

ARTICLE

# Beyond Positive and Negative: New Perspectives on Feedback Effects in Public Opinion on the Welfare State

Marius R Busemeyer<sup>1\*</sup>, Aurélien Abrassart<sup>2</sup> and Roula Nezi<sup>3‡</sup>

<sup>1</sup>Department of Politics and Public Administration, University of Konstanz, <sup>2</sup>Department of Economics, University of Berne and <sup>3</sup>Department of Politics, University of Surrey, Guildford, UK

\*Corresponding author. Email: Marius.Busemeyer@uni-konstanz.de

(Received 4 August 2016; revised 20 July 2018; accepted 16 October 2018; First published online 22 March 2019)

## Abstract

The study of policy feedback on public attitudes and policy preferences has become a growing area of research in recent years. Scholars in the tradition of Pierson usually argue that positive, self-reinforcing feedback effects dominate (that is, attitudes are commensurate with existing institutions), whereas the public thermostat model developed by Wlezien and Soroka expects negative, self-undermining feedback. Moving beyond the blunt distinction between positive and negative feedback, this article develops and proposes a more fine-grained typology of feedback effects that distinguishes between accelerating, self-reinforcing and self-undermining, specific and general, as well as long- and short-term dynamic feedback. The authors apply this typology in an analysis of public opinion on government spending in different areas of the welfare state for twenty-one OECD countries, employing a pseudo-panel approach. The empirical analysis confirms the usefulness of this typology since it shows that different types of feedback effects can be observed empirically.

**Keywords** policy feedback; welfare state policies; public opinion; institutionalism

A recent commentary on the reaction of public opinion to policies of the newly installed Trump administration in the United States hints at the existence of powerful feedback effects: across the board, Teixeira (2017) observes a resurgence of ‘liberal’ (in the US sense) values such as support for free trade, more public spending on different areas of the welfare state and a generally larger role for government. This represents strong support for the ‘public as thermostat’ model and its conception of negative policy feedback (Soroka and Wlezien 2010; Wlezien 1995), which argues that public opinion usually moves in the opposite direction of policy making, balancing out policy makers’ urges to move too strongly in a particular direction. However, Teixeira also notes that, for the first time since 2010, there are more supporters for Obama’s health care reform – the Affordable Care Act (ACA) – than opponents, while Republican proposals to reform or replace the ACA are generally unpopular. This, in turn, indicates strong support for the notion of positive or self-reinforcing policy feedback as promoted by Pierson (1993, 1994, 2000) and others. According to this argument, once established, policies create their own political support, making their abolishment unlikely.

Critical citizens may welcome these data. Researchers, however, are confronted with different conceptions of the notion of ‘policy feedback’, which – according to their particular logic – are supported by stylized facts. Often coined in terms of positive (Piersonian) versus negative (thermostatic) feedback (Fernández and Jaime-Castillo 2013), these literatures suggest seemingly

<sup>‡</sup>The original version of this article was published with errors in the author names and affiliations. A notice detailing this has been published and the errors rectified in the online and print PDF and HTML copies.

contradictory findings: the dominance of self-reinforcing feedback effects on the one hand, and the dominance of self-undermining effects on the other (Jacobs and Weaver 2015). In this article, we deal with the conceptual confusion that has emerged in the debate about policy feedback in recent years. Thus, the main goal of this article is to contribute to an integration of different theoretical and empirical perspectives on the notion of policy feedback, which so far have largely ignored each other. As a first step towards this goal, we propose a novel typology of feedback effects,<sup>1</sup> which might serve as a common reference point for future research on policy feedback that moves beyond the blunt distinction between positive and negative feedback usually found in the literature.

We argue that this distinction is prone to misunderstandings, since the terms are used in very different ways (Baumgartner and Jones 2002; Jacobs and Weaver 2015; Kumlin and Stadelmann-Steffen 2014a; Kumlin and Stadelmann-Steffen 2014b; Soroka and Wlezien 2005). As an alternative, we propose a more fine-grained typology that distinguishes three dimensions: the direction of feedback (accelerating, self-reinforcing or self-undermining), the broadness/scope of feedback (general, specific), and the time dimension (long-term/institutional, short-term/dynamic). In a short empirical section, we also demonstrate the applicability and usefulness of the typology in describing the multitude of feedback effects that have been observed empirically. We also comment briefly on the hitherto neglected empirical implications and challenges associated with testing the notion of policy feedback in quantitative research designs. While acknowledging that a significant share of recent US literature is on the impact of policies on patterns of political participation (Campbell 2012; Mettler and Soss 2004), we focus on policy feedback on popular attitudes and preferences (following Svallfors 2012).

The remainder of the article is structured as follows. In the next section, we provide a brief overview of the existing literature, including some thoughts on its shortcomings and open questions. Next, we develop our theoretical argument and present our more fine-grained typology of feedback effects. This is followed by an empirical analysis of public opinion data from the International Social Survey Programme (ISSP) from the 1996 and 2006 waves.

## Literature Review

The increasing availability of large-scale cross-national surveys of public opinion such as the European Social Survey (ESS) or the ISSP has spawned a vibrant literature studying the feedback effects of existing institutions and policies on public attitudes towards the welfare state. Pierson's (1993, 1994, 2000, 2004) seminal work on 'policy feedback effects', building on previous work by Skocpol (1992) and Esping-Andersen (1990), provided the theoretical framework that inspired this new wave of research in the 2000s and after. Whereas Esping-Andersen (1990) was primarily interested in why different types of welfare state regimes have emerged, Pierson was more concerned with explaining why they are sustainable politically in the long term and why large-scale reforms, primarily cutbacks, are increasingly unlikely the longer welfare states persist.<sup>2</sup> Pierson distinguished between resource/incentive effects on the one hand (that is, the effects of institutions on the distribution of material resources across different welfare state beneficiary groups), and interpretive effects on the other, which refer to the effect of institutions on 'cognitive processes of social actors' (Pierson 1993, 610) and – one could add – their normative orientations vis-à-vis the welfare state.<sup>3</sup> He further identified three levels on which feedback effects could materialize: government elites, 'social groups' and the mass public (Pierson 1993, 624).

Pierson's basic argument is that feedback effects reinforce existing institutions, and contribute to their long-term political sustainability and path dependency in policy development. In the

---

<sup>1</sup>In this, we follow previous contributions, which raised a similar criticism and started to move in a direction that is similar to ours (Jacobs and Weaver 2015; Kumlin and Stadelmann-Steffen 2014b).

<sup>2</sup>See also the work of Brooks and Manza (2006, 2007), which follows in the tradition of Pierson.

<sup>3</sup>See the important work of Rothstein (1998) and Svallfors (2010) on this point.

literature, this type of feedback is often labeled ‘positive feedback’ (Weaver 2010, 137), since popular attitudes and actors’ strategies positively reinforce the institutional status quo. In Pierson’s original theory (Pierson 2000, 2004), however, the term positive feedback is actually used to describe the process of path formation in the initial stages of development. Since Pierson is (also) interested in explaining the diversity of existing welfare state regimes, he needs to explain why countries, beginning from a similar starting point in the past, move away from this ‘equilibrium’ to develop distinct sets of institutions. In this sense, positive feedback effects occur when the likelihood of moving away from some kind of pre-existing equilibrium increases the further one moves away (a similar use of the term positive feedback can be found in the literature on political agendas, see Baumgartner and Jones 2002).

In the comparative welfare state literature, the notion of reinforcing feedback inspired a new wave of scholarship on the impact of welfare state regimes on public attitudes in the 2000s and beyond, which made use of the newly available cross-national survey data (Andresß and Heien 2001; Arts and Gelissen 2001; Blekesaune and Quadagno 2003; Finseraas 2009; Jæger 2006; Jæger 2009; Jæger 2013; Svallfors 1997; Svallfors 2004; Svallfors 2010; for recent overviews, see Kumlin and Stadelmann-Steffen 2014a and Svallfors 2012). This literature starts with the expectation that public attitudes towards the welfare state should be congruent with institutions, either because institutions have shaped attitudes in the way Pierson claimed or because policies and institutions reflect prevailing patterns of public opinion (Rehm 2012). According to this logic, public support for the welfare state should be strongest in the universalist social democratic welfare states in the Scandinavian countries and weakest in the liberal Anglo-Saxon world, with continental European countries located in between (Arts and Gelissen 2001). The empirical evidence for this claim is mixed. In a recent update of the literature and using a sophisticated methodological approach, Jæger (2009), for example, finds little overall support for the expected pattern. One reason for these mixed findings might be that the early contributions to the literature often used simple dummy variables to capture the influence of welfare state regimes, whereas more precise measures of different institutional dimensions of welfare state policies should be used instead (Busemeyer 2013; Gingrich and Ansell 2012; Jakobsen 2009). Another reason, however, could be that existing institutions do not necessarily and always create patterns of public opinion that are commensurate with the prevailing status quo.

This idea of patterns of public opinion moving in ways that do not reinforce, but instead undermine, existing institutions and policies is often discussed under the label of ‘negative feedback effects’ (Weaver 2010; see also Jacobs and Weaver 2015). According to Weaver’s understanding, negative feedback effects occur when existing institutions and policies have ‘negative socio-economic consequences’ that ‘undermine political support for those policies’ (Weaver 2010, 139; see also Fernández and Jaime-Castillo 2013). This definition is problematic, because whether a certain social phenomenon is considered to be a problem with implications for policy making is itself a matter of public debate. For example, a large and generous welfare state might be perceived as an economic burden in one context or as an important social safety net in another.

The notion of ‘negative feedback effects’ also occupies a central place in the ‘*public as thermostat*’ model developed by Wlezien and Soroka (Soroka and Wlezien 2005; Soroka and Wlezien 2010; Wlezien 1995; Wlezien 2004), but it is used in a very different way compared to Weaver (2010). The ‘public as thermostat’ approach is rooted in the literature on public opinion and responsiveness of policy making (for example, Erikson, MacKuen and Stimson 2002; Page and Shapiro 1983; Stimson, MacKuen and Erikson 1995), and therefore is not well connected to the scholarly debates in the welfare state literature, even though these literatures study similar issues. According to the ‘public as thermostat’ model, the relationship between public opinion on the one hand and the world of policy making on the other is dynamic: citizens perceive and react to changes in public policies by expressing support or opposition, which is then picked up by responsive policy makers. In the words of Wlezien (1995, 982): ‘In effect, the public would behave like a thermostat, where a

departure from the favored policy temperature (which itself can change) produces a signal to adjust policy accordingly, and once sufficiently adjusted, the signal stops.’ Couched in different terms, this means that if there are feedback effects from policies on public opinion, they should be negative (Soroka and Wlezien 2010, 28): if public spending on a particular policy item is too high/too low, the public will react by expressing support for less/more spending. Policy makers interested in winning elections pick up these signals and change their spending priorities accordingly. Even though the connection between citizen-voters and policy makers is mediated by political institutions and differences in the salience of particular policies, Soroka and Wlezien find strong empirical support for their model in general and across a large number of country cases (Soroka and Wlezien 2010; Wlezien and Soroka 2012; see also Bartle, Dellepiane-Avellaneda and Stimson 2011).

### Theoretical Inconsistencies in the Existing Literature

Previous studies in the comparative welfare state and public opinion research literatures have evaluated both the feedback effects of existing institutions and policies on public opinion, but they come to different – even contradictory – findings. The most obvious contradiction is that researchers working in the tradition of Pierson (1993) would argue that ‘positive’ (that is, self-reinforcing) feedback effects dominate. In contrast, Wlezien and Soroka (2012, 1409) expect a dominance of ‘negative’, self-undermining feedback effects. At the same time, the terms ‘positive’ and ‘negative’ have different meanings in these literatures, as we explain in greater detail below. This obvious contradiction, however, is just the starting point and masks deeper underlying challenges for any theory of policy feedback. Some of these challenges are mainly methodological questions of how to measure and operationalize feedback effects (Wlezien 2016).

While we comment on empirical challenges below, the focus of this article is to point out theoretical inconsistencies and problems that are often overlooked or underestimated in the pertinent literature. Jacobs and Weaver (2015) and Kumlin and Stadelmann-Steffen (2014b) have recently made important contributions to reassessing the conceptual foundations of policy feedback theory. This article builds on these contributions, and moves beyond existing work by developing a more fine-grained and rigorous typology of policy feedback effects, which is applicable in both quantitative and qualitative research designs.<sup>4</sup>

In our view, the very success of the concept of policy feedback as it is applied in different literatures has compromised its ‘internal coherence’ and therefore also its differentiation potential, which according to Gerring (1999) are central quality criteria in concept formation. Even though scholars *within* the respective literatures introduced above usually managed to provide coherent and convincing definitions of the concept of policy feedback *in their respective domains*, the ubiquitous usage of the concept has led to blurred conceptual boundaries and confusion. In the next section, we first provide a short account of the most relevant theoretical inconsistencies (from our point of view), followed by a proposal for a more differentiated *typology* of policy feedback effects, which – we hope – might help to mitigate some of the aforementioned inconsistencies. In this exercise, we side with Collier, LaPorte and Seawright’s (2012, 228) claim that the development of typologies can contribute to ‘useful conceptual innovation’ if they are clear about their underlying theoretical concepts and dimensions.

### Generalizability of Feedback Theories

The first inconsistency relates to implicit (or sometimes explicit) claims about the generalizability of the implications of policy feedback theories. Admittedly, this is not necessarily attributable to the scholars who originally coined these concepts in their particular domains and are usually very

---

<sup>4</sup>The contribution by Jacobs and Weaver (2015) is focused on qualitative research designs only, whereas Kumlin and Stadelmann-Steffen (2014b) only present a short outlook on the contours of future conceptual debates about policy feedback without going into much detail.

careful about setting out their scope conditions. But there is certainly a tendency in the literature to determine which type of feedback is the dominant one (see, for example, Fernández and Jaime-Castillo 2013). This is particularly the case in quantitative research designs, in which the terms ‘positive’ and ‘negative’ can be immediately operationalized in terms of regression coefficients.

There are several issues to be discussed with this approach. The first is that the different literatures on policy feedback capture very different aspects of the dynamics of policy making: the policy consequences of long-term historical processes of institutional change and development, which may be ‘slow-moving’ (Pierson 2004: 90); the rather short-term dynamics of the interaction between policy makers and citizens (Soroka and Wlezien 2010; Wlezien 1995); and agenda-setting processes (Baumgartner and Jones 2002). Hence, the different literatures – without necessarily reflecting on this – use very different time scales when talking about policy feedback (Kumlin and Stadelmann-Steffen 2014b, 317). As a consequence, the terms ‘positive’ and ‘negative’ feedback have been used in different ways. Although there are some parallels, subtle differences in the underlying concepts get lost in most research designs used in the literature cited above. For instance, ‘negative feedback’ in the sense of Baumgartner and Jones (2002) and Soroka and Wlezien (2010) is associated with the balancing out of contravening forces, that is, a kind of ‘self-correcting mechanism’ (Baumgartner and Jones 2002, 8). At the same time, this kind of dynamic might contribute to the stabilization of historical development paths and hence reinforce the notion of ‘positive’ feedback in the Piersonian (1993) sense. Pierson (1993, 1994, 2000) employs the notion of ‘positive feedback’ to explain why certain countries deviated from a common path at critical junctures in historical development, and which mechanisms ensure that they remain committed to these development paths in the future. Similarly, for Baumgartner and Jones (2002), positive feedback sets in motion a self-reinforcing dynamic in agenda-setting processes, which indicates a break with the (more common) stability in political agendas, eventually promoting policy change. Thus for Pierson (1993), positive feedback means policy stability, whereas it is expected to increase the likelihood of policy change according to Baumgartner and Jones (2002).

The second issue regarding the generalizability of feedback theories is that if one type of feedback strongly dominates, this would have empirical implications that are hard to sustain. For instance, if self-reinforcing feedback dominated, existing welfare state regimes would never (or hardly) change. In fact, this is exactly what Pierson (1994, 2011) and others (Brooks and Manza 2006; Brooks and Manza 2007) posit: the effects of path dependency are so strong that policy change is incremental and path dependent at best.

There are a number of reasons to doubt this strong version of path dependency theory, however. First, it contrasts with new theoretical perspectives in historical institutionalism (Streeck and Thelen 2005), which argue that even seemingly minor and incremental institutional change can accumulate and become path-changing transformative change over long periods of time. Secondly, in welfare state research, many have challenged Pierson’s claims, pointing out significant instances of change and reform in contemporary welfare states in the last decade, which often combine retrenchment in one part of the welfare state with cautious expansion in others (Bonoli 2013; Häusermann 2010; Hemerijck 2013). Furthermore, the empirical evidence on attitudes mentioned above (such as Jæger 2006, 2009) reveals that there is no straightforward correlation between welfare state generosity and the level of public support: that is, there are cases where citizen-voters demand changes in the status quo rather than stability.

If the opposite is true – that is, if negative/self-undermining feedback effects dominate – we should observe a convergence of policy preferences and eventually policy output in the long term. Admittedly, the ‘public as thermostat’ model (Soroka and Wlezien 2010) is mostly concerned with studying the dynamic process of feedback and responsiveness between the public and policy makers, and less with the long-term historical trajectories of policy change. Thus, different from work in the tradition of historical institutionalism, it is less concerned with explaining the

emergence of cross-country variation in policies and institutions as such. But if the notion of negative feedback is generalized beyond the short-term dynamic process of feedback and responsiveness, it seems to imply a gradual convergence of citizens' preferences and eventually policies as low-spending countries are catching up to high-spending countries, which remain more or less in place. Independent of whether convergence is actually happening in contemporary welfare states, the more important issue for our purposes is to point out the limitations of the thermostat model in theorizing about differences in 'preferred levels' of policies across country contexts and their long-term development.

### *Reinforcing vs. Accelerating and Absolute vs. Relative Preferences*

The second inconsistency is about the distinction between self-reinforcing and 'accelerating' feedback effects, which is also related to the distinction between 'absolute' and 'relative' preferences discussed in the feedback literature (Kumlin and Stadelmann-Steffen 2014b, 317; Wlezien 2016). Again, there are several issues to discuss.

First, the feedback literature in the tradition of Pierson (1993, 1994, 2000) is ambivalent about whether existing institutions set in motion an accelerating, positive feedback dynamic or merely a self-reinforcing one. In the former case, citizens living in and getting used to generous universal welfare states develop preferences in support of an even larger and more encompassing model (Kumlin and Stadelmann-Steffen (2014b, 318) describes this as 'adaptive' feedback). If feedback effects are merely reinforcing, citizens would cherish existing institutions and policies but not necessarily exhibit support for a further extension of the scope of the welfare state. Jacobs and Weaver's discussion of this point mirrors this ambivalence: 'Accounts of policy feedback have emphasized the *self-reinforcing* effects of policies on their own politics: the ways in which policies bolster their own bases of political support, *yielding either policy stability or an expansionary dynamic over time*' (Jacobs and Weaver 2015, 441, emphasis added). In developing a comprehensive theory of policy feedback, it is obviously important to be able to state whether existing institutions merely promote policy stability or whether they set in motion an accelerating, expansionary dynamic.

There is support for both sides of the argument: citizens might value existing policies if they derive concrete benefits from them, but not support the expansion of other welfare state benefits and services out of self-interest (these are Pierson's (1993) 'resource effects'). Alternatively, citizens may adopt their beliefs and cognitive frames and support the normative underpinnings of universalist welfare state policies even if they do not benefit from them directly (Pierson's (1993) 'interpretive effects' or Svallfors' (2010) 'normative feedback effects'). Solving this issue is an empirical question, which we do not directly address here. Instead, we want to highlight that we are currently *lacking the conceptual tools needed to differentiate and distinguish between these different types of feedback effects*.

Secondly, we also note a certain lack of attention in the literature to the distinction between 'absolute' and 'relative' preferences. According to Wlezien (2016), absolute preferences refer to citizens' 'preferred levels of policy', whereas relative preference captures preferences *relative to the current policy status quo*, that is, whether citizens want 'more' or 'less' spending/government activity on a particular policy. Roosma, Oorschot and Gelissen (2014) also distinguish between 'attitudes about what people believe the welfare state should do, and attitudes towards its actual performance',<sup>5</sup> which partly overlaps with Wlezien's distinction. Distinguishing between absolute and relative preferences holds particular empirical challenges, as extensively discussed by Wlezien (2016). On a more conceptual level, the important point is to distinguish between citizens' preferences on 'preferred levels of policy' (Wlezien 2016, 7), independent of the current status quo, and their preferences regarding changes to this status quo. A similar point has been

<sup>5</sup>See also Meuleman and Chung 2012; Toikko and Rantanen 2017 for similar arguments.

made by Kumlin and Stadelmann-Steffen (2014b, 318), who posit that ‘reactive feedback concerns attitudes to policy change’, whereas ‘adaptive [feedback] processes may concern basic norms about which absolute policy “level” is desirable’.

The confusion occurs when researchers do not clearly spell out which type of preferences are addressed, which could lead to inaccurate conclusions about the relative importance of accelerating, self-reinforcing and self-undermining feedback effects. For instance, a negative association between lagged levels of spending and support for additional spending on a given policy might be interpreted as a negative thermostatic effect, but this does not necessarily imply that it would be a self-undermining effect in the sense that citizens fundamentally change their views about the appropriate role of government in the welfare state regime. It might simply indicate a stabilization effect at a high level of welfare state generosity. Policy developments in some Scandinavian welfare states in recent years might be a good real-world example of this dynamic. Partial liberalization and marketization of the provision of welfare state services has occurred – often in response to public demands for more decentralized provision – but a whole-scale dismantling of the Nordic welfare state model has not taken place (Gingrich 2011). In a similar vein, Roosma, Oorschot and Gelissen (2014) as well as Meuleman and Chung (2012) show that dissatisfaction with the current status quo in the provision of welfare state services does not necessarily correlate with general support for the welfare state: citizen dissatisfaction with welfare state performance can lead to support for higher public spending (in order to ameliorate the deficiencies) or lower support (due to distrust in the government). Thus, following Roosma, Oorschot and Gelissen (2014) and Wlezien (2016), it is important to distinguish between general attitudes towards the underlying normative principles of welfare states on the one hand, and more concrete demands regarding particular changes in the status quo on the other.

### **Towards a Synthesis: A Typology of Feedback Effects**

The previous section discussed a number of theoretical inconsistencies in the literature, which have emerged as a consequence of conceptual confusion related to the widespread use of the notion of policy feedback in different literatures. We believe the typology of different types of feedback effects, which we develop below, can contribute to ‘conceptual innovation’ (Collier, LaPorte and Seawright 2012, 226) by improving the internal coherence of the concept of policy feedback and therefore its usefulness in studying different phenomena. In this exercise, we avoid the ambiguous and often misleading labels of ‘positive’ and ‘negative’ feedback effects, since they often mean different things in different contexts and literatures (Jacobs and Weaver 2015). We explore the empirical relevance and usefulness of this typology in the empirical section of this article.

Devising a more complex and differentiated typology of feedback effects implies that these different types of effects can also be observed in the real world. Instead of assessing whether positive or negative feedback effects dominate in principle or ‘on average’, scholars should instead strive to identify the contextual conditions that explain why certain types of feedback occur in some instances, but not in others. As the example mentioned at the beginning of the introduction shows, different kinds of feedback dynamics might be going on at the same time. The purpose of this article is not to provide a definitive theory that explains these contextual conditions, although we discuss this issue in the concluding section. The goal is more modest – to give scholars new conceptual tools to describe the multitude of feedback effects in a more coherent manner.

Eventually, this conceptual innovation can contribute to a more encompassing synthesis of the core insights of the various literatures outlined above. This will most likely contribute to developing a deeper understanding of the complex linkages between the macro world of policies and institutions and the micro level of attitudes and preferences. In fact, to some extent, Piersonian and thermostatic feedback theory already provide insights that are complementary to each other: the

**Table 1.** Summary presentation of typology of policy feedback effects

<i>Direction of feedback</i>		
Accelerating (policies creating support for further expansion)	Self-reinforcing (policies creating support for continued provision)	Self-undermining (policies creating opposition to their continued provision)
<i>Broadness/scope of feedback</i>		
General (feedback effects from welfare state regimes in total or large set of government activities)	Specific (feedback effects from particular policies or policy changes)	
<i>Time dimension</i>		
Long-term institutional (feedback effects from past (historical) policy choices)	Short-term dynamic (feedback effects from recent policy changes)	

institutional and political mechanisms that Pierson identifies as being responsible for path dependency (Pierson 1993, 2000) can be expressed as thermostatic (or self-correcting) feedback in the frameworks of Soroka and Wlezién (2010) and Baumgartner and Jones (2002). Historical–institutionalist approaches may, in turn, be more useful in explaining cross-national differences in ‘preferred levels of policy’, which are largely taken as exogenous in the thermostat model. This is because historical–institutionalist approaches are less concerned with understanding the short-term fluctuations and dynamics of public opinion, and focus instead on the long-term impact of institutions on attitudes, that is, absolute preferences in the sense defined above. Again, developing common conceptual tools help to stimulate this kind of exchange between different literatures.

According to Collier, LaPorte and Seawright, good typologies should clearly spell out the ‘overarching concept’ (Collier, LaPorte and Seawright 2012, 222), which serves as a reference point for the various subcategories, as well as the dimensions used in the construction of the typology. The overarching concept is the notion of policy feedback, which can be broadly defined as the effect of existing policies and institutions on individual-level attitudes and preferences.<sup>6</sup> Following Pierson (2006), we conceive of policies as ‘institutions’ in the sense that feedback effects are not caused by the policy output as such (for example, public spending), but by institutional arrangements created and modified by the actions of policy makers. Attitudes refer to general predispositions towards the welfare state, in particular its normative underpinnings, which might not necessarily have immediate policy consequences, for example, support for the principle of distributive justice. In contrast, (policy) preferences are related to concrete policy options and demands, such as support for or opposition to public spending.

The central dimensions in our typology refer to (1) the direction of feedback, (2) its broadness/scope and (3) the time lag of feedback effects. As discussed above, researchers should also pay more attention to the distinction between abstract and relative (or concrete) preferences, but this distinction is not part of the typology, because it refers to characteristics of the dependent variable (attitudes and preferences). Previous contributions by Soroka and Wlezién (2010), Jacobs and Weaver (2015) as well as Kumlin and Stadelmann-Steffen (2014b) have already alluded to some of these dimensions, but they have not yet been combined into an overarching systematic framework. Instead of drawing a three-dimensional matrix and then discussing potential combinations of feedback types in each individual cell, we discuss each of the three dimensions subsequently and leave it to future researchers to theorize in greater detail about more or less plausible combinations of the different characteristics. Table 1 presents a brief summary of the dimensions.

<sup>6</sup>As mentioned above, a broader definition of policy feedback also includes feedback from institutions and policies on political participation and electoral behavior (Campbell 2012; Kumlin and Stadelmann-Steffen 2014a), but we focus on policy feedback on attitudes and preferences in this article.

### **Direction of Feedback**

The direction of feedback is clearly the dimension discussed most in the pertinent literature, but it is certainly not the only one that matters. In order to deal with some of the theoretical inconsistencies discussed above, we refrain from using the terms ‘positive’ and ‘negative’ feedback, because these may actually indicate different directions of feedback depending on the particular context. We also move beyond the dichotomous distinction between ‘self-reinforcing’ and ‘self-undermining’ feedback effects, as this blurs the important distinction between a truly accelerating (or expansionary) feedback dynamic and a merely reinforcing one. Furthermore, it is important to point out that the dimension of direction of feedback is relevant for both absolute as well as relative preferences. Citizens’ preferences can signal a desire for change in current policy settings relative to the status quo (relative preferences), or they can indicate a shift in the underlying ‘preferred level of policy’ (absolute preferences). More specifically, we distinguish between:

- *Accelerating feedback effects* occur when the expansion of a given policy creates further support for an expansion in the provision of this and similar policies. This implies, for instance, even more support for an expansion of the welfare state and increases in public spending in generous, high-spending welfare states, but – since this is merely about the direction of feedback – it could in theory also indicate further support for retrenchment in residual welfare states.
- *Self-reinforcing feedback effects* occur when a given policy creates support for its continued provision, but not for further expansion of this and similar policies. In this setting, public opinion indicates that citizens are largely satisfied with the prevailing status quo and tend to oppose both further expansion as well as significant retrenchment.
- *Self-undermining feedback effects* occur when the promotion of a given policy creates opposition to its continued provision (or that of similar policies). In residual welfare states, this might imply support for increased government spending and welfare state expansion, whereas in universalist welfare states, it would indicate opposition to spending increases or further welfare state expansion.

### **Breadth/Scope of Feedback**

This dimension relates to the question of whether public attitudes reflect changes and levels in specific policies or whether the public simply responds in a more undifferentiated manner to broader changes in welfare state or other, even broader, policy regimes.<sup>7</sup> This also relates to the question of how important policies (and policy changes) are as determinants of individual-level attitudes and preferences relative to other socio-economic factors. Jæger (2013), for instance, finds that socio-economic variables such as GDP growth and inequality are more relevant determinants of preferences than policies (social expenditures). In principle, it would be possible to define the boundaries of the feedback concept very broadly, including feedback effects from policies as well as the broader socio-economic context. However, we believe a more focused definition of policy feedback that refers only to feedback effects emanating from existing policies and institutions is empirically more useful and theoretically more convincing. Thus we differentiate between two types of effects:

- *General feedback effects* refer to the effects of welfare state regimes or a large set of government activities on individual-level attitudes and preferences. Citizens might not be able or willing to process subtle differences in the dynamics of individual policy fields and only react to large-scale changes in welfare state policy making and governmental activities. Wlezien (2004) confirms this expectation for the case of the United States. The work by Stimson and colleagues also shows that citizens’ policy preferences react to broad changes in

<sup>7</sup>See Wlezien (1995, 2004) for a previous discussion on this point.

the ‘policy mood’ of government activities rather than specific policy changes (Bartle, Dellepiane-Avellaneda and Stimson 2011; Erikson, MacKuen and Stimson 2002; Stimson, MacKuen and Erikson 1995). An important point to consider is that general feedback effects might affect broad aggregates of public opinion at the micro level (general left–right positions or redistribution preferences) or the effects might spill over into attitudes and preferences for particular policies, that is, the association between total public spending and support for a particular social policy.

- *Specific feedback effects* refer to the effects of particular policies or policy changes on public attitudes about the policy. It could be argued that citizens are able to perceive and react to changes in specific policy domains, in particular if these policies are very important to their individual well-being. Following in the footsteps of Svallfors (2003, 190) on this point, Jordan (2013) argues (and finds empirical support for the claim) that a disaggregated perspective on policy feedback is empirically and theoretically more convincing, because using broad aggregates of government activity and public opinion masks a significant amount of variation across policy fields. The work by Naumann (2014) on retirement policies and by Van Oorschot and Meuleman (2014) on unemployment shows that citizens in fact perceive and react to specific institutional characteristics of these welfare state programs.

Although we believe this strong distinction between general and specific feedback effects is useful for defining the ideal-typical endpoints of the underlying dimension, we also want to emphasize that from an empirical perspective, medium-specific feedback effects might be particularly relevant. Applied to the example of welfare state policies, this would imply the following: citizens’ attitudes might be particularly sensitive to changes in welfare state policies (for example, levels of total social spending), but not necessarily to changes in an even larger set of government activities, which would also include economic, environmental or defense policies, or to very specific policy changes in particular domains of social policy making.

### *Time Dimension*

The time dimension of policy feedback is treated very differently in different literatures. Policy feedback theories in the tradition of Pierson focus on processes, which take a significant amount of time to unfold. In the general framework of historical institutionalism, cross-national differences in institutions are traced back to policy decisions during critical junctures in historical development. Studies in the tradition of historical institutionalism commonly use qualitative research designs, and document the impact of existing institutions on actors’ strategies and preferences with a focus on organized interests rather than the mass public. The recent wave of scholarship on the impact of welfare state regimes on popular attitudes (for example, Andreß and Heien 2001; Blekesaune and Quadagno 2003; Jæger 2006, 2009; Svallfors 1997; Svallfors 2004) works with the assumption that cross-country comparisons of attitudes within regime types would reveal the accumulated long-term effects of previous changes. In contrast to the *longue durée* perspective of historical institutionalism, the ‘public as thermostat’ model (Soroka and Wlezien 2010; Wlezien 1995) is based on the idea of a dynamic interactive relationship between changes in policy output and changes in public attitudes, which is related to Stimson, MacKuen and Erikson’s (1995) concept of ‘dynamic representation’. Even though these models certainly pay a lot of attention to the time dimension, they focus largely on the relationship between public opinion and policy making in the short term (by historical standards) as they stipulate that the mass public reacts dynamically to policy changes.

In sum, we find different perspectives in the literature on the role of time and timing in policy feedback. It remains an open empirical question whether feedback effects materialize mostly in the short term as dynamic feedback effects or mostly in the long term. In fact, both might be going on at the same time: the public may react in a ‘thermostatic’ manner to policy changes in

the short term, while also being confined to a long-term development path that sets certain cognitive, normative and resource-based limitations on the range of feasible policy options (Kumlin and Stadelmann-Steffen 2014b, 318). This is also in line with empirical work which shows that public attitudes change slowly over time (Stegmueller 2014) as they represent deeply held beliefs and normative orientations, which are often formed during individuals' formative years. Attitudinal change therefore often occurs via generational replacement rather than individuals changing their beliefs and attitudes (Svallfors 2010). To come back to the example mentioned at the beginning of the introduction, American public opinion may slowly turn towards favoring more government involvement in health care in reaction to the policies of the new administration, but it is still unlikely that it will completely turn towards supporting a fully fledged European-style public model. Again, we distinguish between two types of effects.

- *Long-term institutional feedback effects* occur when past policy choices shape popular attitudes over a long period of time via normative, cognitive and resource-based feedback mechanisms.
- *Short-term dynamic feedback effects* refer to the dynamic interaction between policy makers and public opinion in the short term, when the public immediately reacts to (perceived) policy changes.

Table 1 presents a summary of the proposed typology of feedback effects. Again, it is important to highlight that different combinations of feedback effects might be observed empirically in the real world. The primary purpose of this typology is therefore to provide the researcher with a more fine-grained conceptual toolbox, which helps to develop more sophisticated theoretical expectations on the conditional factors that might explain the emergence of one type of feedback or another. The feedback types on the poles of the individual dimensions that make up the typology should be understood as ideal-typical endpoints in a continuum rather than crisply defined and qualitatively distinct categories.

### Empirical Implications and Challenges

In the remainder of the article, we turn from theory to empirics. To be clear, the purpose of our empirical analysis is not to provide a definitive (and final) answer as to which type of feedback dominates the others. Rather, we use it to demonstrate the usefulness and applicability of our proposed typology in describing and conceptualizing the multitude of feedback effects observed empirically. Before we do that, we comment briefly on the challenges that emerge when testing the empirical implications of policy feedback theory. Many of these challenges are seldom discussed in the pertinent literature, and we cannot necessarily provide easy solutions at this point. It is still important to think more thoroughly about this topic in order to understand the apparent discrepancies between the different literatures as mentioned above.

The first point to discuss is *the challenges of testing long- and short-term feedback effects in quantitative research designs*. The most common research design in empirical tests of policy feedback theory is to include a macro-level variable on policies or institutions as an independent variable in regressions of public attitudes. This is plausible in the case of short-term dynamic feedback effects, as respondents are expected to perceive and react to short-term policy changes. But it gets more challenging the longer the time lag. Operationalizing feedback effects as cross-national differences between regimes is the most common way to deal with longer time lags, but it is not a very convincing operationalization of policy feedback theory as it carries the implicit assumption that all kinds of differences between countries are related to feedback of past policy choices. In this broad interpretation, the concept of policy feedback loses analytical rigor and becomes less useful as a conceptual tool to understand the dynamics of the interaction between the mass public and policy makers. Furthermore, it does not seem to make sense – from either a

theoretical or empirical perspective – to include levels of spending from the 1970s (or even before?) in regressions of present-day attitudes, because the causal chain between these past decisions and contemporary public opinion is simply too long.

The opposite is also true: including only recent, short-term policy changes in regressions of public attitudes assumes a high level of attentiveness and political interest on the part of citizens. It might well be the case that citizens need some time to perceive and react to policy changes; that is, there might be a significant, but not indeterminate, time lag until policy changes ‘trickle down’ to the level of popular attitudes. But how long is ‘significant’? Obviously, this question is difficult to answer from a purely theoretical perspective, but from an empirical point of view, it is equally problematic to assume that all kinds of feedback effects will materialize immediately in the short term. This is also because it might take some time for policy changes to have real effects on socio-economic outcomes. In established Organisation for Economic Co-operation and Development (OECD) democracies, regular elections typically take place every four or five years. This might serve as a plausible upper limit to the length of lags operationalizing long-term feedback since the link between past policies and attitudes becomes too tenuous otherwise. Recent work by Adams and Somer-Topcu (2009) confirms the importance of mid-range feedback effects, albeit from a different perspective. They show that voters reward political parties electorally for policy moderation at the *previous* (rather than the current) election, even though the previous election is typically four to five years in the past.

The bottom line of this discussion is that researchers should spend more time thinking about the time dimension when operationalizing feedback effects. The time dimension may also interact with other dimensions of policy feedback. For instance, it may be plausible to assume a short-term dynamic feedback connection in some cases (for instance, in combination with specific feedback effects), but the time lags might be longer in other cases (such as in the case of general feedback effects). Furthermore, testing the historical–institutionalist variety of policy feedback theory in its entirety poses additional challenges as long-term historical feedback effects are difficult to operationalize by way of institutional regime indicators.

The latter is also related to the second point to be discussed: *the fact that feedback effects materialize via both macro- and micro-level factors*. The common operationalization of policy feedback effects via macro-level variables underestimates the fact that policy feedback also occurs at the micro level. In fact, the longer back we go, the more likely it is that the latter channel gains in importance. For instance, Pierson’s (1993) argument about resource-based feedback effects simply implies that beneficiaries of welfare state services and benefits develop an interest in their continued provision. In the literature on welfare state attitudes, this is most commonly referred to as ‘material self-interest’ and usually tested at the micro level.

Furthermore, as was also famously argued by Esping-Andersen (1990), welfare states change the socio-economic fabric of societies, that is, the *size* and *composition* of welfare state beneficiary groups. This (quite obvious) point is often neglected in quantitative analyses of feedback effects since these typically focus on average effects of micro-level variables. Even if the average effect of micro-level variables is similar across regimes, the simple fact that the *size* of beneficiary groups varies across regimes is an important manifestation of feedback effects as it has implications for the general levels of political support for a particular policy. A common finding in the literature on welfare state attitudes is that attitudes and preferences are also shaped by norms and values above and beyond pure self-interest (Kangas 1997; Margalit 2013). The relative size of partisan constituencies subscribing to a particular ideology is likely to vary across countries (and over time) as well, even if the *average* effect of left–right ideology on welfare state attitudes and preferences is the same across countries. There might simply be more people identifying themselves with a ‘left-wing’ ideology in a country like Sweden compared to the United States. This is an important empirical implication of policy feedback theory, which gets lost when focusing on the average effects of micro-level variables only or when assuming that the entire feedback effect will only materialize via macro-level variables.

## Empirical Analysis

In this section, we perform an empirical analysis of commonly available public opinion data. As mentioned above, the purpose of this analysis is not to settle the debate between historical-institutionalist and ‘thermostat’ perspectives on policy feedback by identifying one (or even: *the*) dominant type of feedback effect. Rather, we engage in a broad analysis of different types of feedback effects, and demonstrate the applicability and usefulness of the new typology developed above. Thus in some respects the empirical analysis fits well with the theoretical concepts introduced above, while at the same time, we fully acknowledge that the empirical analysis cannot address all the issues and problems discussed before and needs to be complemented with additional analyses.

Regarding aspects where theory and empirics fit well, our analytical set-up allows us to assess the direction of policy feedback by simply reading off the sign of regression coefficients of macro-level variables, which is arguably the easiest point to address. Furthermore, we operationalize the broadness/scope of feedback dimension by employing measures of total social spending and spending on particular social policies. Finally, we address the time dimension by including different time lags of spending levels (as well as long-term averages). Given the intrinsic difficulties of measuring long-term feedback effects in quantitative research designs as discussed in the previous section, this is still an imperfect solution, but it is an improvement on existing research approaches, which rarely consider long-term lag structures.

There are a number of limitations to our analysis, however. First of all, the empirical analysis is more concerned with measuring feedback effects on relative preferences, since the dependent variables used below capture support for more or less public spending in a given area. Mostly for reasons of space, we cannot consider feedback effects on absolute preferences here. Furthermore, our empirical set-up is helpful to identify accelerating and self-undermining feedback, but it has limitations regarding the assessment of genuine self-reinforcing feedback, that is, support for maintaining the status quo. This is, however, also a limitation of much of the existing literature, although it is rarely discussed as such. Below, we present some very basic descriptive statistics to help determine the extent of self-reinforcing feedback, but if taken seriously, more sophisticated and new survey instruments would need to be developed.<sup>8</sup> What is more, we do not have a strong theory about which factors might explain *why* feedback effects vary across policy fields (that is, why we should expect accelerating, long-term feedback effects in the case of education, but self-undermining, dynamic feedback in the case of unemployment, for instance). In the conclusion we discuss potential conditional factors that might explain which type of feedback is more likely to be observed and how these might be related to policy characteristics. Finally, since our research design is quantitative, we do not illustrate how our typology can be applied in qualitative designs, although we believe it is helpful independently of the choice in methods. We take this point up in the concluding section as well. Despite these limitations, we believe it is useful to confront our theoretical concepts with empirical data.

## Data and Variables

We use survey data from the ISSP Role of Government module. We combine two waves of this survey (from 1996 and 2006) into a comprehensive dataset, covering a total of twenty-one OECD countries, thirteen of which were included in both waves, for a total of thirty-four country-years.<sup>9</sup>

<sup>8</sup>We thank the anonymous reviewers of this article for their remarks on this point.

<sup>9</sup>The countries are: Australia, Canada, Czech Republic, Denmark (only 2006), Germany, Finland (only 2006), France, Great Britain, Hungary, Ireland, Japan (only 2006), the Netherlands (only 2006), New Zealand, Norway, Poland, Portugal (only 2006), Slovenia, Spain (only 2006), Sweden (only 2006), Switzerland and the USA. The reason for the absence of certain countries in the 1996 wave is that Denmark, Finland, the Netherlands and Portugal were only included in the 2006 wave, and the income and occupation variable (ISCO classification) only contain missing values in 1996, respectively, for Italy and Spain, and Great Britain, Japan and Sweden.

Even though policy output and welfare state attitudes can, of course, relate to many different aspects of policy making, we focus on public spending here. This is in line with the approach pursued by Wlezien and Soroka (Soroka and Wlezien 2010; Wlezien 1995; Wlezien and Soroka 2012). The focus on spending has the advantage of creating a direct and clear connection between policy output and attitudes, independently of the question of what should actually be done with additional spending.

The ISSP data provide a comparative measure of individual attitudes towards spending in different policy areas of social spending, which is included in both waves. This is a relative measure of individual preferences, because individuals are asked for their support for deviations from the current status quo. The question is worded as follows:

Listed below are various areas of government spending. Please show whether you would like to see more or less government spending in each area. Remember that if you say 'much more', it might require a tax increase to pay for it.

Respondents' replies to this question are measured on a five-point scale (much more, more, the same as now, less, much less). In the later analysis, this categorical variable is aggregated into a continuous variable of group-level support for spending (see below) and recoded to range from -2 (much less spending) to +2 (much more spending). Since welfare state policies are considered to be the highly salient issues in policy making, we focus on the comparison of the attitudes-output link in four areas of welfare state politics: unemployment, pensions, health care and education.

Our main independent variables are actual levels of government spending as a percentage of GDP for each policy area. To capture the dynamic aspect, we include measures of public spending in the specific area (unemployment, pensions, health care and education) lagged by one to five years, as well as a five-year average starting from the first lag. In addition to public spending in specific areas, we include a measure of total spending on social policy in order to be able to distinguish between specific and general feedback effects.<sup>10</sup>

Complementing the battery of variables measuring actual levels of government spending, we include a set of individual-level variables that have been found to have a strong influence on attitudes towards spending. Previous analyses have stressed the importance of partisan ideology (Jacoby 1994). The ISSP surveys measure ideology, derived from party choice, on a five-point scale variable ranging from far left (1) to far right (5). We expect that left-wing voters will support higher levels of spending on welfare policies. Similarly, income is included as an independent variable at the individual level, since it is expected to have an important impact on attitudes towards spending (Meltzer and Richard 1981). The continuous measure of income provided in the survey has been transformed into a measure based on income deciles in order to improve comparability across countries.

In addition, a set of individual-level characteristics such as age, social class and educational background were included in the construction of groups for the pseudo-panel analysis (see next section). Age was recoded in three main categories (15-34, 35-54, 55 and more). Social class is based on the Erikson, Goldthorpe and Portocarero (EGP) scheme obtained from ISCO88 occupations and is also divided into three main categories: the working class (EGP class V-VII), the middle class (EGP class III and IV) and the upper class (EGP class I and II). Education was recoded as primary/lower secondary (ISCED 1-2), upper secondary (3-4) or tertiary education (ISCED 5-6).

<sup>10</sup>The macro-level data come mostly from the Comparative Political Data Set (Armingeon *et al.* 2016). For the measure of spending on education, data were compiled from OECD *Education at a Glance* reports for the corresponding years. When data points were missing between two available years, the average of the previous and next values was computed and imputed. Finally, data on total spending on social policy for Slovenia between 1990 and 1995 were recoded as missing due to incomplete information.

Finally, we include a number of macro-level control variables. Income inequality (the net Gini index, taken from Solt (2009)) is expected to be positively associated with support for additional public spending, since higher levels of inequality should increase the demand for compensatory social spending (Finseraas 2009; Meltzer and Richard 1981). Higher levels of unemployment should have the same effect, particularly in the case of spending on unemployment, whereas GDP growth should depress demand for social spending, because individuals are less dependent on the welfare state.<sup>11</sup> We use different specifications – only GDP growth as control, only economic variables as controls (GDP growth plus unemployment rate) and the full set of controls including the Gini index – in order to test the robustness of our findings.

### *Methods and Models*

In this article, we move beyond existing scholarship by employing a pseudo-panel analysis. This method has several advantages for our purposes compared to simple multi-level models. Instead of using individuals as the basic unit of analysis, we construct groups of individuals with similar individual-level characteristics. Thus it becomes possible to compare groups of individuals across the two different waves, assuming that these represent ‘typical’ cases, even though we do not deal with real panel data, but two repeated cross-sections. Calculating group-level averages instead of using individuals as the basic unit of analysis and including group dummies absorbs a significant amount of variation at the micro level, which is an advantage if one is primarily interested in the effects of macro-level contexts on attitudes as we are here.

To our knowledge, the pseudo-panel approach has not been applied to the study of welfare state preferences, with the sole exceptions of Olivera (2015) and Jæger (2013). The choice of variables defining the grouping of observations is crucial to the robustness of the results (Jæger 2013). Hence, we follow the lead of Jæger in using age, education and social class as micro-level variables and socio-demographic characteristics to link the two distinct waves of the ISSP included in the analysis. Each of these three variables is then divided into three categories (see above), which results in a matrix of  $3 \times 3 \times 3$  subdivisions, that is, twenty-seven distinct groups for each country and year or survey. For each country-year, we therefore calculate twenty-seven groups with different socio-demographic combinations (young, low-educated, low social class; young, low-educated, middle class, etc.). Obviously, not all groups will have the same number of observations and may even have none (there are few young, low-educated, upper-class people around, for instance). Further, some groups are characterized by a lack of variation of the dependent variable (for example, when all respondents in that group demand more or much more spending) and are therefore excluded from the analysis.

Once the different groups are created, the next step is to estimate the mean values of the dependent and independent variables at the micro level for each sub-group. Since age, education and social class are already used to construct the groups, they are not directly included in the regression. However, we include two additional micro-level control variables (income and partisan ideology), for which we again calculate mean values for each sub-group nested in country-years. In the case of the dependent variables, average values were also computed within each sub-group nested in country-years to determine average levels of support for spending.

Due to the clustering of group-level observations within country-years, we apply multilevel modeling techniques using maximum likelihood estimation. Accordingly, we include random intercepts at the country level to account for the fact that the group-level observations share common characteristics within countries. Because we are mostly interested in cross-country variation in our dependent variables, models with country dummies instead of random intercepts are discarded. Year fixed effects are also included to account for wave-specific effects and to isolate the variation of the dependent variable across countries while controlling for its variation over time. Finally, group dummies are included in all models to deal with potential endogeneity problems resulting from unobserved characteristics correlated with the variables used to create the groups.

<sup>11</sup>The data source for these other macro-level variables is also the Comparative Political Data Set (Armingeon et al. 2016).

More formally, the model can be specified as:

$$\bar{A}_{ijt} = \alpha + \beta_1 \bar{I}_{ijt} + \beta_2 \bar{P}_{ijt} + \beta_3 S_{jt} + \beta_k C_{jt}^k + \zeta_j + Y_t + G_i + \varepsilon_{ijt}, \quad (1)$$

where  $i$  stands for groups,  $j$  for countries and  $t$  for years.  $\bar{A}$  is the average value of preference for more spending for each sub-group, country and year,  $\bar{I}$  the mean income (in deciles), while  $\bar{P}$  stands for the average of the index of partisanship,  $S$  for spending as a percentage of GDP on the different areas of social policy as well as total social spending using the different lags as discussed above.  $C^k$  stands for the  $k = \{1-3\}$  control variables at the macro level. Finally,  $\zeta_j$  are the country random intercepts,  $Y_t$  and  $G_i$ , the year and group fixed effects, respectively.

In interpreting the coefficient estimates of the macro-level spending variables, we adopt a rather broad operationalization of feedback effects, which focuses on the *direction* of feedback effects and therefore does not take into account whether respondents generally support or oppose spending in the first place. Thus, a positive coefficient estimate indicates a higher demand for more spending in more generous countries, but also preferences for less spending in countries that tend to spend less (accelerating effect). When the coefficient estimate is negative, individuals respond to high levels of spending by demanding less of it and react to low levels by demanding more spending (self-undermining). In the robustness section below, we also take into account a narrower definition of feedback effects and discuss why this particular operationalization is more problematic than the broader operationalization, which we use in the main analysis.

## Findings

Before we discuss the results from the regression analyses in detail, we want to briefly comment on the distribution of respondents across the different categories of the dependent variables (see Figure 1). This is important because a high share of respondents indicating a preference for maintaining current spending levels could be evidence of self-reinforcing feedback. Figure 1 shows that the modal category in these distributions is one, which is significantly above the zero mark – with the notable exception of unemployment spending. For education, health and pension spending, citizens – on average – express a desire to increase current levels of public spending rather than maintaining current levels of cutting back. In the case of unemployment, support is more or less normally distributed around the zero mark. Looking at the relative share of respondents in the middle category using the raw individual-level rather than the aggregated group-level data (see Appendix Table A1) reveals similar findings. These data also show that there is significant variation across countries in the size of the middling category, reaching from 5.95 per cent of respondents supporting the status quo in health care to 60.49 per cent of Danes supporting the existing level of spending on unemployment (both figures are for the 2006 wave).

In sum, these findings indicate that self-reinforcing feedback is indeed a meaningful category, different from accelerating and self-undermining feedback. However, interpreting the share of respondents expressing a preference for maintaining current levels of spending as evidence of self-reinforcing feedback is – at best – an imperfect proxy. First of all, there are methodological concerns as respondents might choose the middling category not because this reflects their actual preferences, but because they are uncertain and ambiguous about how to respond and therefore opt for the middle ground (Goerres and Prinzen 2012, 518). Secondly, the lower average for support for unemployment spending might simply reflect the fact that citizens are less fond of supporting the expansion of welfare state benefits in this area as (potential) recipients are regarded as less ‘deserving’ (Van Oorschot 2006) and not a genuine desire for maintaining the current status quo. Given the currently available data, there are no good alternatives to tackle these problems in a convincing manner. Hence, future research should devote more attention to properly measuring support for maintaining the status quo relative to alternatives, for example by employing factorial survey methods. Alternatively, if panel datasets are available, self-reinforcing feedback can be modeled as a process over time, in which respondents gradually

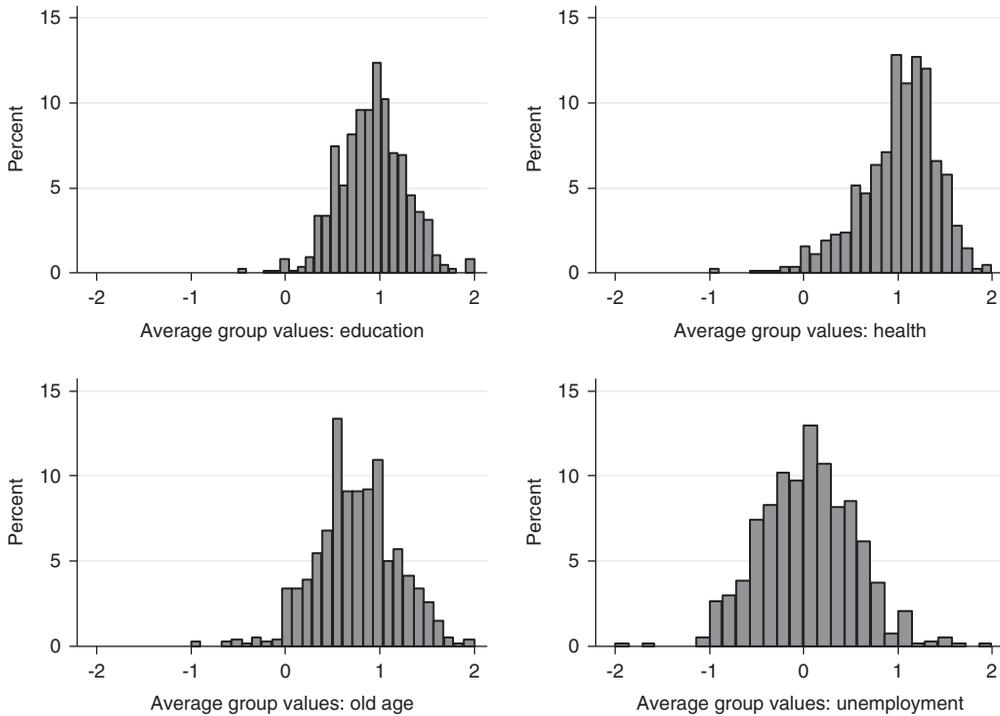


Figure 1. Descriptive statistics about the distribution of respondents for different policy fields

adapt their attitudes after a significant change in spending has occurred.<sup>12</sup> Further analyses could also explore whether support for the status quo is in any way related to welfare state institutions. A consistent interpretation of policy feedback theory suggests that there should be no association between welfare state institutions and the extent of self-reinforcing feedback effects as citizens' attitudes adapt to different institutional contexts.

The first analyses displayed in Table 2,<sup>13</sup> focus on the specific feedback effect across different time lag specifications (from  $t - 1$  until  $t - 5$ , plus the five-year average). We include different types of social spending commensurate with the dependent variable, that is, public spending on education, on health care, on pensions and on unemployment as macro-level independent variables affecting micro-level support for more spending on these policies.

Our first major finding is that the type of feedback effect depends on the field of spending and the time lag. The case of unemployment shows the most consistent and robust evidence of a self-undermining and specific feedback at work. Higher levels of spending on unemployment are associated with lower levels of support for spending increases. The effect holds across different time lags (including the five-year average lag) and different sets of control variables. For education, by contrast, the association between macro-level spending and micro-level support is mostly not significant, except for  $t - 1$ , when we find a negative association, that is, evidence of a self-undermining feedback effect in the short term when we include less than the full set of control variables. For health, we also find evidence of a self-undermining (specific) feedback logic, but the association is less consistent and robust compared to the case of unemployment as it varies depending on the controls included in the models. There is no evidence of any statistically significant association between spending on pensions and popular support for more or less

<sup>12</sup>This approach has already been partly pursued by Soroka and Wlezién (2010) for a select number of countries and with simple time-series data only, although they focus on thermostatic rather than self-reinforcing feedback.

<sup>13</sup>For more detailed results and the full regression tables, see Appendix Tables A4–A11.

**Table 2.** Specific feedback effects across policy fields with different time lags

	Lag 1	2	3	4	5	5-year average
<b>Only GDP growth as control</b>						
Education	-0.0603** (0.0304)	0.0326 (0.0287)	0.0384 (0.0326)	0.0258 (0.0317)	0.00324 (0.0313)	0.0291 (0.0382)
Health	-0.136*** (0.0436)	-0.0787** (0.0307)	-0.120*** (0.0309)	-0.103*** (0.0237)	-0.0550** (0.0247)	-0.114*** (0.0310)
Retirement	-0.0124 (0.0471)	-0.0387 (0.0396)	-0.0473 (0.0377)	-0.0301 (0.0349)	-0.0151 (0.0301)	-0.0300 (0.0377)
Unemployment	-0.0839** (0.0374)	-0.0975*** (0.0349)	-0.0993*** (0.0267)	-0.0950*** (0.0249)	-0.0628* (0.0341)	-0.0639* (0.0360)
<b>Only economic controls</b>						
Education	-0.0608** (0.0305)	0.0312 (0.0285)	0.0292 (0.0329)	0.0194 (0.0336)	-0.0129 (0.0324)	0.0161 (0.0387)
Health	-0.0884* (0.0470)	-0.0111 (0.0334)	-0.0461 (0.0355)	-0.0361 (0.0285)	0.0224 (0.0282)	-0.0320 (0.0363)
Retirement	0.0566 (0.0587)	0.00141 (0.0446)	-0.00588 (0.0434)	0.00255 (0.0384)	0.0161 (0.0329)	0.0118 (0.0429)
Unemployment	-0.0678* (0.0400)	-0.0891** (0.0418)	-0.123*** (0.0367)	-0.105*** (0.0310)	-0.108** (0.0458)	-0.0820* (0.0419)
<b>All controls</b>						
Education	-0.0418 (0.0315)	0.0283 (0.0272)	0.0270 (0.0315)	0.0199 (0.0349)	0.000862 (0.0403)	-0.0265 (0.0472)
Health	-0.0895* (0.0473)	-0.0504 (0.0346)	-0.0653* (0.0351)	-0.0558** (0.0283)	0.000758 (0.0285)	-0.0583 (0.0361)
Retirement	0.0567 (0.0588)	0.00116 (0.0448)	-0.00605 (0.0436)	0.00268 (0.0385)	0.0168 (0.0333)	0.0118 (0.0430)
Unemployment	-0.0594 (0.0405)	-0.0844** (0.0422)	-0.115*** (0.0373)	-0.0972*** (0.0320)	-0.116** (0.0474)	-0.0832** (0.0421)

Note: standard errors in parentheses. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

spending in this policy field. This could be related to the particularly high degree of institutional inertia in this policy field or to a high degree of support for the prevailing status quo (that is, self-reinforcing feedback in the above defined way).

In the next set of regressions, we focus on general feedback effects. For that purpose, we replace the macro-level variables on spending on specific areas of social policy with total levels of social spending to test whether the feedback effect is specific or non-specific. As before, we use one- to five-year lags of spending levels, as well as the five-year average (Table 3). While the effect is now more consistent across time lags, important differences remain in the support for spending across the different fields included. In the case of education, for example, we find more robust evidence of an accelerating feedback effect, in particular for time lags longer than three years as well as for the five-year average. In contrast, overall social spending is not significantly related to popular support for spending on unemployment at all. In the case of health care, there is evidence of a long-term self-undermining feedback effect, but only in model specifications, which include merely GDP growth as a control. By contrast, we find a statistically significant self-undermining feedback effect in the case of pension spending. The effect is consistently significant across model specifications and different time lags. It is interesting to note that in the case of pensions, overall levels of social spending – general feedback effects – are more relevant than specific feedback effects, whereas we observe the opposite for unemployment. Thus it appears that attitudes towards spending on education and retirement are more strongly linked to overall social spending and in a rather long-term dynamic. On the contrary, demand for more or less spending on unemployment appears to be more sensitive to the specific levels of spending in this field.

As briefly mentioned above, we also employ a narrower operationalization of policy feedback.<sup>14</sup> It could reasonably be argued that a positive/negative sign of the coefficient estimate is

<sup>14</sup>We would like to thank an anonymous reviewer for pointing out this possible alternative operationalization.

**Table 3.** General feedback effects across policy fields with different time lags

	Lag 1	2	3	4	5	5-year average
<b>Only GDP growth as control</b>						
Education	-0.0127 (0.0508)	0.0221 (0.0476)	0.0452 (0.0429)	0.0617 (0.0388)	0.0453 (0.0387)	0.0446 (0.0477)
Health	-0.141*** (0.0446)	-0.106*** (0.0382)	-0.126*** (0.0349)	-0.0816** (0.0325)	-0.0667** (0.0332)	-0.107*** (0.0367)
Retirement	-0.113** (0.0457)	-0.125*** (0.0384)	-0.138*** (0.0347)	-0.115*** (0.0313)	-0.0936*** (0.0316)	-0.125*** (0.0364)
Unemployment	-0.0213 (0.0587)	-0.0391 (0.0476)	-0.0475 (0.0424)	-0.0402 (0.0385)	-0.0301 (0.0388)	-0.0404 (0.0454)
<b>Only economic controls</b>						
Education	0.0357 (0.0626)	0.0616 (0.0508)	0.108** (0.0479)	0.0970** (0.0397)	0.0893** (0.0412)	0.0984* (0.0514)
Health	-0.0904* (0.0524)	-0.0545 (0.0423)	-0.0526 (0.0390)	-0.0258 (0.0348)	0.0107 (0.0369)	-0.0391 (0.0413)
Retirement	-0.0957** (0.0488)	-0.113*** (0.0400)	-0.123*** (0.0367)	-0.102*** (0.0327)	-0.0755** (0.0340)	-0.110*** (0.0384)
Unemployment	-0.0195 (0.0597)	-0.0384 (0.0483)	-0.0493 (0.0444)	-0.0403 (0.0394)	-0.0300 (0.0405)	-0.0403 (0.0467)
<b>All controls</b>						
Education	0.0378 (0.0655)	0.0763 (0.0592)	0.134** (0.0523)	0.120*** (0.0438)	0.114** (0.0459)	0.124** (0.0575)
Health	-0.0695 (0.0569)	-0.0420 (0.0431)	-0.0415 (0.0397)	-0.0188 (0.0348)	0.0179 (0.0363)	-0.0263 (0.0419)
Retirement	-0.116** (0.0555)	-0.120*** (0.0416)	-0.130*** (0.0380)	-0.104*** (0.0332)	-0.0766** (0.0345)	-0.116** (0.0399)
Unemployment	-0.0326 (0.0661)	-0.0444 (0.0499)	-0.0552 (0.0458)	-0.0431 (0.0400)	-0.0326 (0.0411)	-0.0461 (0.0483)

Note: standard errors in parentheses. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

merely a necessary, but not sufficient, condition to document accelerating/self-undermining feedback. This is because accelerating feedback would suggest that citizens are in general supportive of spending increases (that is, the value of our dependent variable should be above zero), because otherwise, the positive coefficient estimate could also mean a weakening of opposition to spending increases and not genuine support for more spending. The same logic holds for self-undermining feedback (in reverse). We believe this narrow interpretation of policy feedback is too stringent and opt for a broader operationalization, which focuses merely on the direction of feedback. Also, as Figure 1 shows, in most cases, respondents are in 'positive territory' anyway, expressing some support for spending increases. Hence, a positive coefficient estimate signals a genuine accelerating effect, whereas a negative coefficient estimate signals self-undermining feedback mostly in the sense of weakening opposition to further spending increases. This is plausible in the context of welfare state politics, where citizens rarely express open opposition to welfare state expansion.

Nevertheless, we ran additional analyses applying a narrow operationalization of feedback. More specifically, to measure accelerating feedback, we recoded the dependent variables so that a value of 1 amounts to support for more or much more spending and 0 represents the remaining categories (the same or less). To measure self-undermining feedback, we recorded the dependent variables: 1 captures support for less or much less spending vs. 0 for the remaining categories (the same or more). With these narrower operationalizations, a positive coefficient estimate of the macro-level spending variables constitutes evidence of an accelerating effect in the first case and a self-undermining effect in the second case.<sup>15</sup> The number of observations (groups) is significantly

<sup>15</sup>The coefficient estimate is expected to be positive in the latter case because of the reversed coding: higher levels of spending would then be associated with (positive) support for spending cuts.

**Table 4.** Robustness tests applying a narrower operationalization of feedback effects

	Lag 1	2	3	4	5
<b>Specific feedback effect</b>					
<i>Accelerating FE</i>					
Education	-0.0160 (0.0161)	0.0167 (0.0142)	0.00999 (0.0158)	0.0111 (0.0169)	0.0165 (0.0210)
Unemployment	-0.0281* (0.0166)	-0.0363** (0.0173)	-0.0404*** (0.0154)	-0.0304** (0.0135)	-0.0448** (0.0199)
<i>Undermining FE</i>					
Education	-0.00324 (0.00487)	-0.00838* (0.00485)	-0.00626 (0.00478)	-0.00384 (0.00458)	-0.00244 (0.00460)
Unemployment	0.0255 (0.0167)	0.0369** (0.0175)	0.0437*** (0.0154)	0.0377*** (0.0131)	0.0179 (0.0181)
<b>General feedback effect</b>					
<i>Accelerating FE</i>					
Education	0.0118 (0.0298)	0.0359 (0.0284)	0.0651** (0.0277)	0.0556** (0.0228)	0.0540** (0.0238)
Unemployment	-0.0209 (0.0276)	-0.0287 (0.0212)	-0.0316 (0.0196)	-0.0281 (0.0173)	-0.0294 (0.0180)
<i>Undermining FE</i>					
Education	-0.00153 (0.00595)	-0.00116 (0.00573)	-0.00141 (0.00589)	-0.000150 (0.00553)	-4.97e-05 (0.00555)
Unemployment	-0.0159 (0.0237)	-0.0106 (0.0192)	-0.000400 (0.0182)	-0.000376 (0.0160)	-0.0152 (0.0162)

Note: standard errors in parentheses. \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1

reduced in the second case because only a small number of the respondents expressed support for less or much less spending. This is another, more methodological reason why we stick with the broader operationalization in the main part of the analysis. As a consequence, the variable does not vary at all in many groups (because respondents are either happy or want more or much more spending, yielding a value of 0). These groups have to be excluded from the analysis.

In spite of these limitations, Table 4 shows that applying a narrower operationalization yields similar results to the previous analysis. We focus on the two cases of unemployment and education here for reasons of space and since the previous analysis had also revealed relatively strong feedback effects for these two areas. Table 4 confirms the existence of a statistically significant specific self-undermining effect in the case of unemployment *and* a specific ‘negative accelerating’ feedback – a weakening of accelerating feedback. Put differently, this finding confirms that a weakening accelerating feedback and a genuine self-undermining effect are essentially the same thing; that is, it matters little whether the value of the dependent variable is ‘positive’ or ‘negative’, as the direction of the feedback effect is ultimately the decisive factor. For education, the results in Table 4 confirm previous evidence of a long-term accelerating effect. In this case, the reverse (a weakening self-undermining effect) cannot be observed, most probably for the simple reason that the number of respondents supporting cutbacks in education spending is even lower than in the other policy fields given the strong popularity of education. In short, this additional robustness test strengthens our confidence that the findings obtained using the broader operationalization of feedback in the main part of the analysis are robust to different operationalizations and model specifications.

In order to further boost our confidence in the robustness of the results from the pseudo-panel analysis, we performed the following sensitivity analyses:<sup>16</sup>

<sup>16</sup>Yet another sensitivity analysis was to replicate the results after filling in the data gaps regarding spending on education using alternative data sources, namely the World Bank and national-level sources. Differences in data quality between the sources, however, unfortunately prevented us from drawing any conclusions about the validity of the new results (available upon request).

- A replication of the previous analyses excluding the groups of fewer than twenty;<sup>17</sup>
- A replication of the previous analyses including only countries with data from both waves, 1996 and 2006.

The results (see Appendix Tables A2 and A3) are in general consistent with the findings of the previous models. Some differences emerge from the first replication in the case of the general feedback effect: the effect for education is no longer significant, while the one for unemployment gains in significance (Table A2). These differences might be due to the important loss of information that accompanies the exclusion of rare groups with attitudes that are significantly different from the rest of the population. In contrast, replicating the analyses including only countries with data from both waves yielded very similar results to the original analyses (Table A3).

As a final robustness test, we replicated the pseudo-panel analysis using a Bayesian hierarchical modeling framework. We use the same data as before and the same model specification with the same dependent and independent variables. Bayesian hierarchical modeling has become increasingly popular for studying the association between macro-level contexts and micro-level attitudes. A vexing problem in this field of research is that the number of macro-level units (countries) is usually quite small (less than twenty), casting doubt on the robustness and validity of the findings obtained. Recent work by Stegmueller (2013) has shown that Bayesian approaches may be superior to inference-based maximum likelihood estimation methods when the number of country cases is below twenty, since the latter produces standard errors that are too small. Detailed results as well as some background on the methodological choices can be found in Appendix Tables A12–A19.

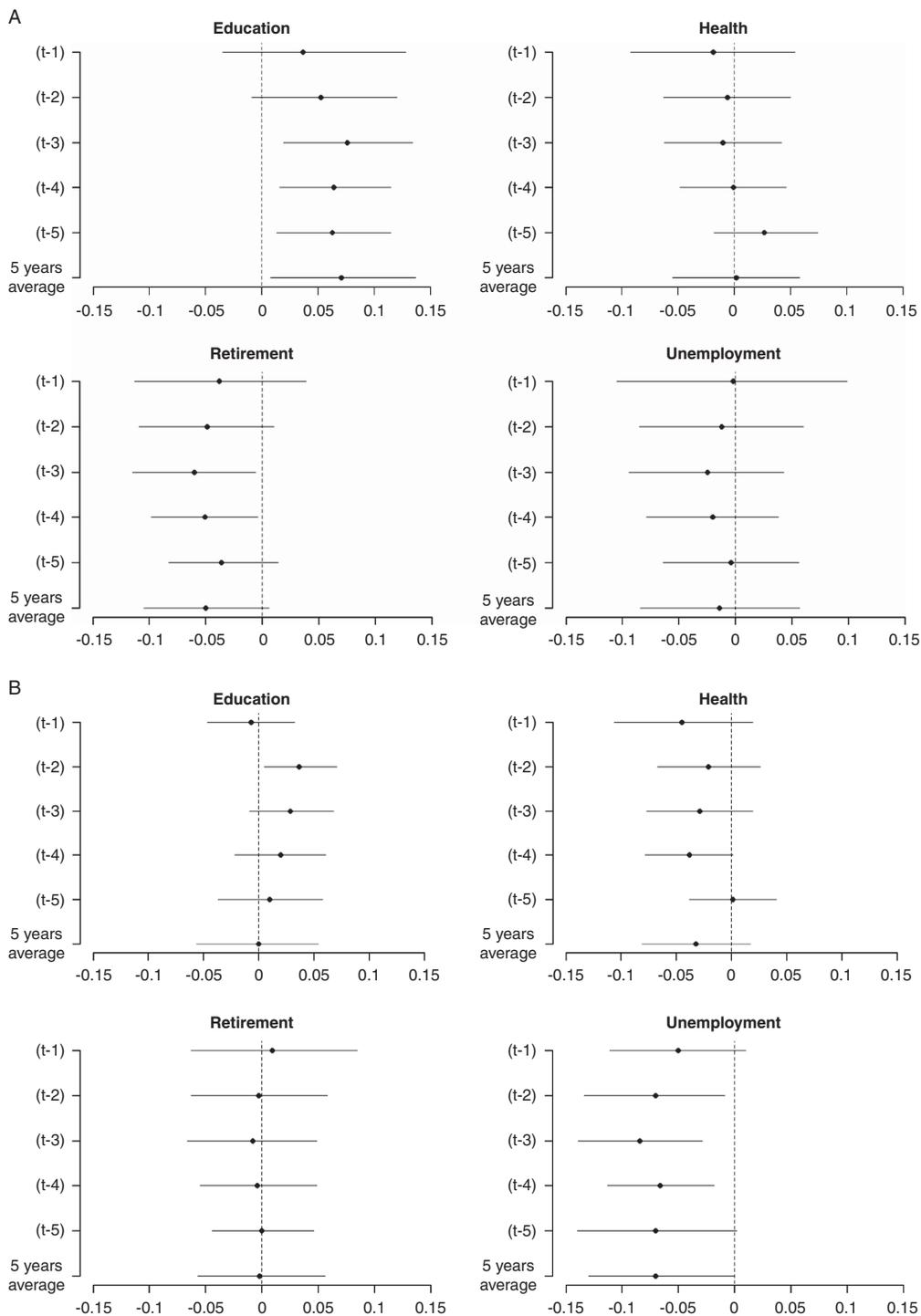
Figure 2 displays the results from the Bayesian hierarchical analysis. In general, the findings are very similar to the previous analysis. Since we are using the full set of controls in the Bayesian replication, there are slightly more cases where the credible intervals include zero. Nevertheless, the most important findings hold. Again, there is a significant positive association between past levels of social spending and support for further education spending (that is, evidence of an accelerating, general and long-term feedback effect) and a significant negative association between past levels of spending on unemployment and further support for spending increases – this could be interpreted as a specific, self-undermining and long-term feedback effect. Similar to above, the association between spending levels lagged by about three to four years and current support for spending changes seems to be more robust than short-term associations, indicating that in general long-term feedback might be more prevalent than dynamic feedback.

## Discussion and Conclusions

To reiterate, the purpose of the empirical analysis was to demonstrate the usefulness of the typology of feedback effects introduced above. The analysis clearly confirms that different kinds of feedback effects can be observed empirically. The three dimensions – direction, broadness/scope and time – are useful for describing this multitude of types. For instance, we identified an accelerating, more long-term feedback effect in the case of education, but a short-term dynamic, specific and self-undermining feedback effect in the case of unemployment. For pensions, we found evidence of a general self-undermining feedback effect, but no evidence of specific feedback. Furthermore, the analysis also revealed a number of non-significant associations, which could indicate that the impact of welfare state policies and institutions on attitudes might be trumped by the influence of other micro- and macro-level factors. However, as discussed above, policy feedback effects can also work through micro-level variables. Finally, we discussed

---

<sup>17</sup>Models in which more groups would have been excluded on the basis of their size (e.g., thirty observations) were not tested, as it would have meant dropping more than half of our sample and thereby also losing precious information on relatively rare groups – combining, for instance, low education and higher status, or high education and lower status.



**Figure 2.** Results from the Bayesian hierarchical analysis. (A) General feedback and (B) specific feedback

descriptive evidence of the extent of self-reinforcing feedback – citizens being satisfied with a particular status quo – while also noting that more sophisticated research designs are needed to properly measure this aspect.

This article does not – and cannot – provide concrete answers as to *why* different types of feedback are observed under different conditions, but this is a useful avenue for future research. Current debates are often too concerned with identifying either positive or negative feedback effects, even though this dichotomous distinction masks the significant variety of feedback effects and contributes to conceptual confusion as explained above. Furthermore, Soroka and Wlezien (2010) point to political institutions such as federalism and parliamentarism/presidentialism to explain why the feedback–responsiveness loop is stronger in some countries than in others. But this approach does not account for why different *types* of feedback should occur.

Speculating about potential conditional factors, we want to highlight two candidates. The first is the relative size and saliency of a particular policy field. If a particular policy affects large parts of the population, it is more likely that general feedback effects are more important than specific feedback. In this case, debates about highly salient policies might become intermingled with broader debates about welfare state reform, whereas reform discourses about less salient and smaller policy areas could become disconnected from these general debates. Furthermore, large-scale policy schemes such as pensions also imply a high degree of complexity in policy making, which could explain why long-term effects take precedence over short-term dynamic feedback.

The second conditional factor could be related to the relative ‘maturity’ of different areas of the welfare state. Self-undermining or self-reinforcing feedback effects seem to be more common in the traditional and well-developed sectors of the welfare state such as pensions, unemployment insurance and (partially) health care. By contrast, we observed accelerating feedback effects in the case of education, suggesting that high levels of spending are related to popular support for even more spending. Even though there is a long tradition of public provision of education in advanced post-industrial welfare states, policies to promote human capital formation have assumed a central place in recent debates about the transformation of existing welfare states towards the social investment model, which might fit better with the requirements of service-based knowledge economies (Busemeyer 2015; Hemerijck 2013). Thus, there remains significant leeway to expand this relatively new (or rediscovered) sector of the welfare state, in particular policies such as early childhood education and care, lifelong learning and higher education. Accelerating feedback effects are therefore more likely in young and less institutionalized elements of welfare state regimes than in the traditional sectors.

**Supplementary Material.** Data replication sets can be found in Harvard Dataverse at: <https://doi.org/10.7910/DVN/OEM7DW> and online appendices at: <https://doi.org/10.1017/S0007123418000534>.

**Acknowledgements.** Research for this article was supported by ERC Starting Grant No. 3117679 (Investing in Education in Europe: Attitudes, Politics and Policies). A previous version of the article was presented at the Annual Meeting of the American Political Science Association in San Francisco in 2015 and at the “Robert Schuman Centre for Advance Studies Seminar “Beyond the Democratic Deficit: Political Representation & Differential Policy Responsiveness in the Union”, 14–15 May 2015, European University Institute”. We thank the participants at this event as well as the anonymous reviewers of this journal for their helpful comments and suggestions.

## References

- Adams J and Somer-Topcu M (2009) Moderate now, win votes later: the electoral consequences of parties’ policy shifts in 25 postwar democracies. *Journal of Politics* 71 (2):678–692.
- Andreß H and Heien T (2001) Four worlds of welfare state attitudes? A comparison of Germany, Norway, and the United States. *European Sociological Review* 17 (4):337–356.
- Armington K, Isler C, Knöpfel L, Weisstanner D and Engler S (2016) *Comparative Political Dataset 1960–2014*, Berne: Institute of Political Science, University of Berne.
- Arts W and Gelissen J (2001) Welfare states, solidarity and justice principles: Does the type really matter? *Acta Sociologica* 44 (4):283–299.
- Bartle J, Dellepiane-Avellaneda S and Stimson J (2011) The moving centre: preferences for government activity in Britain, 1950–2005. *British Journal of Political Science* 41 (2):259–285.
- Baumgartner FR and Jones BD (2002) Positive and negative feedback in politics. In Baumgartner FR and Jones BD (eds), *Policy Dynamics*. Chicago, IL: University of Chicago Press, pp. 3–28.

- Blekesaune M and Quadagno J** (2003) Public attitudes toward welfare state policies a comparative analysis of 24 nations. *European Sociological Review* **19** (5):415–427.
- Bonoli G** (2013) *The Origins of Active Social Policy: Labour Market and Childcare Policies in a Comparative Perspective*. Oxford, New York: Oxford University Press.
- Brooks C and Manza J** (2006) Why do welfare states persist? *Journal of Politics* **68** (4):816–827.
- Brooks C and Manza J** (2007) *Why Welfare States Persist: The Importance of Public Opinion in Democracies*. Chicago, IL: University of Chicago Press.
- Busemeyer MR** (2013) Education funding and individual preferences for redistribution. *European Sociological Review* **29** (6):1122–1133.
- Busemeyer MR** (2015) *Skills and Inequality: The Political Economy of Education and Training Reforms in Western Welfare States*. Cambridge, New York: Cambridge University Press.
- Busemeyer MR, Abrassart A, Nezi S and Nezi R** (2019) Replication Data for: Beyond positive and negative: New perspectives on feedback effects in public opinion on the welfare state, <https://doi.org/10.7910/DVN/OEM7DW>, Harvard Dataverse, V1, UNF:6:HouWHWwg1U4PLIfqzj+2IA== [fileUNF]
- Campbell AL** (2012) Policy makes mass politics. *Annual Review of Political Science* **15**:333–351.
- Collier D, LaPorte J and Seawright J** (2012) Putting typologies to work: concept formation, measurement, and analytic rigor. *Political Research Quarterly* **65** (1):217–232.
- Erikson RS, MacKuen MB and Stimson JA** (2002) *The Macro Polity*. Cambridge: Cambridge University Press.
- Esping-Andersen G** (1990) *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- Fernández JJ and Jaime-Castillo AM** (2013) Positive or negative policy feedbacks? Explaining popular attitudes towards pragmatic pension policy reforms. *European Sociological Review* **29** (4):803–815.
- Finseraas H** (2009) Income inequality and demand for redistribution: a multilevel analysis of European public opinion. *Scandinavian Political Studies* **32** (1):94–119.
- Gerring J** (1999) What makes a concept good? A criterial framework for understanding concept formation in the social sciences. *Polity* **31** (3):357–393.
- Gingrich J** (2011) *Making Markets in the Welfare State: The Politics of Varying Market Reforms*. Cambridge, New York: Cambridge University Press.
- Gingrich J and Ansell B** (2012) Preferences in context: micro preferences, macro contexts, and the demand for social policy. *Comparative Political Studies* **45** (12):1624–1654.
- Goerres A and Prinzen K** (2012) Can we improve the measurement of attitudes towards the welfare state? A constructive critique of survey instruments with evidence from focus groups. *Social Indicators Research* **105**:515–534.
- Häusermann S** (2010) *The Politics of Welfare State Reform in Continental Europe: Modernization in Hard Times*. Cambridge, New York: Cambridge University Press.
- Hemerijck A** (2013) *Changing Welfare States*. Oxford, New York: Oxford University Press.
- Jacobs AM and Weaver KR** (2015) When policies undo themselves: self-undermining feedback as a source of policy change. *Governance* **28** (4):441–457.
- Jacoby WG** (1994) Public attitudes toward government spending. *American Journal of Political Science* **38** (2):336–361.
- Jæger MM** (2006) Welfare regimes and attitudes towards redistribution: the regime hypothesis revisited. *European Sociological Review* **22** (2):157–170.
- Jæger MM** (2009) United but divided: welfare regimes and the level and variance in public support for redistribution. *European Sociological Review* **25** (6):723–737.
- Jæger MM** (2013) The effect of macroeconomic and social conditions on the demand for redistribution: a pseudo panel approach. *Journal of European Social Policy* **23** (2):149–163.
- Jakobsen TG** (2009) Public versus private: the conditional effect of state policy and institutional trust on mass opinion. *European Sociological Review* **26** (3):307–318.
- Jordan J** (2013) Policy feedback and support for the welfare state. *Journal of European Social Policy* **23** (2):134–148.
- Kangas O** (1997) Self-interest and the common good: the impact of norms, selfishness and context in social policy opinions. *Journal of Socio-Economics* **26** (5):475–494.
- Kumlin S and Stadelmann-Steffen I** (2014a) Citizens, policy feedback, and European welfare states. In Kumlin S and Stadelmann-Steffen I (eds), *How Welfare States Shape the Democratic Public: Policy Feedback, Participation, Voting, and Attitudes*. Cheltenham: Edward Elgar Publishing, pp. 3–16.
- Kumlin S and Stadelmann-Steffen I** (2014b) How welfare states shape the democratic public: borrowing strength across research communities. In Kumlin S and Stadelmann-Steffen I (eds), *How Welfare States Shape the Democratic Public: Policy Feedback, Participation, Voting, and Attitudes*. Cheltenham: Edward Elgar Publishing, pp. 311–325.
- Margalit Y** (2013) Explaining social policy preferences: evidence from the Great Recession. *American Political Science Review* **107** (1):80–103.
- Meltzer AH and Richard SF** (1981) A rational theory of the size of government. *Journal of Political Economy* **89** (5):914–927.
- Mettler S and Soss J** (2004) The consequences of public policy for democratic citizenship: bridging policy studies and mass politics. *Perspectives on Politics* **2** (1):55–73.

- Meuleman B and Chung H** (2012) Who should care for the children? Support for government intervention in childcare. In Ervasti H, Anderson JG, Fridberg T and Ringdal K (eds), *The Future of the Welfare State: Social Policy Attitudes and Social Capital in Europe*. Cheltenham, Northampton: Edward Elgar, pp. 107–131.
- Naumann E** (2014) Raising the retirement age: retrenchment, feedback and attitudes. In Kumlin S and Stadelmann-Steffen I (eds), *How Welfare States Shape the Democratic Public: Policy Feedback, Participation, Voting, and Attitudes*. Cheltenham: Edward Elgar Publishing, pp. 223–243.
- Olivera J** (2015) Preferences for redistribution in Europe. *IZA Journal of European Labor Studies* 14 (4):1–18.
- Page BI and Shapiro RY** (1983) Effects of public opinion on policy. *American Political Science Review* 77 (1):175–190.
- Pierson P** (1993) When effect becomes cause: policy feedback and political change. *World Politics* 45 (4):595–628.
- Pierson P** (1994) *Dismantling the Welfare States: Reagan, Thatcher and the Politics of Retrenchment*. Cambridge: Cambridge University Press.
- Pierson P** (2000) Increasing returns, path dependence, and the study of politics. *American Political Science Review* 94 (2):251–267.
- Pierson P** (2004) *Politics in Time: History, Institutions, and Social Analysis*. Princeton, NJ: Princeton University Press.
- Pierson P** (2006) Public policies as institutions. In Shapiro I, Skowronek S and Galvin D (eds), *Rethinking Political Institutions: The Art of the State*. New York, London: New York University Press, pp. 114–131.
- Pierson P** (2011) *The Welfare State over the Very Long Run*. ZeS Working Paper 02/2011. Bremen: University of Bremen.
- Rehm P** (2012) Social Policy by Popular Demand. *World Politics* 63 (2):271–299.
- Roosma F, Van Oorschot W and Gelissen J** (2014) The preferred role and perceived performance of the welfare state: European welfare attitudes from a multidimensional perspective. *Social Science Research* 44:200–210.
- Rothstein B** (1998) *Just Institutions Matter: The Moral and Political Logic of the Universal Welfare State*. Cambridge: Cambridge University Press.
- Skocpol T** (1992) *Protecting Mothers and Soldiers. The Political Origins of Social Policy in the United States*. Cambridge: Cambridge University Press.
- Solt F** (2009) Standardizing the world income inequality database. *Social Science Quarterly* 90 (2):231–242.
- Soroka SN and Wlezién C** (2005) Opinion–policy dynamics: public preferences and public expenditure in the United Kingdom. *British Journal of Political Science* 35 (4):665–689.
- Soroka SN and Wlezién C** (2010) *Degrees of Democracy: Politics, Public Opinion, and Policy*. Cambridge: Cambridge University Press.
- Stegmüller D** (2013) How many countries for multilevel modeling? A comparison of frequentist and Bayesian approaches. *American Journal of Political Science* 57 (3):748–761.
- Stegmüller D** (2014) Bayesian hierarchical age-period-cohort models with time-structured effects: an application to religious voting in the US, 1972–2008. *Electoral Studies* 33:52–62.
- Stimson JA, MacKuen MB and Erikson RS** (1995) Dynamic representation. *American Political Science Review* 89 (3):543–565.
- Streeck W and Thelen K** (2005) *Beyond Continuity: Institutional Change in Advanced Political Economies*. Oxford: Oxford University Press.
- Svallfors S** (1997) Worlds of welfare and attitudes to redistribution: a comparison of eight western nations. *European Sociological Review* 13 (3):283–304.
- Svallfors S** (2003) Welfare regimes and welfare opinions: a comparison of eight Western countries. *Social Indicators Research* 64:495–520.
- Svallfors S** (2004) Class, attitudes and the welfare state: Sweden in comparative perspective. *Social Policy & Administration* 38 (2):119–138.
- Svallfors S** (2010) Policy feedback, generational replacement, and attitudes to state intervention: Eastern and Western Germany, 1990–2006. *European Political Science Review* 2 (1):119–135.
- Svallfors S** (2012) *Contested Welfare States: Welfare Attitudes in Europe and Beyond*. Palo Alto, CA: Stanford University Press.
- Teixeira R** (2017) What right-wing populism? Polls reveal that it's liberalism that's surging. Available from <https://www.vox.com/the-big-idea/2017/5/22/15672530/opinion-polls-liberal-immigration-trade-role-government-aca> (accessed 23 May 2017).
- Toikko T and Rantanen T** (2017) How does the welfare state model influence social political attitudes? An analysis of citizens' concrete and abstract attitudes toward poverty. *Journal of International and Comparative Social Policy* 33 (3): 201–224.
- Van Oorschot W** (2006) Making the difference in Europe: deservingness perceptions among citizens of European welfare states. *Journal of European Social Policy* 16 (1):23–42.
- Van Oorschot W and Meuleman B** (2014) Popular deservingness of the unemployed in the context of welfare state policies, economic conditions and cultural climate. In Kumlin S and Stadelmann-Steffen I (eds), *How Welfare States Shape the Democratic Public: Policy Feedback, Participation, Voting, and Attitudes*. Cheltenham: Edward Elgar Publishing, pp. 244–268.
- Weaver KR** (2010) Paths and forks or chutes and ladders? Negative feedbacks and policy regime change. *Journal of Public Policy* 30 (2):137–162.

- Wlezien C** (1995) The public as thermostat: dynamics of preferences for spending. *American Journal of Political Science* **39** (4):981–1000.
- Wlezien C** (2004) Patterns of representation: dynamics of public preferences and policy. *Journal of Politics* **66** (1):1–24.
- Wlezien C** (2016) Public opinion and policy representation: on conceptualization, measurement, and interpretation. *Policy Studies Journal* **45** (4):561–582.
- Wlezien C and Soroka SN** (2012) Political institutions and the opinion–policy link. *West European Politics* **35** (6): 1407–1432.