What counts? Survey designs to measure citizens' social policy priorities comparatively

Silja Häusermann, Macarena Ares, Matthias Enggist, Michael Pinggera,
all University of Zurich

Paper prepared for the Fifth International Meeting on Experimental and Behavioral Social Sciences, IMEBESS 2018
Florence, May 3-5 2018
Panel: Attitudes toward Policy

Abstract:
While most comparative surveys provide data on citizens' positions regarding social policies, these questions are increasingly unable to validly measure both key concepts of the main theories to be tested, as well as real-life opinions citizens form in the context of welfare politics. The problems of validity are mainly linked to the fact that in today's context of constrained public budgets, there is a need to measure citizens' *priorities* in the field of social policy and not only their positions. Contrary to positions, priorities reflect the relative weight citizens attach to different social policy fields, instruments or reform proposals.
In the context of the ERC project "welfarepriorities", we explore and test a variety of experimental and observational ways to measures and explain such priorities comparatively across 8 European countries. This paper presents the rationale of the project and the different ways in which we approach the measurement of priorities in a comparative public opinion survey as well as country-specific surveys, which will all be fielded in the course of 2018. We present and discuss the benefits and costs of various measurement strategies: conjoint experiments, rating questions, direct trade-off questions, as well as country-specific vignette experiments regarding the determinants of priorities. Moreover, a key challenge we address in the paper is how to fruitfully use experimental survey designs in comparative research.

1. Introduction

In this paper, we propose strategies to measures citizens' priorities between different social policy fields and social policy reforms. The goal is to discuss and evaluate the costs and benefits of a range of observational and experimental design strategies for measurement. While the paper substantively focuses on welfare state research, its questions and claims are more general in nature. How to best measure relative importance and priorities regarding political choices? How to combine findings on priorities – from experimental (conjoint) and observational designs – with measures of position to get information on weighted positions? How to test causal determinants of priorities? And how to implement (experimental) surveys for comparative purposes?

Welfare state research is a field in which these questions have become extremely important, because individual-level research on preferences has gained so much theoretical and empirical prominence over the past two decades (e.g. Iversen and Soskice 2001; Brooks and Manza 2008, Blekesaune and Quadagno 2003; Rehm 2009; Svallfors 2012). At the same time, both the key theories and the real-life choices citizens and politicians make in this field are poorly reflected in the available survey indicators of social policy preferences that are so widely used in the field. Hence, we argue that the field faces serious problems in terms of concept validity – the mismatch between theory and concept – and measurement validity – the mismatch between the concept and its indicator.

Therefore, one ambition of the ERC project "welfarepriorities" is to improve concept and measurement validity of social policy preferences by focusing on *priorities*, i.e. the relative importance citizens attach to different types of welfare state reform. The overarching goal of the project is comparative, i.e. to test if such new data allows for a better understanding of why welfare states develop differently across countries.

This paper is structured as follows: the next section provides first the motivation by discussing the key problems of existing social policy preference data: concept validity on the one hand and measurement validity on the other hand. It then discusses the recent efforts that have been made at measuring priorities and the remaining problems we identify in these newer studies. A third section outlines the various strategies we propose to overcome these problems with both observational and experimental approaches. It addresses both the challenge of measuring priorities and of identifying the explanatory factors of these priorities. The final section gives an overview of the planned surveys in terms of sampling, design and planned fieldwork.

2. Theoretical motivation, objectives and state of research

2.1. Validity problems in existing research

The daunting challenges – such as demographic changes, labor market failures the recent economic crisis – that threaten the political balance between social solidarity and sound public finances, and the emergence of new social and political divides that may undermine and reshape existing social policies are among the main stakes that motivate comparative welfare state research. The goal in this research field is to understand the scope and impact of political agency in a tight network of structural and institutional constraints: whether and how welfare states and welfare state reforms differ between countries.

Since roughly the 1990s, research on whether and how welfare states are reformed has been theorized and studied in a different intellectual framework than before. While welfare politics previously took place in a context of expansion, the structural and political context turned to fiscal constraints and austerity since then (Pierson 1996, 2001). This changing context, in combination with methodological trends and data availability, have put the study of individual social policy preferences and public opinion as determinants of welfare state change to the forefront, because public opinion is considered the key factor constraining both a retrenchment of existing welfare state benefits, as well as, more generally, the recalibration of the welfare state. Therefore, understanding social policy preferences has become the key endeavor in this research field and a massive and still growing literature has emerged that studies the determinants of individual attitudes regarding social spending, redistribution, and various social policy instruments (e.g. Iversen and Soskice 2001; Estevez-Abe et al. 2001; Brooks and Manza 2008, Blekesaune and Quadagno 2003; Rueda 2005; Rehm 2009, 2011; Walter 2010; Rehm et al. 2012; Svallfors 2012; Margalit 2013; Häusermann et al. 2014, 2016; Gallego and Marx 2015; Fossati and Häusermann 2014, Beramendi and Rehm 2016; Busemeyer et al. 2017, Dimick et al. 2017).

These studies are oftentimes valid and insightful when it comes to understanding individual-level preference formation as such. However, either explicitly or implicitly, studying social policy preferences is a means rather than an end in itself. Ultimately, the goal is to understand politics and policy development at the aggregate level, and to explain differences between governments, periods and countries. To illustrate this point, think of the recently emerged literature on how multicultural diversity and immigration affect redistribution preferences (e.g. Finseraas 2008, Huber 2017; Rueda forthcoming). Of course, there is a genuinely social-psychological interest to this research. However, the main question this research speaks to is whether the politicization of immigration in recent times is likely to undermine the existing welfare states. Other studies make this link explicitly, by relating patterns of public opinion to policies and reforms (e.g. Brooks and Manza 2008, Häusermann 2010, Rehm 2016).

Hence, the contribution of political science research on social policy attitudes ultimately lies in its capacity to help us understand variance in politics and policy. From this perspective, we see *two main problems* in the validity of the available cross-sectional survey data on spending, redistribution or generosity preferences that most findings rely on: first, these questions on policy positions increasingly *fail to capture the choices that politicians and citizens face*. While politicians and citizens have to evaluate reform proposals destined at restricting social policies at the benefit of either other policies or sounder public finances, surveys are usually inquiring about expansion preferences and in an unconstrained way (see Häusermann et al. 2017). The second problem is more theoretical. Upon close reading, the most prominent recent theories of welfare state development – e.g. the new politics of the welfare state (Pierson 2001, Jensen 2012), varieties of capitalism/risk-theories (Hall and Soskice 2001, Rehm 2016) and dualization theories (Rueda 2005) – rely on *micro-foundations that relate to the relative importance individuals attribute to different policies*, rather than their absolute positions on these policies. Yet, the empirical studies tend to use position-measures to test the theories. This mismatch creates obvious problems of validity.

In the following, we briefly review the two issues sequentially. The first problem – surveys not reflecting the political choices salient in politics today – may be illustrated most easily with an example. Over the years of the Great Recession, huge budgetary pressure emerged regarding distributive politics, especially in the countries most affected by the crisis. These pressures only intensified the structural constraints on welfare state spending that had become increasingly salient since the 1990s, such as demographic ageing or structural unemployment. In such a context, welfare politics increasingly resembles a negative- or zero-sum game. Hence, the main question governments face is whether to cut back on social spending and on what areas. These are the reforms that are debated in politics and in election campaigns, i.e. they are relevant for voters and supposedly for their choices. However, in existing surveys, citizens are usually asked if they would prefer more or less spending, more or less generous policies. In the answers to these questions, variance with regard to people's is generally low and attitudes may even be contradictory (Goerres and Prinzen 2012; Giger and Nelson 2013, Busemeyer 2014), because most people support generous social spending for redistribution, pensions, education, unemployment etc. and they also support sound public finances and low levels of taxation. Being unconstrained in their answers, respondents seem to favor all the desirable outcomes equally. Consequently, the findings then usually look similar to the ones in figure 1, displaying average support levels for increased ("more" and "much more") spending on pensions and education and for opposition to retrenchment ("less" and "much less" spending), as measured in the ISSP 2016. The numbers in figure 1 are informative and probably valid, in the sense that most citizens across all countries would indeed prefer increased spending in both key areas of welfare policy and virtually all citizens reject retrenchment. However, these numbers do not reflect the policy decisions at stake in these countries and it is highly doubtful that citizens indeed expect their parties and governments to spend "much more" when the tone of the political debate for over two decades has centered on the need for fiscal consolidation.

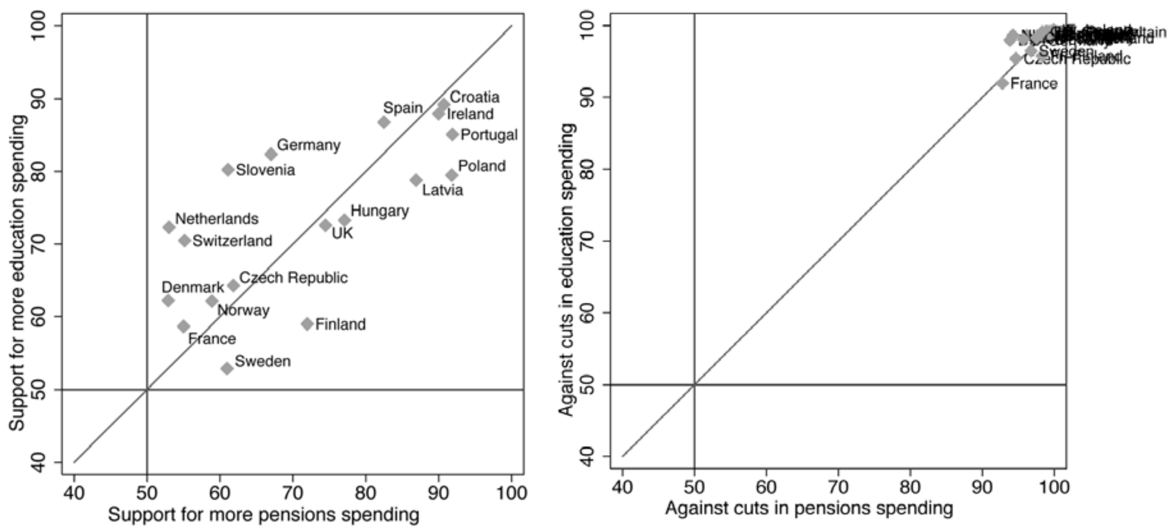

Figure 1: average levels of support for social spending in % (ISSP 2016)

It is obviously doubtful that these data are able to explain cross-national variation in policy development – but not because public opinion does not matter. Rather, the data used is unable to carry more information than a general "mood" regarding welfare state reform. What we would want to know is how citizens evaluate the ways in which politicians prioritize different benefits, how they handle the fiscal pressure, and how citizens themselves prioritize different areas of social spending.

The second problem relates to the mismatch between theories and measurement. We illustrate it with reference to three of the most prominent welfare state theories of the past two decades. First, theories on welfare state retrenchment. In the wake of Pierson's groundbreaking studies on the ways in which existing institutional settings bias democratic politics against welfare state retrenchment, innumerable studies have investigated the extent to which governments have or have not cut back on welfare generosity. One key claim that has found resonance is that the stability of benefit levels is inversely related to the concentration of the risk the benefit addresses. In other words: if unemployment is highly concentrated among the low-skilled, public opinion (or the median voter) will be more permissive towards retrenchment than if unemployment spreads widely across occupations, classes and generations (Jensen 2012, Rehm et al. 2012, Rehm 2016, Pontusson and Weisstanner 2017). The dispersion of risk varies across time and space, but also across policy fields, which is why old age pensions are supposed to be more stable than disability pensions. However, the point we want to make here is that the theoretical argument is based on the *relative importance* citizens attribute to particular social benefits. The more likely they are to incur the risk, the more intensely they are supposed to defend existing levels of generosity in social schemes that address precisely that risk. Hence, while general attitudes on welfare generosity and redistribution are most strongly linked to political ideology, individual interest is supposed to explain how much relative importance individuals attribute to particular social policies (within a general bandwidth of social spending). This is why Jensen (2012) argues an interaction effect of left-right partisanship and social policy attitudes: left-wing citizens favor a higher equilibrium of taxation and spending, but *within* this equilibrium, they are *more attached* to those policies they benefit from directly, and the same goes for the right. Hence, while we do find a link between risk incidence and position, this argument would be more validly tested with an indicator of relative importance and the empirical evidence should be much stronger. Ideally, we would be able to weight position with indicators of importance.

The second, related example is the hugely influential re-thinking of welfare politics that the theory on "Varieties of Capitalism" (Hall and Soskice 2001) entailed: contrary to ideology- and class-based arguments, this approach suggested that the risk-profile of citizens and firms in terms of skills predicts their social policy preferences. Firms with high risk incidence prefer large risk-pools (to which other firms with lower risk incidence also contribute) to smaller risk-pools; employees with general skills and discontinuous labor market biographies have a stronger interest in universal, redistributive social policy, while employees with specific skills who remain in long, stable employment relationships fare best with a (sector-based) social insurance regime. From this contention a long and controversial academic debate emerged: do employers really support large and generous risk pools? Do workers with general skills really reject social insurance? (e.g. Mares 2003, Korpi 2006, Paster 2013). Introducing relative importance and priorities to the argument largely unveils this debate as mistaken: of course, firms *generally* favor lower levels of taxation and spending than employees. But if in a particular reform they are put before the choice between larger and smaller risk pools, they *prioritize* one design over the other. And any employee is better off with a generous rather than a small welfare state. But a generally skilled employee would arguably prefer redistribution over insurance. Survey questions that inquire about the level of spending or generosity of a social policy are unable to provide a valid test of the theoretical arguments made.

Insider-outsider theories of the labor market provide a final telling example. Rueda (2005, 2007) made waves about a decade ago with his research that pointed out the dilemmas left-wing parties face in the recalibration of labor market policies. While labor market insiders' main interest consists in employment protection and generous social insurance, labor market outsiders have an interest in low "barriers to entry" into the labor market, active labor market policies and redistribution. While Rueda (2005) indeed showed significant differences in the average positions of insiders and outsiders regarding employment security policies and activation, many subsequent studies pointed out the very small substantive differences in these positions, and they countered the initial argument with evidence showing that labor market outsiders *also* support employment protection (e.g. Emmenegger 2009). Again, a close reading of the theoretical argument reveals that data on *positional* attitudes regarding social policies is not valid to accurately test it. Rather, given their employment profile, we would expect labor market insiders to *prioritize* employment protection over redistributive compensatory and activating policies and vice versa for outsiders. Dualization theory does not claim that labor market insiders would actively fight redistribution or that outsiders would mobilize to demand lower employment protection regulation. The divide that one would theoretically expect is one over priorities, not over position. Several studies have tried to operationalize such attributed relative importance indirectly, for instance by observing on which policy reforms collective actors did speak up or remain silent (Palier and Thelen 2010), but to date, most of the research on labor market dualization and preferences has approximated these expected preferences with position indicators that may be unable to bring the debate forward.

The two problems discussed in this section – the failure of position data to reflect the actual politics of welfare state reform and the mismatch between many theoretical claims and the data used to test them – motivate the objective of our project and this paper to explore different ways of measuring citizens' priorities regarding welfare state policies. The focus of this paper is therefore on measurement of relative importance as a dependent variable. Beyond this, we briefly discuss how measures of relative importance might be used as independent variables or as weights for positional indicators in order to explain welfare state reforms. Moreover, an analytical shift of attention from position to priorities also raises the question about determinants of priorities and how to test them. In our project, we are interested both in individual-level determinants of priorities as well as in contextual determinants. Therefore, in section 3 we also discuss selective ideas for analytical strategies and survey designs to test these explanatory factors. In the next section, however, we start by reviewing recent studies that we build on and how our objectives and strategies relate to them.

2.2. Recent improvements in measurement and how our own project relates to them

Recent research has made considerable advancements in the measurement of policy preferences—both regarding priorities and position—in line with the developments in the literature on welfare policies. These improvements have been made on two directions. First, new survey questions ask directly about citizens' preferences for *specific social policies* that are far more concrete than general preferences for state intervention or aggregate levels of redistribution. Second, and more closely linked to the key ambitions of our project, recent developments have also included survey items that explicitly capture *citizens' preferences in terms of relative priorities when faced with a trade-off or constraint*. Through different means respondents are made aware that they (or someone else) have to bear the cost of expanding state intervention. In our own surveys, we plan to build on these two advancements by presenting respondents with the choice between several specific forms of social policy reform profiles. This will be done both in the form of conjoint experiments, as well as via more direct survey items. The advantage of these choice-designs is that reform choices involve a cost or at least opportunity costs for respondents.

One strategy to address the lack of constraints in traditional survey items has been to include a reference to the (implicit) cost of policy intervention in the formulation of the question. Frequently, this has been done by including a reference to budgetary constraints and indicating that expansion in a specific policy will come at the cost of higher taxation, a raise in public debt or cutbacks in public expenditure in other areas (see e.g. Hansen, 1998). In social policy, this kind of trade-off has not only been formulated between different policy domains (e.g. expansion of education at the cost of pensions), but also between different forms of intervention or between specific programs within a broader domain (see e.g. Busemeyer & Garritzmann, 2017; or Fernández & Jaime-Castillo, 2013).

Although this approach represents a considerable improvement from unconstrained measures of policy preferences, it still presents some limitations, especially if our interest lies in studying conflicts around welfare state priorities within the electorate. First, including a reference to the fiscal cost of a specific policy ("higher taxes") usually does not convey any information about who exactly will be the main bearer(s) of this cost. Although these items present an explicit trade-off, they do not hint at whose interests might be at stake. This drawback is partially accounted for when the choice is between spending in different policy domains (e.g. childcare vs. pensions), where respondents can more easily identify the constituencies benefitting from different programs. However, these kind of survey items do not convey information about (continuously quantified) relative importance of different domains and forms of intervention. This is precisely why we set out to measure policy priorities. This kind of fine-grained information can be better obtained by the implementation of items that ask respondents to *simultaneously evaluate and choose between welfare profiles or policies that vary on different dimensions*. We propose to do this via conjoint experiments, as well as through items that ask respondents to allocate points to different options of policy reform.

Recently, conjoint studies have shown great promise in assessing how traits of social policies or spending profiles relate to support for these policies. Conjoint analyses have a long tradition in marketing research but their implementation in political science studies is rather recent. The advantage of conjoint designs is that they allow us to evaluate how different traits of a specific object contribute to the likelihood that a respondent will choose that object. Many of the characteristics of this design appear to match some of the necessary improvements to the measurement of welfare preferences. First, conjoint profiles *differ on a number of attributes*, which allows us to account for the multidimensionality of welfare policies. Second, respondents are asked to *choose* between different profiles, which introduces a constraint.

In the realm of social policy and public opinion, conjoint analyses have so far been implemented to study preferences for different public spending profiles (e.g. Bremer & Bürgisser, n.d.; Kölln & Wlezien, 2016) or for different specific reform profiles within a specific policy domain (Gallego & Marx, 2017 regarding unemployment policy; Häusermann et al. 2017 regarding pension policy). We do not find, however, conjoint analyses that present welfare state profiles that vary in the specific policy actions (e.g. targeted or universal benefits) implemented on different domains (e.g. pensions, health, education). This is, precisely, what we propose in our set of conjoint experiments.

Our main *objective* with this design is to capture distributive choices that welfare reforms entail, as well as the priorities citizens have when it comes to these choices. By comparing the priorities of different sub-groups of the population, we want to get at the *political divides* over these priorities. Hence, we e.g. want to know if party groups or socio-demographic groups differ in their priorities about how scarce resources should be used, or about who should be affected by cuts and who should be sheltered. If they differ significantly in these priorities, we speak of a divide. This is a different conceptualization of welfare reform conflict: whereas in the past, we thought of conflict as opposing groups that favor and groups that reject a certain policy (e.g for and against generous old age pensions), we think that divides in terms of priorities capture political conflict around the welfare state in the 21st century more adequately: everybody may deem generous old age pensions desirable, but the question is whether people think generous old age pensions are so important that they are willing to let go of other (also desirable) social policy benefits for the sake of these pensions.

The main distributive choices we focus on in our project are class (whether a reform should be more or less progressive), multiculturalism (whether a reform should treat migrants more, equally or less generously than citizens), dualization (whether a reform should prioritize social protection or activation), and the choice between social investment and social consumption (whether a reform should rather invest in employability or replace income). If different groups value the alternatives differently, we speak of a political divide. For illustrative examples, think of different age cohorts differing in the relative importance they attribute to investment and consumption; of insiders and outsiders differing in their priorities regarding insurance and activation or different party constituencies differing in their evaluation of the choices between universalism and welfare chauvinism.

Thus, we need to develop measures that capture divisions within the citizenry in terms of their priorities and, as we mentioned above, these priorities encompass both the notion of constraint and of multidimensionality. These specificities of our research agenda separate our design from some previous proposals. Some recent designs (Bremer & Bürgisser, n.d.; Kölln & Wlezien, 2016) have varied levels of spending across different policy areas (e.g. increase/reduce spending on education, healthcare, defense, child benefits by 2, 5 or 10 percent). Hence, respondents are asked to choose among different profiles that differ in terms of the fields in which they observe an increase or decrease of spending. Our strategy proposes not to vary spending levels, but rather to present expansive and retrenching measures that are more explicit in terms of the action that is taken, in order for respondents to be able to understand the distributive effects of the proposed reform. To illustrate this with an example, instead of proposing that spending on pensions will increase by an amount of *x*, we would specify that the government will increase minimum pensions. Both measures indicate expansion in the provision of pensions, but there are two reasons that justify our approach.

First, in specifying attribute levels that differ from one another qualitatively, we are better able to reflect the complex multidimensionality of social policy. Changes in spending levels will have qualitatively different distributional implications depending on how they are implemented. Trade-offs do not only take place across policy areas—i.e. spending on pensions instead of education—but also *within* specific policy fields – i.e. increasing minimum pensions significantly vs. increasing all pensions slightly. Being more specific about the *constituencies* that benefit from a specific policy can provide more nuance to our measures of citizens' preferences. Recent studies have found that general increases in expenditure on pensions or education appear as highly popular (Bremer & Bürgisser, n.d.). Moreover, these increases frequently trump rises in spending in other areas. This support for education and pensions may not be as broad when policies in these areas are targeted to specific groups. Providing respondents with information about *who mainly benefits from policy reforms* bring us closer to the truly multidimensional nature of welfare policies.

Second, including this more detailed information about the specific nature of welfare expansion on retrenchment provides a better fit with the objective of identifying the *divides* that characterize conflict over welfare politics today. Raising pension spending by a certain percentage does not tell the respondent much about the distributional effect of this reform. Conversely, our conjoints provide more information about which groups are more likely to benefit from it. This strategy should better capture constituencies that differ in the social policies they prioritize. As explained in further detailed below, we design the conjoints based on the potential sociodemographic divides we want to identify. We specify directions of policy change that affect particularly certain groups, which should display greater disparities in priorities since they face different kinds of risks. Our proposed conjoints can be considered as a middle ground between previous examples that have assessed preferences for levels of spending in different policy areas (e.g. Bremer & Bürgisser, n.d.; Kölln & Wlezien, 2016) and those that have focused on very specific individual reforms within an area (see e.g. Franchino & Segatti, 2017, for varying features of economic policy and European fiscal union; Gallego & Marx, 2017, for labor market reforms; Häusermann et al. 2017 for pension reforms).

The main advantage of implementing this conjoint design is that it allows us to account jointly for the multidimensionality and the inclusion of constraints. This strategy also allows us to measure priorities on a large set of different policies since, because profiles are randomly generated, respondents do not need to evaluate every possible combination of attributes and levels (Hainmueller, Hopkins, & Yamamoto, 2013). The conjoint profiles, which are presented as a package, also neatly capture the notion of *welfare reform profiles*, i.e. of potential directions of welfare state reform.

In spite of these advantages, the conjoints also present certain drawbacks which lead us to include other items measuring priorities in our survey. As we discuss in further detail in the next section, one of the drawbacks of conjoint designs is that – since, by definition, respondents do not evaluate all attribute-level combinations – the coefficients they produce are aggregates that apply only to a group (full sample or subgroup) of respondents, but not to individual respondents themselves. Hence, conjoint designs deliver findings that are not easily used as explanatory factors for other variables (think, for example, of testing the relationship between policy priorities and electoral choice). Many previous analyses have resorted to analyzing the conjoint results via sub-samples of and interactions with meaningful explanatory variables that ground differences in preferences (with and without controls). This approach, however, limits the possibilities of the analyses that can be carried out, especially if we seek to formulate more complex models where priorities are expected to vary as a consequence of the interactions between individual-level covariates, as is frequent in comparative political economy. The number of parameters estimated in a conjoint analysis also increases complexity in comparative designs like ours (eight different countries). For all these reasons, conjoint designs appear as a very valid tool to estimate the relative contribution of specific reform elements on overall welfare reform support, but not the panacea for explanatorily oriented theoretical research questions. Therefore, we propose additional measures to capture policy priorities on specific reforms and areas of intervention: respondents are also asked to *allocate a limited number of points* to different kinds of reforms, and also to position themselves on *direct trade-off questions* where expansions in one policy come at the cost of retrenchment in another. These survey items are poorer in terms of dimensionality and causal identification, but they deliver actual micro-level data specific to each respondent that can be used in further research questions.

3. Different ways of measuring priorities

3.1. Measurement of priorities, DV

3.1.1. Conjoint experiments

In our survey we ask respondents to evaluate different proposals for changes to social policies through different items. We start by presenting them with two types of sets of conjoint profiles. Each conjoint profile corresponds to an entire welfare reform strategy that represents a direction or profile of development for the welfare state. Each of these profiles is constituted by specific policy reforms (the conjoint levels) to be implemented in different policy fields (the conjoint attributes). Each comparison asks respondents to choose the most preferred of the two profiles they are presented, and then to rate their support for both profiles (on a 0 to 10 scale). Respondents are presented separately with two kinds of conjoints: (i) one kind asks respondents to choose between two scenarios where policy change is in the direction of expansion of benefits and services, (ii) the other kind faces respondents with a choice between two proposals that refer to a retrenchment of welfare policies. Both retrenchment and expansion profiles vary on the same six attributes (the areas of social policy intervention): old-age pensions, childcare, tertiary education, income replacement for the unemployed, employment reintegration services for the unemployed, and social services. Each attribute varies on three to four levels, with each level representing a specific change in the corresponding policy field. Because we do not want to present respondents with profiles that entail across-the-board expansion or retrenchment, each attribute includes one level that maintains social policy at its current level (status quo). Hence, in the expansion scenario benefits in each policy field can be either expanded or kept at the status quo, while in the retrenchment scenario they can be either reduced or maintained. The profiles are generated randomly but we constrain them to both show an equal number of policy retrenchments (or expansions, depending on the conjoint) so that respondents always have to choose between options that are balanced in the extent to which each of them expands (or retrenches) the welfare state.

There are different reasons for why we separate the retrenchment and expansion scenarios. First, there are theoretical grounds to believe that priorities for expansion and retrenchment are different. Priorities regarding where to increase levels of intervention and provision are not necessarily symmetric to those concerning where to avoid retrenchment. Endowment effects make people value a good or a service more if they already possess it and this is likely to ground an asymmetry in priorities for expansion and retrenchment (Chapman, Dean, Ortoleva, Snowberg, & Camerer, 2017). The different nature of these two choices might also be mirrored in the kind of explanatory factors behind these choices. Measuring priorities for welfare state expansion and retrenchment separately will allow us to study potential differences in these two decisional contexts.

Other than the different nature of priorities regarding gains or losses, there are also other more practical reasons to keep expansion and retrenchment separate. An analysis of recent developments in welfare policies (from 2010 onwards) in the eight countries under study reveals substantive differences in the direction that social policy reforms have taken across different contexts. While some countries have gone, almost exclusively, through a period of stark retrenchment (mainly Italy and Spain), others (e.g. Germany or Sweden) have experienced mostly expansion—even if more limited in the size of the reforms. Thus, presenting a combination of expansion and retrenchment would be quite unrealistic in several of the countries considered. Moreover, these same analyses also indicate that debates around recent social policy reforms do not seem to explicitly refer to recalibration (or retrenchment in one are being compensated by expansion in another or vice versa). Keeping the two scenarios separately means that, in most countries, at least one of the proposed situations will be realistic for respondents. To avoid any order-effects, we randomize whether respondents reply first to the retrenching or the expansive conjoint.

Tables 1 and 2 present, respectively, the attributes and levels of our expansion and retrenchment profiles. This design responds to our two objectives of capturing citizens' priorities and the divides structuring conflict around them. Having balanced comparisons in the conjoints—i.e. an equal number of expansions (or retrenchments) in the pairs compared—means that respondents are faced with a constrained decision. At the same time, the combination of attributes and levels accounts for the multidimensionality of social policy. Because welfare policies can vary on many dimensions, we need to follow certain criteria for selecting the relevant attributes and specifying the levels within them. We do so by defining attribute-level combinations that correspond to our objective of capturing divides among the citizenry. At the same time, we seek to find a balance between the cross-country comparability and the realism of the levels presented to respondents.

Tables 1 and 2 here: Attributes and levels of the conjoint designs

The attributes presented vary on three to four levels. One level always refers to the status quo. Then, each attribute includes one level in which expansion (or retrenchment) is universal. In other words, the specific reform is defined so that it will affect all (potential) beneficiaries in a specific policy field. The third and fourth level (where there is one) introduces a reform that is targeted, i.e. that affects particularly certain groups of the population. The decision as to which groups are benefitted (or harmed) by the specific policy action is based on the divides that will likely articulate social policy conflict.

To illustrate this with an example, the first attribute presented in table 1 corresponds to old-age pensions. For this attribute the level for the status quo keeps pension benefits unchanged. The level referring to an action with universal implications proposes to support complementary pension savings for everyone, thus all respondents will potentially benefit from this reform in the area of pensions. The other two levels are group-targeted. One level attempts to capture class divisions and specifies that minimum pensions will be raised. The other level attempts to get at the generational divide by specifying that additional savings will only benefit the younger cohorts. This level excludes older citizens from the pool of beneficiaries. We apply this same logic across all attributes. Through the targeted reforms we seek to get at divides in priorities based on social class, age, citizenship, and labor market status, and we specify levels of attributes accordingly. We attempt to measure these divisions in the policy fields where they are more likely to emerge (e.g. employment integration policies benefitting labor market outsiders, means-tested policies favoring low-income groups in tertiary education). Purposely matching the divides to certain fields of social policy where they are more pertinent also means that the levels are more realistic. In fact, striking the balance between realism and generalizability across different national contexts is one of the main challenges in this measurement strategy. Designing levels that refer to specific reforms conveys more information about who is to benefit (or be harmed) from them. At the same time, the wording of the levels needs to be general enough so that it is reasonable for respondents living in different institutional contexts. We combine this specific measures in the conjoint with more generic items in the survey, which capture, generally, priorities regarding fields of intervention (pensions, unemployment…). While the latter are more comparable, but they do not provide any information regarding support for specific policies.

As we mentioned above, conjoints present certain limitations as to the type of individual-level analyses that can be implemented. Two of the main objectives in this project are to study the impact of individual socio-economic factors, and to compare priorities and divides across countries. At the individual level we are limited in that it is not simple to study the correlates of priorities measured via conjoints. Moreover, the relatively large number of attribute-level combinations in the conjoint make the comparison of results across countries difficult, since we will have 15 coefficients estimated for each type of conjoint (expansion or retrenchment) in each country. For these different reasons we developed additional survey items that also seek to capture citizens' priorities. Clearly, these alternative measures are limited in the complexity of the reform profiles they can present to respondents. The advantage is that replies to these other items are available for each respondent, thus simplifying the individual-level analyses that can be estimated on these data.

Before being presented with the conjoint profiles, respondents are asked to position themselves on a four-point agree/disagree scale regarding items that replicate some of the attribute-level combinations included in the conjoints. Half of these items refer to expansions and the other half to retrenchments. In this case, replies do not reflect priorities but positions since there are no trade-offs implicit in these questions. Respondents are unconstrained to support expansion or retrenchment of public provision in all of the items. We include these positional items for the purpose of designing composite measures that weight positional items by the information on priorities stemming from the conjoints. The logic behind this strategy is to arrive to a measure that takes into account the position that a respondent takes on an issue as well as to what extent this is a priority. For instance, an individual may strongly favor improving the quality of public childcare but this is not a priority for him (which would be evidenced in a low impact of this item on the probability of choosing a conjoint profile). In our project we attempt to find a strategy for combining the information from these two kinds of questions to get to priority-weighted measures of position. There are certain challenges in pursuing this strategy: (i) the AMCEs in the conjoints are estimates and they are always relative to a reference category hence it is not straightforward how to translate this information into weights, (ii) the estimates of the AMCEs already contain some information on position since the attribute levels contain specific information about the development of a policy in a certain direction, and (iii) the weights would have to be computed for the entire or sub-groups of the sample but not for individual respondents.

3.1.2 Point distribution questions

As an alternative strategy for measuring social policy priorities we ask respondents to allocate a limited amount of points (100) to different survey items (Table 3). Through this type of question, we are able to introduce both the multidimensionality and the constraint relevant for measuring priorities. Because distributing points is a relatively demanding task for respondents, we reduce the amount of items that are presented to them. We rely on a selection that mirrors some of the attribute-level combinations designed for the conjoint. As before, we separate expansion and retrenchment. This approach for capturing priorities will be replicated in a survey to be fielded among political elites (members of parliament of the different countries under study). Thus, these scoring questions allow us to gather comparable data at the individual and the elite level.

Table 3 here: Rating questions

To simplify the task that respondents need to complete but still measure priorities on different issues we implement two kinds of questions. First we ask respondents to allocate points to statements similar to those included in the conjoint. These items present social policy reforms that affect specific groups (e.g. increase minimum pensions). Individuals have to allocate 100 points to six such items. To simplify future analyses of these items, we restrict the dimensions on which the statements vary. These items refer to the generational and class divides, as well as consumption vs. social investment policies. The task involves distributing points instead of ranking because points allow respondents to attribute the same level of priority to more than one item, and to express the intensity of their priorities. To have additional information about priorities on other areas of welfare state intervention, in a second step, we ask respondents to rate social policy fields (e.g. old age pension, higher education…) without providing information on the nature of reforms in these fields. Through these different items we are able to gather information that is available for each respondent in the sample, that is comparable across countries, and that is replicated for political elites in each country.

3.1.3 Direct trade-off questions

Our last strategy for measuring citizens' welfare priorities relies on direct trade-off items. In these questions respondents are asked to express their opinion on a four-point agree-disagree scale regarding situations in which expansion in a specific policy comes with the direct and explicit cost of retrenchment in another area (the exact wording is included in table 4 below). To illustrate this with an example, individuals are asked to state to what extent they agree with the statement that the state should guarantee stable old age pensions for future generations, even if this comes at the cost of lower benefits for the currently retired. Respondents face a four-category response scale. In contrast to the items presented before, these questions introduce expansion and retrenchment simultaneously, and directly related to each other. We developed five different items that attempt to capture the social divides we are interested in. Responses to these questions should be guided by the gains attributed to the expansion and the costs attributed to the retrenchment included in each item. These items, alone, do not provide information about which of the two judgements have a greater weight in guiding this decision.

However, by combining the information from these questions with previous items that did separate the expansion and retrenchment scenarios we will be able to estimate to what extent aversion to a retrenching reform or preference for expansion guide responses to the trade-off. Moreover, having one trade-off capturing each divide and fielding identical items across all countries can also facilitate cross-country comparisons.

Table 4: Policy trade-off questions

Overall, the different strategies implemented in this survey (each with their advantages and limitations) will allow us to overcome some of the challenges in measuring and studying priorities over social policies. Moreover, the combination of the information gathered through the different items will provide more nuances about respondents' preferences when faced with constraints and with the multidimensionality of these policies.

3.2. Explaining priorities

Our project has the ambition both to *measure* priorities, as well as to *explain* priorities across individuals and across contexts and possibly also to *use priorities* as weights or independent variables to account for political preferences and behavior. This section therefore discusses the analytical strategies that we plan to implement in the explanatory parts of the project.

3.2.1. Individual-level determinants of priorities

Our survey will contain 30-40 items measuring attitudes (a.o. electoral choice and political participation; social policy positions; attitudes on state intervention, cultural liberalism, EU integration and migration; left-right self-positioning, evaluation of the welfare state, association membership), as well as a battery of socio-demographic items (a.o. household situation; employment situation; education and income; occupation; intra-familial transfers).

The use of the individual-level variables as independent variables is rather straightforward and does not require too much elaboration. For the rating- and the trade-off-questions in particular we will have individual-level responses that can be used directly or recoded into individual-level dependent variables (i.e. the points attributed to a particular welfare reform option or the difference between this option and an alternative) to be analyzed in standard regression models or conditional logit regressions.

The analysis of individual-level characteristics as explanatory variables driving the results of the conjoint-experiment is somewhat less straightforward, because the findings of the conjoint (in particular the *average marginal component effects AMCE*, see Hainmüller et al. 2014) are estimated only at the group-level, not at the individual level. Beyond simply comparing the AMCEs[1] estimated for different subsamples, we plan to use interaction effects between respondent characteristics and attribute levels (i.e. the *conditional AMCE* of an attribute level) to estimate the extent to which different groups of respondents differ in the importance they attribute to different attributes of the welfare reform strategies to be evaluated. To give an example: if raising the retirement age has a clear and significant negative effect for younger respondents to support a reform strategy while this item does not significantly affect the way elderly respondents think about the acceptability of a welfare reform strategy, then we see a generational divide over the importance of this item and the type of reform (consumption policy, retrenchment) it stands for.

One open question we still ponder regards how to best analyze conjoint findings in more complex models that involve the interaction between different respondent characteristics. Imagine a hypothesis that would suggest that highly educated women hold priorities that differ from those of lower-educated women. Testing this hypothesis would imply a three-way interaction between gender, education and the conjoint attribute-levels. Similarly, the interaction of two attribute-levels and a respondent characteristic would also immediately result in a three-way interaction (e.g. if women and men hold different priorities on retrenching childcare spending given that spending on the elderly is untouched). Despite the fact that such theoretical expectations seem very plausible and straightforward given the literature, we hardly

[1] The average marginal component effect (AMCE), represents the marginal effect of attribute-level l averaged over the joint distribution of the remaining attribute-levels (Hainmüller et al. 2014, 10). Its interpretation relates to the effect of an attribute-level on the probability of a set being chosen, relative to a set including the status quo on that attribute (and independent of the values of all other attribute-levels). We interpret this coefficient as an indicator of the relative importance of a specific element of a welfare reform agenda on the probability of this agenda being supported.

see them in published conjoint experiments. Why is that? Are there alternative more pragmatic ways to explore conjoint-findings that go beyond the overall pooled samples?

The conjoint experiments will indicate how specific welfare reform elements contribute to increase or decrease the extent to which respondents support or reject an overall welfare reform strategy. We hence plan to read the conjoint coefficients (AMCE) as indications of the relative importance these individual elements have for the respondents (overall or sub-groups). However, we ideally also would like to use these coefficients beyond the experiment itself. On the one hand, we would like to use them as *weights* for the positions respondents hold on these attribute levels and on the other hand, we could imagine using the (sub-group based) coefficients as independent variables to explain meso- and macro-level outcomes such as implemented reforms or elite agendas. However, the coefficients being estimations, they entail uncertainty linked both to the model of estimation itself and to the sub-group chosen to estimate them, which is why this use of the AMCE-coefficients is not straightforward. We also have not encountered published work that utilizes the coefficients in this way, which is why input and comments on this idea would be most helpful.

3.2.2. Contextual determinants of priorities

The ERC project in the context of which we field our surveys is comparative in nature. It comprises eight European countries (Spain, Italy, UK, Ireland, Sweden, Denmark, the Netherlands and Germany), which represent key cases of the four institutional welfare regimes types in Europe and will allow us to build cumulatively on existing research. Hence, when explaining priorities citizens hold for particular welfare reform strategies, we are also interested in contextual, macro-level determinants linked to the institutional or partisan context of national welfare politics. Straightforward hypotheses from the literature e.g. point to the relevance of universalist welfare institutions on the saliency of welfare chauvinism in the preoccupations of citizens or to the fact that high levels of regulatory employment protection increases the importance citizens attribute to consumption policies (Beramendi et al. 2015).

The number of countries included is large for experimental studies, but small for observational ones. Hence, one key question in our project is how we may be able to use experimental research designs well for the purpose of comparative designs. Observationally, the number of countries will not allow us to go beyond relatively broad cross-sectional comparisons of findings. These will be important and telling, especially since we have deliberately chosen to work on countries for which we have a very large stock of established theorizing and empirical research. Nevertheless, the observational comparative design will be limited in its capacity to reveal the causal effect of institutional and other contextual differences on priorities.

Therefore, we plan – in addition to the 8-country survey – to field 2-4 additional country-level surveys, which will include separate, specific priming experiments that should allow us to get closer to some the contextual effects we are interested in. Thereby, macro-level hypotheses should be tested via cross-sectional micro-level comparisons. We hope that by combining the observational findings from our eight countries with these priming experiments, we will be able to produce more robust evidence on the relevance of contextual effects.

The following, final, design section provides an overview of the type of effects we are interested in and how we plan to get at them. We plan to implement these priming experiments via vignettes that are shown to a randomly drawn part of the sample only. Two contextual effects are at the center of our interest: *institutional effects* and *party messages*.

There is a wide literature on *institutional feedback effects* (e.g. Kumlin and Stadelmann-Steffen 2014) which are not only notoriously hard to pin down empirically, but also theoretically controversial since both negative and positive feedback effects can be imagined. In terms of positive (reinforcing) feedback effects, one would expect existing institutions to shape the expectations and imaginary of respondents in a way that reinforces the weight they attribute to this existing logic. In this vein, a large literature has shown that support for universalism and redistribution is reinforced by the universalist and redistributive welfare states of Northern Europe (Svallfors 2012). In contrast, negative (undermining) feedback effects are also theoretically plausible: insider-outsider effects are, e.g., to some extent the paradoxical effect of employment security regulation (Rueda 2007). Hence, institutional feedback effects are theoretically complex and priming experiments may provide additional insights. So far, we have discussed two experiments that should ideally be fielded both in Nordic and Southern European welfare states (e.g. Spain and Sweden):

- A first question relates to the effect of welfare universalism (universal access to welfare benefits for all residents). Universalism may both reduce (via positive feedback) or increase (via negative feedback) the importance respondents attribute to welfare chauvinist measures, i.e. measures which selectively limit the access of immigrants to welfare benefits. Hence, before asking respondents about positions and priorities, we would field a vignette to a random sample of respondents in each country stating “*Developing social security for all people in XX has been a major political achievement of the past. The principle of universalism involves giving all people in this country access to welfare, irrespective of who they are and where they are from. Today, the welfare state is at stake. Therefore, we are interested in your opinion on how it should be reformed*.”

The control group would receive the same statement without the underlined sentence whose only purpose is to highlight the principle of universalism. After the treatment, we would ask respondents about their priorities (ranking/points question similar to the one highlighted above) and ask several trade-off questions (putting universalist policies in contrast to those prioritizing immigrants or nationals) to see if highlighting universalism increases or decreases the importance respondents attribute to welfare chauvinist reforms.

The comparison between the experiment results in Sweden and Spain would be highly relevant because Sweden is indeed a universalist welfare state, whereas Spain is not. Hence, we would expect the treatment to have stronger effects in Sweden than in Spain, because of its higher realism/internal validity.

- A second, very similar priming experiment would relate to the question if employment protection legislation reinforces or reduces support for activation policies in the labor market. Again, a vignette would state “*Now we are interested in what your thoughts are regarding labor market reforms. Employment in the labor market is supposed to provide income and security for all. Regular, indeterminate and full-time work contracts provide employees not only with job security, but also with access to social rights. We would like to know what your views are regarding labor market policy reform.*” Again, only half of the respondents would see the underlined sentence in their vignette, whose purpose is to highlight the principle of employment protection.

  We would then ask respondents about their views on reforms that have differential effects on insiders and outsiders (ranking and trade-off questions). The country-comparison here would lead us to expect stronger results in Spain where dualization is a real political issue, whereas it is not in Sweden.

The other category of macro-level effects we are interested in relates to *party messages* about immigration and the consequential salience of this issue. Studies based on observational data have established that anti-immigrant attitudes are negatively correlated with support for the welfare state (Gilens 1999, Finseraas 2008, Mau and Burkhardt 2009, Schmidt and Spies 2014) and have fiercely discussed whether a high share or large influx of foreigners attenuates support for social spending (Brady and Finnigan 2014, Fox 2004, Eger 2010). To identify a causal relationship, recent survey experiments investigate the effect of priming people with the issue of immigration on their support for social spending (Cappelen and Midtbø 2016, Muñoz and Pardos-Prado 2017) and redistribution (Naumann and Stoetzer 2018). Their results, though, do not consistently point to a causal effect of immigration on attitudes for more or less welfare. Rather, they indicate that some social policies are more strongly affected than others (Cappelen and Midtbø 2016) and that the institutional setup (i.e. whether policies are means-tested or universal) matters (Muñoz and Pardos-Prado 2017). These findings point to the fact that thinking about immigration (and therefore the saliency of immigration in the political debate) not necessarily changes *how much* welfare but rather *what kind* of a welfare state respondents prefer. Thus, being acutely aware of immigration and its potential implications on the welfare state might change the *priorities* of expanding or retrenching different social policies rather than *positions*.

To study this effect of the saliency of immigration on welfare priorities, we plan a priming experiment in which a random sample of treated respondents would read a vignette stating:

"The welfare state of [COUNTRY] provides social security against risks such as illness and unemployment, alleviates the material situation and enhances opportunities for various social groups. In recent years, for example, the percentage of welfare recipients of immigrant origin in [COUNTRY] has increased."

The control group would receive only the first sentence. Both groups are then faced with the question asking about which social policy fields they would most strongly like to expand or retrench if forced to do so (see point distribution questions in Chapter 3.1.2). The treatment as proposed here does neither relate immigrants positively (by stating their need) nor negatively (by depicting them as an increasing cost pressure) to the welfare state but just awakens attention to the fact that immigrants make up an increasing share of welfare state recipients.

Theoretically, we would expect that people primed with immigration tend to favor consumption over investment as consumption can be more easily targeted at groups which are perceived deserving. In contrary, investment policies such as active labor market policies or universal childcare benefit immigrants alike if not above average. Moreover, priming immigration might enhance the sympathy for contribution-based over universalist and means-tested social policy programs. To empirically disentangle the effects on the popularity of policy fields and on the support for a principle (insurance, universalism, means-tested) we would optimally field the experiment in two countries where the same policy fields are organized differently. For example, a comparison between the UK and the Netherlands would allow us to exploit differences in the health system (insurance based in the Netherlands, universal in the UK) as well as concerning family and childcare allowances (universal in the Netherlands, means-tested in the UK). Moreover, in both countries immigration has been brought up in the public debate in relation to the welfare state. Hence, they would constitute a realistic context for the treatment as the public can be expected to be more aware about the social policies that benefit immigrants more or less.

4. Planned surveys

In this final section, we provide some very brief indications regarding the surveys we plan to field to implement the items and designs presented in this paper. Fieldwork is planned for summer/fall 2018 (we want to avoid fielding the survey shortly prior to an election, which is why they will not be in the field at identical times).

We plan to field two types of surveys. The main survey, in which we will field the conjoint experiments as explained above, the trade-off and ranking questions, as well as a range of attitudinal and socio-demograhic questions will be fielded in 8 European countries (2 from each regime type: ES, I, UK, IRE, GER, NL, SW, DK) with identical wording. However, our preparatory work for this survey – in particular a detailed qualitative content analysis and quantitative media analysis of the welfare reform agendas in all 8 countries – has made it very clear that the current social policy reform agendas in the countries look quite different. While Ireland e.g. recovers from massive austerity by re-expanding spending in health, education and active labor market policies, the German social policy agenda has been almost exclusively about spending expansion in family and unemployment policy in the recent years and the reform debate in Spain is still clearly dominated by retrenchment and an almost exclusive focus on labor market issues. Therefore, we will also field country-specific surveys in 2-4 countries, in which we will adopt more context-specific wordings for the conjoints, ranking- and trade-off questions. The idea will be to measure the same concepts with more context-sensitive wording. One goal will be to test the robustness of the findings from the identically worded survey as compared to the country-specific surveys. However, the second goal of the country-specific surveys will be to field the vignette/priming experiments outlined in the last section on context-effects. We do not want to include these experiments in the comparative survey, for reasons of length and in order not to affect the findings, but we want to compare the observational comparative findings with the experimental ones.

The identically worded, comparative survey of about 25 minutes will be fielded among 1500 respondents drawn from large country-specific online panels (we are currently evaluating different quotes from survey companies), probably including quotas for gender, age, education and labor market status. The design of the pre-tests is not decided yet, but will certainly include a soft launch (among 100-200 respondents) in each country. We might possibly also consider qualitative pre-tests for the more complex parts of the questionnaire (and we are happy to receive advice on the best strategy).

The questionnaire will start with questions regarding the positions of respondents on the reform items that will later on figure in the conjoint designs and on items asking them about their perception of fiscal and political constraints (in order to establish if they think the scenario we put them in is realistic). We will then field the conjoint surveys, the ranking questions, the trade-off items and a battery of questions on other political attitudes before asking about electoral preferences and a battery of socio-demographic questions. One open question is if we should rely on the socio-demographic information that survey companies have of their panel participants (e.g. age, gender, income, household composition, employment status, sector of employment) or if we should ask this information (again). The country-specific surveys will also include about 100-1500 respondents and will be slightly shorter.

References

Beramendi, P., Häusermann, S., Kitschelt, H., & Kriesi, H. (eds.) (2015): *The Politics of Advanced Capitalism*. New York: Cambridge University Press.

Beramendi, P., & Rehm, P.. 2016. "Who Gives, Who Gains? Progressivity and Preferences." *Comparative Political Studies* 49 (4): 529–63.

Blekesaune, M., & Quadagno, J. (2003). Public Attitudes toward Welfare State PoliciesA Comparative Analysis of 24 Nations. *European Sociological Review*, *19*(5), 415–427.

Brady, D. & Finnigan, R. (2014). Does Immigration Undermine Public Support for Social Policy? *American Sociological Review 79*(1): 17-42.

Bremer, B., & Bürgisser, R. (n.d.). Public Opinion on Welfare State Recalibration in Times of Austerity: Evidence from a Conjoint Experiment.

Brooks, C., & Manza, J. (2008). *Why Welfare States Persist: The Importance of Public Opinion in Democracies*. University of Chicago Press.

Busemeyer, Marius (2014): Trade-offs between social investment and passive transfers in the new welfare state: New political coalitions in European welfare states? *Manuscript.*

Busemeyer, M. R., & Garritzmann, J. L. (2017). Public opinion on policy and budgetary trade-offs in European welfare states: evidence from a new comparative survey. *Journal of European Public Policy*, *24*(6), 871–889.

Cappelen, C. & Midtbø, T. (2016). Intra-EU Labour Migration and Support for the Norwegian Welfare State. *European Sociological Review 32*(6): 691-703.

Cavaillé, C., & Trump, K.-S. (2015). The Two Facets of Social Policy Preferences. *The Journal of Politics*, *77*(1), 146–160.

Chapman, J., Dean, M., Ortoleva, P., Snowberg, E., & Camerer, C. (2017). *Willingness to Pay and Willingness to Accept are Probably Less Correlated Than You Think* (No. w23954). Cambridge, MA: National Bureau of Economic Research.

Dimick, M., Rueda, D., & Stegmueller, D. (2017). The Altruistic Rich? Inequality and Other-Regarding Preferences for Redistribution. *Quarterly Journal of Political Science*, *11*(4), 385–439.

Eger, M. (2010). Even in Sweden: the effect of immigration on support for welfare state spending. *European Sociological Review 26*(2): 203-217.

Emmenegger, P. (2009): Barriers to Entry: Insider/Outsider Politics and the Determinants of Job Security Regulations. *Journal of European Social Policy* 19(2): 131-146.

Estevez-Abe, M., Iversen, T. & Soskice, D. (2001): Social Protection and the Formation of Skills: A Reinterpretation of the Welfare State. In: Peter A. Hall and David Soskice (eds.): *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press, 145-183.

Fernández, J. J., & Jaime-Castillo, A. M. (2013). Positive or Negative Policy Feedbacks? Explaining Popular Attitudes Towards Pragmatic Pension Policy Reforms. *European Sociological Review*, *29*(4), 803–815.

Finseraas, H. (2008). Immigration and Preferences for Redistribution: An Empirical Analysis of European Survey Data. *Comparative European Politics*, *6*(4), 407–431.

Fossati, F., & Häusermann, S. (2014). Social Policy Preferences and Party Choice in the 2011 Swiss Elections. *Swiss Political Science Review*, *20*(4), 590–611.

Fox, C. (2004). The changing color of welfare? How Whites' attitudes toward Latinos influence support for welfare. *American Journal of Sociology 110*(3): 580-625.

Franchino, F., & Segatti, P. (2017). Public opinion on the Eurozone fiscal union: evidence from survey experiments in Italy. *Journal of European Public Policy*, *0*(0), 1–23.

Gallego, A., & Marx, P. (2017). Multi-dimensional preferences for labour market reforms: a conjoint experiment. *Journal of European Public Policy*, *24*(7), 1027–1047.

Giger, N. & Nelson, M. (2013): The Welfare State or the Economy? Preference, Constituencies, and Strategies for Retrenchment. *European Sociological Review* 29(5): 1083-1094.

Gilens, M. (1999). *Why Americans hate welfare. Race, media, and the politics of antipoverty policy*. Chicago: University of Chicago Press.

Goerres, A. & 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.

Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2013). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis*, *22*(1), 1–30.

Hall, P. & Soskice, D. (2001). Varieties of Capitalism. Institutional Foundations of Comparative Advantage. Oxford: Oxford University Press.

Hansen, J. M. (1998). Individuals, Institutions, and Public Preferences over Public Finance. *American Political Science Review*, *92*(3), 513–531.

Häusermann, S., Kurer, T., & Schwander, H. (2015). High-skilled outsiders? Labor market vulnerability, education and welfare state preferences. *Socio-Economic Review*, *13*(2), 235–258.

Häusermann, S., Kurer, T., & Schwander, H. (2016). Sharing the Risk? Households, Labor Market Vulnerability, and Social Policy Preferences in Western Europe. *The Journal of Politics*, *78*(4), 1045–1060.

Häusermann, S., Thomas K. & Traber, D. (2017). «The Politics of Trade-offs: Studying the Dynamics of Welfare State Reform with Conjoint Experiments», manuscript.

Huber, J. D. 2017. Exclusion by Elections: Inequality, Ethnic Identity, and Democracy. Cambridge: Cambridge University Press.

Iversen, T., & Soskice, D. (2001). An Asset Theory of Social Policy Preferences. *American Political Science Review*, *95*(4), 875–893.

Jensen, C. (2012). Labour market- versus life course related social policies: understanding cross-programme differences, *Journal of European Public Policy*, 19(2): 275-291.

Kölln, A.-K., & Wlezien, C. (2016). Conjoint Experiments on Political Support for Governmental Spending Profiles.

Korpi, Walter (2006). “Power Resources and Employer-Centered Approaches in Explanations of Welfare States and Varieties of Capitalism: Protagonists, Consenters, and Antagonists”, *World Politics* 58(2): 167-206.

Kumlin, S. & Stadelmann-Steffen, I. (2014): Citizens, policy feedback, and European welfare states. In: Staffan Kumlin and Isabelle Stadelmann-Steffen (eds.): *How welfare states shape the democratic public: Policy feedback, participation, voting, and attitudes.* Cheltenham, UK and Northampton, MA, USA: Edward Elgar, 3-16.

Margalit, Y. (2013). Explaining Social Policy Preferences: Evidence from the Great Recession. *American Political Science Review*, *107*(1), 80–103.

Mau, S. & Burkhardt, C. (2009). Migration and welfare state solidarity in Western Europe. *Journal of European Social Policy* 19(3): 213-229.

Muños, J. & Pardos-Prado, S. (2017). Immigration and Support for Social Policy: An Experimental Comparison of Universal and Means-Tested Programs. *Political Science Research and Methods.* 1-19.

Naumann, E. & Stoetzer, L. (2018). Immigration and support for redistribution: survey experiments in three European countries. *West European Politics 41*(1): 80-101.

Palier, B. & Thelen, K. (2010). "Institutionalizing Dualism: Complementarities and Change in France and Germany," *Politics & Society* 38(1): 119-148.

Paster, T. (2013). “Business and Welfare State Development: Why Did Employers Accept Social Reforms?” *World Politics* 65(3): 416-451.

Pierson, P. (1996). The New Politics of the Welfare State. *World Politics*, *48*(2), 143–179.

Pierson, P. (2001). *The new politics of the welfare state*. Oxford: Oxford University Press.

Pontusson, J. & David W. (2017). “Macroeconomic conditions, inequality shocks and the politics of redistribution, 1990-2013”, *Journal of European Public Policy* (online first, doi: 10.1080/13501763.2017.1310280).

Rehm, P. (2009). Risks and Redistribution: An Individual-Level Analysis. *Comparative Political Studies*, *42*(7), 855–881.

Rehm, P. (2011). Social Policy by Popular Demand. *World Politics*, *63*(2), 271–299.

Rehm, P. (2016). *Risk Inequality and Welfare States: Social Policy Preferences, Development, and Dynamics*. Cambridge University Press.

Rehm, P., Hacker, J. S., & Schlesinger, M. (2012). Insecure Alliances: Risk, Inequality, and Support for the Welfare State. *American Political Science Review*, *106*(2), 386–406.

Rueda, D. (2005). Insider-Outsider Politics in Industrialized Democracies: The Challenge to Social Democratic Parties. *American Political Science Review*, *99*(1), 61–74.

Rueda, D. (2007). *Social Democracy Inside Out. Partisanship and Labor Market Policy in Advanced Industrialized Democracies*. Oxford: Oxford University Press.

Rueda, D. (2017). Food Comes First, Then Morals: Redistribution Preferences, Parochial Altruism, and Immigration in Western Europe. *The Journal of Politics*, 000–000.

Schmidt, A. & Spies, D. (2014). Do Parties “Playing the ‘Race Card” Undermine Natives’ Support for Redistribution? Evidence from Europe. *Comparative Political Studies* 47(4): 519-549.

Svallfors, S. (2012). Welfare States and Welfare Attitudes. In S. Svallfors (Ed.), *Contested Welfare States: Welfare Attitudes in Europe and Beyond*. Stanford University Press.

Walter, S. (2010). Globalization and the Welfare State: Testing the Microfoundations of the Compensation Hypothesis. *International Studies Quarterly*, *54*(2), 403–426.

Tables

Table 1: Attributes and levels of the expansion conjoint profiles

Attribute	Level
Old age pensions	Pension benefits remain unchanged
	Additional means are directed to the pension system to stabilize current pension levels for all
	Minimum pensions are increased
	Public subsidies are introduced to support the pension savings of those who are now below the age of 45
Childcare services	public childcare services remain unchanged
	The quality of public childcare services is increased through additional resources
	Availability and quality of public childcare services is improved for low-income families
Higher education	higher education remains unchanged
	The quality of higher education is improved through additional resources
	Support for students from low-income families to access higher education is increased
Benefits for the unemployed	unemployment benefits remain unchanged
	unemployment benefits are increased for all
	unemployment benefits are increased for those who have contributed most
	unemployment benefits are increased for those who need it most even if they have not contributed much
Reintegration services for the unemployed	reintegration services for the unemployed remain unchanged
	the quality of reintegration services offered to all job-seekers is improved through additional resources
	reintegration services are expanded for long-term unemployed and precarious workers
	reintegration services are expanded to support immigrants' labor market integration
Social assistance	social assistance remains unchanged
	Social assistance services and benefits are improved for all people in need
	Social assistance services and benefits are improved for immigrants

Table 2: Attributes and levels of the retrenchment conjoint profiles

Attribute	Level
Old age pensions	pension benefits remain unchanged
	the statutory age of retirement increases by 2 years
	the statutory age of retirement increases by 2 years for people with above-average income only
	people below the age of 45 have to pay higher contributions to the pension system
Childcare services	public childcare services remain unchanged
	fees of public childcare services increase for everyone
	fees of public childcare services increase for people with middle and upper incomes
Higher education	higher education remains unchanged
	fees for higher education increase for all students
	fees for higher education increase for students from families with with middle and upper incomes
Benefits for the unemployed	unemployment benefits remain unchanged
	unemployment benefits are lowered for all
	unemployment benefits are lowered for people with above-average incomes prior to becoming unemployed
	the minimum time that people need to pay contributions in order to receive unemployment benefits is increased
Reintegration services for the unemployed	reintegration services for the unemployed remain unchanged
	reintegration services are cut back for all unemployed
	reintegration services are restricted to long-term unemployed and precarious workers only
	reintegration services are restricted to national citizens only
Social assistance	social assistance remains unchanged
	social assistance is restricted to national citizens only
	the conditions for receiving social assistance and their control become stricter

Table 3: Rating questions for the expansion and retrenchment scenarios

3.1	**Rating reforms, EXPANSION**
	Imagine not all is possible, which improvements of social benefits do you consider most important? You have 100 points. Give most points to the reforms that you consider most important and least points to the areas you consider least important.
	Minimum pensions are increased
	Additional means are directed to the pension system to stabilize current pension levels for all
	The quality of public childcare services is increased through additional resources
	Availability and quality of public childcare services is improved for low-income families
	The quality of tertiary education is improved through additional resources
	Support for students from low-income families to access tertiary education is increased

3.2	**Rating fields, EXPANSION**
	Imagine not all is possible, in which fields do you consider improved benefits most important? You have 100 points. Give most points to the areas in which you would most like to see improvements and least points to the areas where you think improvements are least important.
	old age pension
	public childcare services
	higher education
	benefits for the unemployed
	reintegration services for the unemployed
	social and labor market integration of immigrants

3.3	**Rating reforms, RETRENCHMENT**
	Imagine the government has to cut back on social security. Which reductions do you find most acceptable? You have 100 points. Give most points to the reforms that you consider most acceptable and least points to those reforms that you find least acceptable.
	the statutory age of retirement increases by 2 years
	the statutory age of retirement increases by 2 years for people with above-average income only
	fees of public childcare services increase for everyone
	fees of public childcare services increase for people with above-average income
	fees for tertiary education increase for all students
	fees for tertiary education increase for students from families with above-average income

3.4	**Rating fields, RETRENCHMENT**
	Imagine the government has to cut back on social security, in which fields do you consider reductions most acceptable? You have 100 points. Give most points to the areas in which you would find a reduction of benefits most acceptable and least points to those areas in which you would find reductions least acceptable.
	old age pensions
	public childcare services
	higher education
	benefits for the unemployed
	reintegration services for the unemployed
	social and labor market integration of immigrants

Table 4: Social policy trade-off questions

3.5	**Trade-off generational divide**
	To what extent would you agree to the following proposal? Guarantee stable old age pensions for future generations, at the cost of lowering benefits for the currently retired.
1	disagree
2	rather disagree
3	rather agree
4	agree
5	NA

3.6	**Trade-off investment/consumption**
	To what extent would you agree to the following proposal? Improve early childcare services for working parents at the cost of lower child allowances.
1	disagree
2	rather disagree
3	rather agree
4	agree
5	NA

3.7	**Trade-off insider-outsider divide**
	To what extent would you agree to the following proposal? Improve support for the unemployed at the cost of reducing maximum pension levels.
1	disagree
2	rather disagree
3	rather agree
4	agree
5	NA

3.8	**Trade-off class** To what extent would you agree to the following proposal? Increase support for students from lower-income families to access higher education, at the cost of higher tuition fees for students from middle income families.
1	disagree
2	rather disagree
3	rather agree
4	agree
5	NA

3.9	**Trade-off multiculturalism** To what extent would you agree to the following policy? Improve the support for labor market integration of migrants at the cost of slightly less generous old age pensions for everyone.
1	disagree
2	rather disagree
3	rather agree
4	agree
5	NA