Investigating Welfare State Change. The 'Dependent Variable Problem' in Comparative Analysis*. Cheltenham/Northampton: Edward Elgar.
- Coglianesi, Cary (2009). 'The Transparency President? The Obama Administration and Open Government', *Governance*, 22:4, 529-544.
- De Haan, Jakob, and Klomp, Jeroen (2013). 'Conditional political budget cycles: a review of recent evidence', *Public Choice*, 157: 3-4, 387-410.
- Devries, Pete, Jaime Guajardo, Daniel Leigh, and Andrea Pescatori (2011). 'A New Action-based Dataset of Fiscal Consolidation'. Washington: International Monetary Fund (WP/11/128).
- Dominguez, Casey, and Byrne Knudsen (2005). 'Is It a Honeymoon? An Empirical Investigation of the President's First Hundred Days', *Congress & the Presidency*, 32:1, 63-78.
- Döring, Holger, and Philip Manow. 2018. *Parliaments and governments database (ParlGov): Information on parties, elections and cabinets in modern democracies*. Development version. <http://www.parlgov.org>.
- Duval, Dominic, and François Pétry (2019). 'Time and the Fulfillment of Election Pledges', *Political Studies*, 67:1, 207-223.
- Esping-Andersen, Gøsta (1990). *The three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- Fernández, Juan J. (2012). 'Economic crises, population aging and the electoral cycle: Explaining pension policy retrenchments in 19 OECD countries, 1981-2004', *International Journal of Comparative Sociology*, 53:2, 77-96.
- Flora, Peter (1986). *Growth to limits : the Western European welfare states since World War II*. Berlin u.a.: de Gruyter.
- Gelpi, Christopher, Peter D. Feaver, and Jason Reifler (2009). *Paying the Human Costs of War: American Public Opinion and Casualties in Military Conflicts*. Princeton: Princeton University Press.

- Giger, Nathalie, and Moira Nelson (2011). 'The electoral consequences of welfare state retrenchment: Blame avoidance or credit claiming in the era of permanent austerity?', *European Journal of Political Research*, 50:1, 1-23.
- Giger, Nathalie, and Moira Nelson (2013). 'The Welfare State or the Economy? Preferences, Constituencies, and Strategies for Retrenchment', *European Sociological Review*, 29:5, 1083-1094.
- Häusermann, Silja (2010). *The Politics of Welfare State Reform in Continental Europe: Modernization in Hard Times*. Cambridge: Cambridge University Press.
- Hallerberg, M., and Scartascini, C. 2017. Explaining Changes in Tax Burdens in Latin America: Do Politics Trump Economics? *European Journal of Political Economy* 48(1), 162-79.
- Healy, Andrew, and Gabriel Lenz (2014). 'Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy'. *American Journal of Political Science*, 58:1, 31-47.
- Hering, Martin (2008). 'Welfare State Restructuring without Grand Coalitions: The Role of Informal Cooperation in Blame Avoidance', *German Politics*, 17, 165-183.
- Hetherington, Marc (2005): *Why Trust Matters. Declining Political Trust and the Demise of American Liberalism*. Princeton: Princeton University Press.
- Hinterleitner, Markus, and Fritz Sager (2015). 'Avoiding Blame—A Comprehensive Framework and the Australian Home Insulation Program Fiasco', *Policy Studies Journal*, 43:1, 139-161.
- Hobolt, Sara, James Tilley, and Susan Banducci (2013). 'Clarity of responsibility: How government cohesion conditions performance voting', *European Journal of Political Research*, 52:2, 164-187.
- Hofferbert, Richard I., and Ian Budge (1992). 'The Party Mandate and the Westminster Model: Election Programmes and Government Spending in Britain, 1948–85', *British Journal of Political Science*, 22:2, 151-182.
- Horn, Alexander, and Carsten Jensen (2017). 'When and why politicians do not keep their welfare promises', *European Journal of Political Research*, 56:2, 381-400.
- Hooghe, Marc (2018). 'Trust and Elections', in Uslander, Eric M. (ed.), *The Oxford Handbook of Social and Political Trust*. Oxford: Oxford University Press.
- Huber, John D. (1996). *Rationalizing Parliament: Legislative Institutions and Party Politics in France*. Cambridge: Cambridge University Press.
- Huber, Evelyne, and John D. Stephens (2001). *Development and Crisis of the Welfare State*. Chicago/London: Chicago University Press.
- Hübscher, Evelyne (2016). 'The politics of fiscal consolidation revisited', *Journal of Public Policy*, 36:4, 573-601.
- Hübscher, Evelyne, and Thomas Sattler (2017). 'Fiscal consolidation under electoral risk', *European Journal of Political Research*, 56:1, 151-168.
- Jahn, Detlef (2010). 'The Veto Player Approach in Macro-Comparative Politics: Concepts and Measurement', in König, Thomas, George Tsebelis, and Marc Debus, (ed.), *Reform Processes and Policy Change. Veto Players and Decision-Making in Modern Democracies*. New York: Springer.
- Jensen, Carsten (2014). *The Right and the Welfare State*. Oxford: Oxford University Press.
- Jensen, Carsten, Christoph Arndt, Seonghui Lee, and Georg Wenzelburger (2017). 'Policy instruments and welfare state reform', *Journal of European Social Policy*, 28:2, 161-176.
- Kangas, Olli (2006): 'Finland: Labor markets against politics', in Immergut, Ellen, Karen M. Anderson, and Isabelle Schulze, (eds.), *The Handbook of West European Pension Politics*. Oxford: Oxford University Press, 248-296.
- Klingemann, Hans-Dieter (2006). *Mapping policy preferences. Estimates for parties, electors, and governments in Eastern Europe, European Union, and OECD 1990 - 2003*. Oxford: Oxford Univ. Press.
- Klingemann, Hans-Dieter, Andrea Volkens, Judith L. Bara, Ian Budget and Michal D. McDonald (2006). *Mapping Policy Preferences II. Estimates for parties, electors, and governments in Eastern Europe, European Union, and the OECD 1990-2003*. Oxford: Oxford University Press.
- Klingemann, Hans-Dieter, Richard I. Hofferbert and Ian Budge (1994). *Parties, Policies, and Democracy*. Boulder/San Francisco/Oxford: Westview.
- Klitgaard, Michael B., and Christian Elmelund-Præstekær (2014). 'The partisanship of systemic retrenchment: tax policy and welfare reform in Denmark 1975–2008', *European Political Science Review*, 6:1, 1-19.

- König, Pascal (2016). 'Communicating austerity measures during times of crisis: A comparative empirical analysis of four heads of government', *British Journal of Industrial Relations*, 18:3, 538-558.
- König, Pascal, and Georg Wenzelburger (2014). 'Towards a Theory of Political Strategy in Policy Analysis', *Politics & Policy*, 42:3, 397-427.
- König, Pascal, and Georg Wenzelburger (2017). 'Honeymoon in the Crisis? A comparative analysis of the strategic timing of austerity policies and their effect on government popularity in three countries', *Comparative European Politics*, 16:6, 991-1015.
- Kühner, Stefan. (2007). 'Country-level comparisons of welfare state change measures: another facet of the dependent variable problem within the comparative analysis of the welfare state', *Journal of European Social Policy* 17:1, 5-18.
- Lind, Jo Thori, and Halvor Mehlum (2010): 'With or without U? The appropriate test for a U shaped relationship', *Oxford Bulletin of Economics and Statistics*, 72:1, 109-118.
- MacKuen, Michael, Robert S. Erikson, and James A. Stimson (1992). 'Peasants or Bankers? The American Electorate and the U.S. Economy', *American Political Science Review*, 86:3, 597-611.
- Nordhaus, William D. (1975). 'The Political Business Cycle', *Review of Economic Studies*, 42:2, 169-190.
- Pierson, Paul (1994). *Dismantling the Welfare State? Reagan, Thatcher, and the Politics of Retrenchment*. Cambridge: Cambridge University Press.
- Pierson, Paul (1996). 'The New Politics of the Welfare State', *World Politics* 48:2, 143-179.
- Royed, Terry J., Kevin M. Leyden, and Stephen A. Borrelli (2000). 'Is 'Clarity of Responsibility' Important for Economic Voting? Revisiting Powell and Whitten's Hypothesis', *British Journal of Political Science*, 30:4, 669-698.
- Rudolph, Thomas J. (2017). 'Political trust as a heuristic', in Zmerli, Sonja, and Tom W.G. van der Meer, (eds.), *Handbook on Political Trust*. Cheltenham, UK: Edward Elgar.
- Schedler, Andreas (1998). 'The Normative Force of Electoral Promises', *Journal of Theoretical Politics*, 10:2, 191-214.
- Schumacher, Gijs, Barbara Vis, and Kees van Kersbergen (2013). 'Political parties' welfare image, electoral punishment and welfare state retrenchment', *Comparative European Politics*, 11:1, 1-21.
- Scruggs, Lyle, Detlef Jahn and Kati Kuitto (2013). *Comparative Welfare Entitlements Data Set 2, Version 2013*. <http://cwed2.org/>.
- Shi, Min, and Jakob Svensson (2006). 'Political Budget Cycles: Do They Differ Across Countries and Why?' *Journal of Public Economics*, 90:8-9, 1367-89.
- Smith, W. Rand (1990). 'Nationalizations for What? Capitalist Power and Public Enterprise in Mitterrand's France', *Politics & Society*, 18:1, 75-99.
- Strobl, D., Bäck, H, Müller, Wolfgang C., and Angelova, M. 2019. *Electoral Cycles in Government Policy Making: Strategic Timing of Austerity Reform Measures in Western Europe*, *British Journal of Political Science*. Forthcoming.
- Svallfors, Stefan (2012). 'Welfare States and Welfare Attitudes', in Svallfors, Stefan (ed.), *Contested Welfare States*. Stanford: Stanford University Press.
- Tepe, Markus, and Pieter Vanhuysse (2010). 'Who Cuts Back and When? The Politics of Delays in Social Expenditure Cutbacks, 1980–2005', *West European Politics*, 33:6, 1214-1240.
- Thomson, Robert, and Rory Costello (2016). 'Governing together in good and bad economic times: the fulfilment of election pledges in Ireland', *Irish Political Studies*, 31:2, 182-203.
- Thomson, Robert, Terry Royed, Elin Naurin, Joaquín Artés, Rory Costello, Laurenz Ennser-Jedenastik, Mark Ferguson, Petia Kostadinova, Catherine Moury, François Pétry and Katrin Praprotnik (2017a). 'Replication Data for: The Fulfillment of Parties' Election Pledges: A Comparative Study on the Impact of Power Sharing', Harvard Dataverse. <https://doi.org/10.7910/DVN/YJUJIBL>.
- Thomson, Robert, Terry Royed, Elin Naurin, Joaquín Artés, Rory Costello, Laurenz Ennser-Jedenastik, Mark Ferguson, Petia Kostadinova, Catherine Moury, François Pétry and Katrin Praprotnik (2017b). 'The Fulfillment of Parties' Election Pledges: A Comparative Study on the Impact of Power Sharing', *American Journal of Political Science*, 61:3, 527-542.
- Timonen, Virpi (2003). *Restructuring the Welfare State: Globalization and Social Policy Reform in Finland and Sweden*. Cheltenham: Edward Elgar.



- Tsebelis, George (1995). 'Decision Making in Political Systems: Veto Players in Presidentialism, Parliamentarism, Multicameralism, and Multipartyism', *British Journal of Political Science*, 25:3, 289-325.
- Vis, Barbara (2016). 'Taking Stock of the Comparative Literature on the Role of Blame Avoidance Strategies in Social Policy Reform', *Journal of Comparative Policy Analysis*, 18:2, 122-137.
- Volkens, Andrea, Pola Lehmann, Pola, Theres Matthieß, Nicolas Merz, Sven Regel, and Bernhard Weßels (2017). The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR). Version 2017b. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB). <https://doi.org/10.25522/manifesto.mps.2017b>
- Weaver, Kent R. (1986). 'The politics of blame avoidance', *Journal of Public Policy*, 6:4, 371-391.
- Wenzelburger, Georg (2011). 'Political Strategies and Fiscal Retrenchment: Evidence from Four Countries', *West European Politics*, 34:6, 1151-1184.
- Wenzelburger, Georg (2014). 'Blame avoidance, electoral punishment and the perceptions of risk', *Journal of European Social Policy*, 24:1, 80-91.
- Zohlnhöfer, Reimut (2007). 'The Politics of Budget Consolidation in Britain and Germany: The Impact of Blame Avoidance Opportunities', *West European Politics*, 30:5, 1120 -1138.

## Acknowledgements

We thank Alexander Horn and Carsten Jensen for sharing the data on welfare pledges. We also thank the anonymous reviewers for pointing out the relevance of portfolio allocation and the very helpful and thorough comments on this paper. Earlier versions of this paper have been presented at the Political Economy Research Colloquium at the University of Constance and at the Colloquium of the University of Kaiserslautern. Malin Grüninger, Kathrin Hartmann, Anja Konanec, Lasse Leipziger, Fabienne Müller and Kristian Nicolaisen provided first-class research assistance. We acknowledge funding by the Danish Council for Independent Research, grant no. 4003-00013.