

This article was downloaded by: [Vrije Universiteit, Library]

On: 1 June 2011

Access details: Access Details: [subscription number 907218003]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Comparative Policy Analysis: Research and Practice

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713672306>

The Matching Problem within Comparative Welfare State Research: How to Bridge Abstract Theory and Specific Hypotheses

Sabina Stiller^a; Kees van Kersbergen^b

^a Radboud University Nijmegen, The Netherlands ^b Vrije Universiteit Amsterdam, The Netherlands

To cite this Article Stiller, Sabina and van Kersbergen, Kees(2008) 'The Matching Problem within Comparative Welfare State Research: How to Bridge Abstract Theory and Specific Hypotheses', *Journal of Comparative Policy Analysis: Research and Practice*, 10: 2, 133 – 149

To link to this Article: DOI: 10.1080/13876980802028057

URL: <http://dx.doi.org/10.1080/13876980802028057>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The Matching Problem within Comparative Welfare State Research: How to Bridge Abstract Theory and Specific Hypotheses

SABINA STILLER* and KEES VAN KERSBERGEN**

*Radboud University Nijmegen, The Netherlands; **Vrije Universiteit Amsterdam, The Netherlands

ABSTRACT *This paper draws attention to the problem of matching abstract theory and specific hypotheses within welfare state research, which reinforces the dependent variable problem and entails methodological difficulties. We show that matching theory and hypotheses is a ubiquitous problem in the literature. We further elaborate and illustrate the argument with an empirical example from our research on structural welfare state reform. We observe two methodological problems: 1) the risk of drawing conclusions about one level of analysis using evidence from another; 2) the problem of translating causal mechanisms formulated at a high level of abstraction to a lower level.*

Introduction

Scholarly work on the welfare state is characterized by many theoretical and empirical schools. Recently, research attention has been redirected to explaining variation in welfare state *reform*. The conceptualization of the dependent variable “reform”, however, has caused considerable theoretical, empirical and methodological confusion and, accordingly, scholarly debate, both within mainstream approaches and between the mainstream literature and other approaches (such as

Sabina Stiller is Assistant Professor of Comparative Politics at the Department of Political Science at Radboud University Nijmegen. She works on comparative welfare state reform, the role of ideas and political agency in policy change, and methodological questions in comparative welfare state research. Her PhD dissertation is entitled “Innovative Agents versus Immovable Objects. The Role of Ideational Leadership in German Welfare State Reform” (2007).

Kees van Kersbergen is Professor of Political Science at the Department of Political Science of the Vrije Universiteit Amsterdam. His main research interests lie in comparative political sociology, politics and economy and broadly concern issues of change in politics and governance: the changing power of the national state; the changing power and position of political parties in advanced democracies; the development of welfare states and social policies in OECD countries; European integration and Europeanization. He is the author of *Social Capitalism* (Routledge 1995) and co-editor (with Philip Manow) of *Religion, Class Coalitions and Welfare State Regimes* (Cambridge University Press 2008, forthcoming).

Correspondence Address: Sabina Stiller, Nijmegen School of Management, Department of Political Science, Radboud University Nijmegen, P.O. Box 9108, NL-6500 HK Nijmegen, The Netherlands.
Email: s.stiller@fm.ru.nl

the regulation school with its focus on the changes in production regimes and issues of “workfare”).

The debate on the dependent variable problem (see Green-Pedersen 2004, Kühner 2007), that is, the conceptual and operational ambiguity about how the explanatory problem of welfare state reform is formulated for research purposes, has considerably improved our understanding of reform. We now have sharper specifications of welfare state reform and its various *dimensions* (see Pierson 2001a and below), more valid operationalizations and more reliable measurements of crucial concepts (such as decommodification, see Allan and Scruggs 2004), better datasets (see the dataset collected by Allan and Scruggs: <http://sp.uconn.edu/~scruggs/wp.htm>), important debunking efforts of overly speculative theories (for example Swank 2002, Castles 2004) and promising methodological innovations (for example Jaeger and Kvist 2003, Vis 2007).

Part of the solution to the dependent variable problem surely lies in further improving the quality of the data, operationalizations, measurements and methods as well as disaggregating research attention to the programme level (Kühner 2007). But even if we would have perfect measures and operationalizations, we are still left with the problem that so many different and competing (explanatory) theories seem to fit the same data on welfare state reform. In order to explain welfare state reforms, a great variety of explanatory approaches has been formulated. This variety reinforces the dependent variable problem as different theories suggest different conceptualizations and operationalizations of the welfare state or reform as the dependent variable. In any case, these rival theories cannot all be correct at the same time. But how to test such competing theories empirically? We think that in research practice there has been insufficient awareness of the necessity, but also of the complexity and difficulty, of competitive hypothesis testing. It involves a comparison of types of explanation, for instance whether political or economic factors are driving welfare state reform. Next, one needs to establish the explanatory status of the theories involved. Does the theoretical model specify causes and effects in a correlational approach or does the theory mainly describe causal mechanisms linking cause and effect in a historical approach? Also, the level of analysis at which theories specify their hypotheses differ. For competitive testing, it matters whether one aims to explain changes *of* or *in* welfare state regimes or social policy programmes. Finally, there is the issue of the operational and empirical specificity of the hypotheses that can vary between theories and that may complicate comparison between theories with respect to their empirical validity.

In this paper we cannot hope to even start solving all these issues, but propose to focus on the operational and empirical specificity of hypotheses. The reason for this choice is that frequently the potential problems involved here are either not recognized or not well addressed. We believe that a better appreciation of the problem is critically important as different operationalizations tend to lead to different conclusions about the hypotheses that are deduced from the theory being tested. We define the issue as a problem of *matching*, which refers to the difficulty of matching propositions formulated at a high(er) level of abstraction with hypotheses than can be tested empirically. We call attention to this phenomenon, particularly as the matching problem tends to reinforce the dependent variable problem, a problem

that is far better known and dealt with in the literature (Green Pedersen 2004, Clasen and Siegel 2007, Kühner 2007).

The structure of the paper is as follows. In the next section we present a short overview of recent welfare state research to provide the background for an identification of the problem of the dependent variable and the re-specification of welfare state reform in terms of different dimensions. In the third section we elaborate the matching problem, arguing that cause and effect tend to be specified on different levels of analysis, as a result of which the conceptualization and operationalization of the causal mechanism(s) and the dependent variable becomes complicated and the dependent variable problem is reinforced. By discussing examples using different analytical techniques, we show that matching is a problem of the field as a whole. Some circumvent the problem by remaining at a high level of abstraction and offer only loosely connected empirical illustrations. Others evade the problem by discarding theory and focusing on technically advanced quantitative tests of loosely connected (if at all) models that include variables from widely diverging theories. Most empirical researchers find practical solutions, for instance by combining a variable-oriented and a case-oriented approach, but still fail to make explicit how they have tried to match theoretical statements at one level with empirical statements at another level.

Our point is that we would like to raise awareness of the problem and spell out the intricacies involved. Instead of offering further detailed criticisms of other people's work, we present as a final example the work of one of us (Stiller 2007) to track in some detail how one travels from (abstract) theory, the level at which the causes of reform are formulated, to empirically testable hypotheses on causal mechanisms. In other words, to clarify better the matching problem and illustrate how to deal with it explicitly, we discuss one example of research practice in some more detail rather than present a sketchy discussion of a multitude of studies. Finally, the last section draws together the lessons of the foregoing discussion and ends with some recommendations on how to proceed from here.

Welfare State Reform: the Dependent Variable Problem

Until roughly 1990, the welfare state (its emergence and expansion) was routinely conceptualized in terms of welfare effort and operationalized as social spending. The literature of the early 1990s began to rethink what precisely welfare state theory should explain. The social spending variable was rightfully criticized for its loose correspondence to the theoretical issues of social democratization (Esping-Andersen 1990). Esping-Andersen argued that there was a striking conceptual indifference in the literature with respect to the dependent variable: the welfare state. Starting from the judgement that "expenditures are epiphenomenal to the theoretical substance of welfare states" (1990: 19), and the reflection that "it is difficult to imagine that anyone struggled for spending per se" (1990: 21), he suggested that the study of welfare states should focus on the quality of social rights, the typical patterns of stratification, and the manner in which the state, the market and the family interacted in the production of social welfare.

By looking beyond spending patterns, Esping-Andersen was able to distinguish three types of welfare state regimes: a social democratic, a liberal and a corporatist or

conservative regime. These regimes differed with respect to the major institutions guaranteeing *social security* (the state, the market or the family); the kind of *stratification* systems upheld by the institutional mix of these institutions (the extent of status and class differentiation, segmentation and inequality typically implied in social security systems); and the degree of *de-commodification*, that is to say “the degree to which individuals, or families, can uphold a socially acceptable standard of living independently of market participation” (Esping-Andersen, 1990: 37). This inspired a whole new generation of research that focused on the explanation of variation between regimes, provoking discussion on the dependent variable “welfare state regime” and the correct typology to use as a suitable standard in comparative analysis.

Overlapping with this new way of looking at the welfare state, there was a gradual shift in the research attention away from “expansion” and “variation” of welfare state regimes towards “reform” and “retrenchment” (or their absence) of the welfare state as the main explanatory problem. A groundbreaking historical institutionalist study of welfare state *change* was Paul Pierson’s (1994) *Dismantling the Welfare State?* His main finding was that, in spite of mounting pressures from liberal forces symbolized by the names of Reagan and Thatcher and in contrast to changes in the arenas of macroeconomic policy, industrial relations or regulatory policy, “the welfare state stands out as an island of relative stability” (1994: 5). Welfare states resisted change. In explaining this unexpected phenomenon, Pierson focused on institutional structures and electoral mechanisms. In his view, the former included networks of welfare bureaucracies and services in the policy areas of social housing, health care, education, public assistance and social security, the very existence of which was bound to the status quo in social policy and which therefore mounted powerful resistance against attempts of retrenchment. These professional networks were created by postwar welfare state development, and once established they were able to muster substantial veto power against reform efforts (Pierson, 1996: 147). Because these structures stood for path continuity, a weakening of the historical supporters of welfare state expansion (for example social democracy) did not necessarily translate into commensurate weakening of social policy.

Moreover, Pierson argued, “frontal assaults on the welfare state carry tremendous electoral risks” (1996: 178). While welfare expansion usually generated a popular politics of credit claiming for extending social rights and raising benefits to an increasing number of citizens, austerity policies affronted large groups of voters. Even “retrenchment advocates . . . confront a clash between their policy preferences and their electoral ambitions” (Pierson, 1996: 146) and, as a rule, the latter prevailed.

In the literature on reform, there arose considerable confusion around the question “what is to be explained” and the specification of the dependent variable. Pierson (2001b) observed that there was a lack of consensus on welfare state outcomes, particularly with respect to the issue of how much welfare states had actually changed since the end of the Golden Age of growth, that is to say, roughly since the 1980s. The controversy over the dependent variable was first of all a result of the indistinctness of the concept of the welfare state itself. Too many and quite divergent phenomena were being discussed under the same heading. Welfare state research seemed to suffer from a weakness well known in comparative politics and comparative policy analysis: concept stretching. Related to this was the problem of

which indicators to use for the operationalization of “the welfare state” or reform. Finally, Pierson also noticed theoretical weaknesses that concerned the implicit assumption in many studies that one could measure welfare state change on a single scale. He observed that there was a tendency to reduce the problem of welfare state retrenchment and reform to a dichotomy of “less” versus “more” and “intact” versus “dismantled”, which was an unwarranted theoretical simplification.

Pierson proposed to solve the dependent variable problem and improve our understanding of welfare state change by looking at three dimensions:

- (1) **Recommodification:** the attempt “to restrict the alternatives to participation in the labour market, either by tightening eligibility or cutting benefits” (Pierson 2001b: 422), that is to say, to strengthen the whip of the labour market;
- (2) **Cost containment:** the attempt to keep balanced budgets through austerity policies, including deficit reduction and tax moderation;
- (3) **Recalibration:** “reforms which seek to make contemporary welfare states more consistent with contemporary goals and demands for social provision” (Pierson 2001b: 425).

Making use of the strengths of Esping-Andersen’s (1990, 1999) regime approach and the country and policy area studies he brought together in his edited volume, Pierson (2001b) inferred that each regime (social democratic, liberal or conservative) was characterized by its own specific “new politics” of welfare state reform. For instance, in the liberal regime voters are less likely to be attached to the welfare state than in the conservative or social democratic models. In this regime recommodification is the pivotal dimension of welfare state reform. In the social democratic welfare regime, voters are highly attached to, and dependent on, the welfare state. Recommodification is not so much on the political agenda of reform, but – if only because of the sheer size of the public sector – cost containment is. The conservative regime is probably the most ill-adapted model of the three worlds of welfare capitalism, as a result of which recalibration and cost containment are the two dominating dimensions of reform. Here, the issue is how to stimulate job growth in the underdeveloped service sector and how to contain the exploding costs of pensions, disability benefits, and health care.

Contemporary research seems to have taken seriously Pierson’s suggestion of the multi-dimensionality of the dependent variable (see Wincott 2003). For example, the issue of *cost containment* and *retrenchment* (or *regress*, as Korpi 2003 calls it) is taken up, for instance, in the broad study by Huber and Stephens (2001) and by Green-Pedersen (2004). Olli Kangas’ (2004) study of sickness benefits in 18 OECD countries indicates that – in spite of continuing institutional distinctiveness – cost containment and retrenchment in Scandinavia causes convergence towards the continental welfare states. Korpi and Palme (2003) have shown that retrenchment has become a significant phenomenon in many countries.

With respect to *recommodification*, one can find at least five different notions in the literature. First, as already noted, Pierson (2001b) defines recommodification as the effort to restrict the alternatives to participation in the labour market, either by tightening eligibility or cutting benefits. Second, there are scholars who link recommodification to changes in the international economy (Geddes 1994). Third,

Breen (1997) looks at recommodification from a risk and stratification perspective and argues that recommodification occurs whenever hedging institutions, mechanisms or arrangements are weakened. Fourth, Bonoli (1998) sees recommodification as the opposite of decommodification and analyzes measures that are intended to weaken the position of commodified workers. Finally, Holden (2003) holds that recommodification occurs when states withdraw from the field of social welfare.

Recalibration seems to be a much less studied topic, but there are scattered (mainly German) examples, such as Lamping and Rüb's (2004) study of the restructuring of the German pension system and Leibfried and Obinger's (2003) analysis of the direction of social policy reforms in Germany.

Finally, a dimension of welfare state reform that Pierson did not distinguish (probably because he saw it as an aspect of recommodification) concerns *activation* and (market-driven) *workfare*. There is a host of literature also on this dimension, particularly prominent in the regulation approach (for an overview see Vis 2007), but the issue also seems particularly "hot" among British social scientists who study welfare policies under New Labour in a comparative perspective (for example Clasen and Clegg 2003, Lindsay and Mailand 2004, Taylor-Gooby *et al.* 2004, Wright *et al.* 2004). Moreover, it has inspired studies of activation in cases such as Germany and Austria (Ludwig-Mayerhofer and Wroblewski 2004), the Netherlands (Van Oorschot 2004), and the Netherlands and Denmark (Van Oorschot and Abrahamson 2003).

Although Pierson considerably improved our understanding of welfare state reform, his analyses still suffered somewhat from a weakness that is inherent to the institutional approach. Institutional analyses are very well capable of explaining institutional *resilience*, but have much more difficulties with understanding institutional and policy *change* (see Taylor-Gooby 2002). There seems to be a growing awareness of this explanatory problem of institutional approaches.

The Matching Problem in Comparative Welfare State Research

Depending on the type of theory proposed, scholarly work highlights different aspects of welfare state reform, that is, of the dependent variable. Equally important is to note that there are various explanatory approaches that compete with each other, but nevertheless seem to fit the same empirical data. In other words, it seems exceedingly difficult to properly assess their theoretical and empirical value. We have already introduced four aspects of this issue that we now will briefly elaborate: the type of explanation; the explanatory status of the theory (in terms of the specificity of the underlying causal model); the level of analysis; and the operational and empirical specificity of the hypotheses.

Looking at *types of explanation* for welfare state reform, we find a) economic explanations, including studies of internationalization and globalization (for example Glatzer and Rueschemeyer 2005, Kemmerling 2005, Kittel and Winner 2005), b) institutional explanations (for example Pierson 2001a, Obinger *et al.* 2005), party-political explanations (Kitschelt 2001, Green-Pedersen 2002, Burgoon and Baxandall 2005), d) discourse and framing-related explanations (Béland 2005, Schmidt 2000, Schmidt *et al.* 2005), and e) hypotheses on how European level policies affect national welfare states (Ferrera and Gualmini 2004, Natali 2005).

Next, the *explanatory status* refers to the precise type of causal argument employed in a study. Macro-level quantitative analyses usually employ a correlation approach focusing on the relation between variables that represent cause(s) and effect(s) (for example Castles 2004). Meso-level studies usually study causal mechanisms in addition to specifying causes and effects (for example, the many case studies we find in the field; see for an overview Starke 2006). Rare but important are the studies that combine both, as Huber and Stephens (2001) do.

If we focus on the *level of analysis* we note that welfare state reform is usually either analyzed at the macro-level of the determinants of regime change via comparative case studies or cross-national (pooled) quantitative studies, or at the meso-level of the determinants of changes in single programmes, such as unemployment insurance, sickness benefits, pension benefits and so on (for example on pensions: Bonoli 2001, Immergut *et al.* 2007; or on unemployment: Clasen 2005).

The final category is the *operational and empirical specificity of the hypotheses*. This is the aspect we focus on further because it is all too often neglected or not properly addressed in the literature. Different operationalizations may lead to different conclusions about the hypotheses that are deduced from the theory that is being tested. Remember Esping-Andersen's (1990: 19) insight on the operationalization of the dependent variable: "Most . . . studies claim to explain the welfare state. Yet their focus on spending may be misleading. Expenditures are epiphenomenal to the theoretical substance of welfare states." His operationalization remained closer to the theoretical substance in terms of the various qualitative dimensions of the welfare state. This not only opened up a whole new way of looking at welfare state development, but also in one stroke made redundant a whole series of theories that had focused on social spending alone.

The issue on which we focus is the matching problem: frequently, causes and effects are specified at different levels of analysis, which has consequences for the conceptualization and operationalization of the causal mechanisms and the dependent variable. The crucial point is that this matching problem reinforces the dependent variable problem. We believe that the matching problem is a general phenomenon in various welfare state research traditions. If we make a distinction between, on the one hand, the mainstream literature that uses welfare state regime theory as an analytical instrument, and, on the other hand, the regulation approach to political economy (see Vis 2007), we observe that – in spite of their different levels of abstraction – they struggle with the same matching problem.

One way of "solving" the problem is to circumvent it by remaining at a high abstract theoretical level. This implies that theory and hypotheses are not matched at all because no (clear) empirical test strategy is offered. Torfing (2001), for instance, takes up the difficulty of institutional analysis referred to above, namely how to deal with institutional and policy change. He posits that in four important areas change has occurred in the welfare state: 1) demand-side macro-economic policy has been replaced by a supply-side orientation; 2) there has been a shift from welfare to workfare; 3) the scale and level of social policy making has shifted because authority has been displaced away from the national level; and 4) decentralized governance networks have taken over hierarchical state intervention. His proposal is to recognize that change is evolutionary and the result of the interplay between path

shaping and path dependency. He offers a well informed and refined comparison between various institutional approaches and discusses how to analyze institutional change. To this end, Torfing proposes new theoretical concepts such as dislocation and path shaping, embedded in a more general neo-Gramscian framework that stresses institution-building as hegemonic projects. He then speculates how path shaping and path dependency interact to produce change.

His theoretical position is formulated at a very high level of abstraction:

The political answer of the social and political forces to the obstructing effect of the institutional context on both policy output and policy outcome is the development of self-reflexive strategies. These aim to reduce or eliminate the effect of path dependency by deliberately restructuring the institutional framework for adopting and implementing new policies. Such strategies involve second-order reflections, since the hegemonic forces are not concerned with the form and content of the new policy, but rather with the tactical question of what they can do in order to remove obstacles to their path-shaping strategy and to turn new conjunctural opportunities into institutionalized supports of a new policy path. (Torfing 2001: 291)

After a thoughtful discussion of how the various institutionalisms have employed the concept of path dependency, he restates his position at an ever higher level of abstraction:

the premise for analyzing path dependency is that a dislocating event disrupts the structured coherence of a policy path and thereby creates room for an effective agency engaged in processes of disarticulation and rearticulation which are not determined by any structural necessity. The dislocation of the old policy path creates a space for inchoate decisions about how to shape future policy regulations. These decisions are taken against a background of institutional fluidity. Whereas a relative institutional unfixity is the condition of possibility for path-shaping strategies, a relative institutional fixity is the condition of possibility of path-dependent policy reform. (ibid.: 298–299)

The empirical part consists of an *illustration* of the theoretical account via a description of failed and successful attempts to reform the Danish welfare state. Torfing (2001: 306) concludes that the new institutionalisms and his own theoretical contribution “help us to explicate the mechanisms of path dependency”. However, he does not offer an empirical operationalization of the theoretical notions nor does he specify any hypotheses. He gives no conditions of refutation nor does he reflect upon the research design, the case selection, the data, the sources, and so on. As a result, he circumvents the problem of matching his highly abstract notions with hypotheses. He seems quite aware of this feature of his work, but discards it as unproblematic when he concludes that the “arguments about path dependency prove their value even in a brief and sweeping analysis of social policy reform carried out at a high level of abstraction” (ibid.: 306). Our point is that no matter how valuable his theoretical reflections are, by entirely circumventing the matching problem Torfing offers us one narrative about welfare state reform among the many

that are possible and cannot convince that his theoretical notions are of empirical relevance.

At the other extreme, we find another strategy of evading the matching problem, that is narrowly focusing on what King *et al.* (1994) have called the observable consequences of various theories. This strategy is most often applied in quantitative comparative studies. It boils down to bringing together various hypotheses present in the literature, and translating these into a single regression model in order to test these hypotheses competitively. The matching problem arises here, because no attempt is made to connect the theories to the hypotheses derived from them, even though one would need such information in the first place for specifying the hypotheses clearly and correctly. As a result, it is impossible to judge the consistency of the theories and the model that results from it. It is also impossible to see whether uncritically adopting variables from very different approaches into a single model actually makes theoretical sense for the problem (for example social policy reform) at hand. In short, in this approach, which has a long tradition in comparative public policy studies, especially social policy or welfare state studies (one early example is Wilensky 1975, a more recent one Castles 2004), researchers usually make quite an effort to explain empirical operationalizations, hypotheses, the conditions of refutation, research design, case selection, the quality of the data, sources, and so on. But their models consist of loosely connected (if at all) hypotheses that have been isolated from contrasting and even conflicting theoretical perspectives.

In this context, for example, Castles and Obinger's (2007) recent study of social spending trends sums up the main theoretical approaches to welfare state research in a single table by simply isolating the main variables, predicted outcomes, measurement procedures and data sources. They go about as follows. First, they formulate a baseline model of three variables (level of GDP per capita, average rate of economic growth, average degree of bourgeois party cabinet incumbency). Second, they include other variables from various approaches on top of the incumbency variable already taken from the power resources approach included in the baseline model. This then becomes a difficult to grasp mixture that has no real theoretical grounding. It includes the impact on social expenditure of social needs (taken from functionalist modernization theories), of the timing of welfare state consolidation (taken from path dependency theories), and of political institutions (taken from political institutionalism). Finally, they offer a best-fit model by eliminating the statistically insignificant variables.

The goal of such designs is usually to debunk widely held but empirically unsubstantiated beliefs or speculations about the impact of certain developments on the welfare state, for instance globalization (see especially Castles 2004). To the extent that the hypotheses tested are accurate matches of the theoretical intentions and to the extent that they are in fact rejected, debunking is an important result. But since no attention is paid to how theories and hypotheses are matched, it is difficult to judge whether the successful debunking effort is the result of a fair test or an artefact of the design. To give but one illustration, theories of socio-economic development or modernization are fairly complex and multifaceted accounts – rooted in the theory of structural-functional differentiation – of why and how industrialization and its social and cultural correlates, in interaction with democratization, lead to higher social spending. It is highly questionable whether

such a theory can be characterized by simple correlations or causal statements such as “economic affluence leads to more expenditure” and “problem pressure leads to more spending” or that it can be operationalized in two key variables; GDP per capita and a social need index. Such operationalizations do not match the high level of abstraction at which modernization theory is formulated.

These are extremes of a continuum between which most empirical/analytical studies of welfare state development and change could be placed. Theoretically informed thick descriptions, for instance, would be closer to the Torfing approach discussed above. But those studies that provide a combination of variable-oriented quantitative comparisons and case-oriented qualitative studies (for example Huber and Stephens 2001, Swank 2002) or approaches that make use of the Qualitative Comparative method or fuzzy sets (for example Kvist 2007, Vis 2007) would be more in line with the Castles and Obinger example. Concerning the former group of studies, the use of method triangulation may serve to alleviate the matching problem, as the problems linked to very abstract quantitative studies may be addressed in the case study part. Concerning the latter, the chances of studies using set-theoretic methods to avoid the matching problem ultimately depend on the carefully and theoretically motivated choice of how to operationalize set memberships of causal conditions. It is important to note that all of these studies face the same matching problem, but that they do make serious efforts, although not always explicitly, to deal with it.

The main advantage of the case-oriented approach over the variable-oriented one is that the former seems better capable of dealing with the matching problem. This is because thick descriptions tend to remain much closer to the conceptual framework than when proper names have been turned into variables (Przeworski and Teune 1970). Let us illustrate the point. A widely accepted insight is that political institutions of federalism have tended to obstruct the growth of social solidarity. Federalism is commonly assumed to be an institutional design to make sure that some level of national unity is preserved by decentralizing power and thus allowing a larger extent of social, cultural, economic and political diversity between subnational territorial units than would normally be tolerated in a unitary state. The welfare state’s function, by contrast, is normally argued to derive from the goal to ensure equal social rights for all citizens. Federalism and the welfare state, then, apparently have conflicting goals. The received wisdom is that federalism has considerably held back the development of social policy. As regards the politics of reform, the argument is that federalism has an impeding effect too, as the multiple veto points in federal systems make it extremely hard to overcome resistance against changing the status quo. Quantitative studies time and again reported this effect (see Obinger *et al.* 2005). But it seems questionable that such studies have matched the theory and the hypotheses adequately, because not only do comparative historical analyses indicate that under certain conditions federalism also encourages social spending, but also that the causal link between federalism and the welfare state runs both ways. In other words, the causal theory on the relationship between the two concepts becomes complicated and the demands on matching accordingly rise. The message here is that we need to become more conscious of how such problems of endogeneity, an all too common phenomenon in social science (see King *et al.* 1994), influence the problem of matching. Even in the light of lacking definite solutions, at least paying attention to such issues would be an improvement.

The best way to develop the feedback link between federalism and the welfare state is in careful studies of single country experiences. For instance, Keith Banting (2005) argues that the development of the welfare state has had a decisive effect on the centralization of the Canadian federation. Moreover, social programmes played a key role in establishing the constituent political communities of the federal state and their political identities. “Political identities are highly contested in Canada”, writes Banting (2005: 90), “and social programmes have emerged as instruments of nation-building. For the central government, social policy has been seen as an instrument of territorial integration, part of the glue holding together a vast country subject to powerful centrifugal tendencies”. But also the subnational units, most obviously Québec, used social policy to reinforce their own distinctive regional community and identity, based on culture and language. For both the federation and the regions, “social policy has been an instrument not only of social justice but also of statecraft” (ibid.). The equalization programme of the Canadian fiscal system through which resources are redistributed from the poorer to the wealthier regions are to enhance national unity. The point of the analysis is that federalism shapes social policy and social policy shapes federalism and reform of social policy reforms federalism. Here highly abstract theoretical concepts (such as territorial integration, political identity and nation-building) and their interconnections play a crucial role in the description itself. Theory and empirics are closely matched: the concepts have a descriptive function.

However, at some point researchers who prefer this type of theoretically informed descriptions of welfare state change will face the problem of generalization. If one does not try this, or if scope conditions have been formulated that exclude further generalization, one ends up with only description and little is learned about the dynamics of welfare state change. Many case studies for this reason refrain from generalization. However, if one does try to generalize the findings, then in fact a new theory is formulated that needs to be tested against new data and this posits the matching problem all over again. Yet, better to be bold in generalization and create a new challenge than to refrain from trying to improve the scope of one’s theory.

In order to further substantiate our argument, we finally present an example taken from the work of one of the current authors, focusing on ideational leadership as a cause of welfare state reform. In the welfare state literature that focuses on historical institutionalism and regime theory, strong institutional and electoral forces prevent states from conducting far-reaching, that is structural (or institutional) reforms in social policy. Then, the question arises how such reforms occur at all. The composite concept of ideational leadership can help to explain why welfare states – despite the obstacles identified – *do experience* at least some far-reaching reforms (Stiller 2004). The dependent variable is defined as far-reaching or structural reforms (measured by shifts in the financing, benefit, or management structure of a policy area), and the independent variable is ideational leadership (IL) of key policy makers. IL is characterized by four aspects: a) rejection of the status quo, b) consistent legitimization of new policy principles, c) an appeal to reform critics to rethink their resistance against reform, and d) efforts to build political coalitions for reform without resorting to tactical “games”. The claim is that the combination of these aspects leads to a substantial reduction in reform resistance, enabling the passing of structural reforms. According to the resilience literature, such an event presupposes

the overcoming of institutional and electoral obstacles. The causal mechanisms linking the separate aspects and structural reform draw on insights from the literature on ideas in policy making, institutional change and policy change. In combination, these mechanisms create an “institutional break-out” situation that reverses lock-in effects and policy stickiness, phenomena implied by path dependency. Engineering an institutional break-out enables ideational leaders to adopt structural reforms that transform the policy status quo in any one social policy area.

The problem of matching occurs in the following way. Cause and effect are situated on different levels of abstraction: on the one hand, IL and its aspects concern the level of actors, as they relate to the communicative or political behaviour of individual policy makers. The effects of such behaviour are conceptually situated at a higher level of abstraction dealing with institutional structures, relating to institutional lock-in mechanisms surrounding individual policies. Institutional structures are underlined by the values and interests of those defending policy arrangements: for instance, political parties and interest groups (and their perception of switching costs when choosing alternative institutional arrangements). The causal mechanisms linked to the different aspects of IL are: a) policy failure or loss of effectiveness brings on the search for alternatives; b) creating insights into the logics of appropriateness and necessity behind the innovation help lowering switching costs and redefining values underlying old policies; c) reform-critical interest groups are made to “face the facts” or redefine (the perception) of their interest, lowering switching costs; and d) forging consensus based more on policy-seeking than power-seeking motives reduces switching costs.

These mechanisms should ideally bridge the gap between the independent variable, IL, its operationalization in concrete patterns of argumentation and behaviour of individual policy makers, and the much more abstract concept of “institutional break-out” that enables structural reform to materialize. Conceptually, this does not immediately cause major problems, but empirically it is problematic. The main question here is twofold: 1) how to operationalize and match causal mechanisms linking an explanation on the individual actor level with an outcome that is conditional on the loosening of institutional constraints on the meso level of structure, and 2) how to find data to illustrate the working of each of the theorized causal mechanisms. For instance, which data should one use to show the effect of a policy maker criticizing the status quo and credibly promoting innovative reform on the perception and re-definition of interest by some crucial opposition party or interest group? If such data are not available, an alternative would be to use data that indirectly support (or refute) the mechanism in question. For instance, a trade union changing its preference during the reform process may be assumed to have undergone a reconsideration of interests. Another alternative could be to illustrate the effect of IL by first demonstrating its existence with the help of several qualitative indicators, and subsequently show its role in the adoption of reform by excluding alternative actor-related strategies (that is concession making) of overcoming reform obstacles. However, it would be preferable to improve the quality of the argument by illustrating the causal mechanism that links IL to the acceptance of a reform proposal. The ideal strategy in evaluating the effects of IL would be first to look for indicators of its several aspects in empirical material and establish whether these

aspects were all present in a specific case of reform. Second, suitable data could illustrate the connection between politicians' actions and the outcome of structural reform.

Empirical research on Germany has shown that demonstrating the link between the theorized causal mechanisms and a specific reform outcome often proves difficult. For instance, the case of the adoption of the unemployment insurance reform "Hartz IV" illustrates well the difficulties related to the matching problem. The central question was whether ideational leadership by the Minister of Labour Affairs contributed to the eventual adoption of this structural reform. When tracing the causal mechanisms elaborated above, the first such mechanism, concerning the link between policy makers qualifying the status quo as failure, was hard to prove. Notably, policy failure (of existing unemployment insurance arrangements) had already been acknowledged by the opposition parties before the government announced its "Hartz IV" reform plan. Regarding the effect of cognitive and normative arguments about the policy alternative on the eventual breakthrough (the second causal mechanism), it could be shown that some sceptics, especially in the governing coalition parties, had been convinced. However, it could not be directly shown that the parliamentary opposition gave up its policy position because of the government's efforts to argue for its centralistic solution for providing and administering the new benefit. Third, despite the minister's appeals to reform opponents to co-operate constructively and refrain from detrimental blockades (third causal mechanism), the appeals to give up reform resistance did not change the opposition's negative attitude so that a crucial agreement was only achieved in the parliamentary reconciliation procedure. Fourth, regarding the link between the minister's efforts of political consensus-building and the outcome of structural reform, the picture was different: there was abundant evidence of the minister actively seeking consensus with reform opponents during the agenda-setting and legislative preparatory phases. Even when, during a mediation procedure in the final negotiation phase, the minister's consensus-building efforts met institutional constraints, he remained centred on achieving a final compromise in order to avoid an outright failure of negotiations.

To sum up, the analysis focused on demonstrating the presence of indicators of IL rather than direct effects of IL on reform adoption. It did establish, by way of indirect argument, that the fourth aspect of IL contributed to the breakthrough of reform, while the effects of the remaining three indicators could not be demonstrated directly. This nicely illustrates the difficulties of finding appropriate data to demonstrate causal mechanisms, that is to bridge the gap between levels of analysis of various variables, related to the matching problem.

Conclusion

In this article, we have presented a brief overview of contemporary issues in comparative welfare state research and demonstrated that there has been real progress in our understanding of the multi-dimensionality of the dependent variable, that is, the conceptualization of welfare state change via various dimensions of reform. We also found that the debate on welfare state change and reform is not yet over. Specifically, we highlighted that the matching problem reinforces still

unresolved issues with respect to the conceptualization and operationalization of the dependent variable. Moreover, we found that – in addition to identifying causes and effects – there are difficulties with specifying the causal mechanisms that link causes and effects. Furthermore, we have tried to demonstrate, by referring to several examples from the literature using different analytical techniques, that matching theory and hypotheses is a ubiquitous problem in the literature. As a final example, we further elaborated and illustrated the argument with an empirical example from our research on structural welfare state reform, observing two methodological problems: 1) the risk of drawing conclusions about one level of analysis using evidence from another; 2) the problem of translating causal mechanisms formulated at a high level of abstraction to a lesser level.

We identified the risk of committing either an ecological or an individualist fallacy as a result of the matching problem. These fallacies occur when inferences are drawn about one level of analysis using evidence from another, for example, when on the basis of an analysis of aggregate-level data inferences are drawn about the behaviour of individuals (ecological fallacy) or vice versa (individual fallacy) (Landman 2000: 53). In addition, the problem of “measuring” causal mechanisms has to do with difficulties of “translating” them to a lesser level of abstraction which might involve finding a whole chain of mechanisms that provide together a plausible link between the two variables. However, even if such a causal link can be plausibly made theoretically, data to illustrate its workings may be difficult to obtain or non-existent.

As a first step to deal with the matching problem, we propose that researchers pay more attention to the difficulty of bridging the difference in levels of analysis between the independent and dependent variables. Furthermore, we need to address the question of how to use an actor-centred account to explain changes of structure. The second and related problem concerns the conceptualization and operationalization of the causal mechanism between these variables. There may be evidence that ideational leadership is capable of overcoming institutional resistance against change. But how can we specify precisely how this works?

We believe that questions of this sort are relevant to welfare state scholars and policy analysts more generally, deserve more attention in the welfare state literature and certainly merit further discussion. Recognizing such questions and therefore showing awareness of the matching problem and related methodological difficulties would be a first step. We hope our paper will stimulate researchers to think about potential answers and ways to deal with the matching problem so that we can improve our theoretical and empirical understanding of mechanisms of welfare state change.

Acknowledgements

An extended version of this paper was first presented in the panel “Comparative analysis of welfare reform: the dependent variable problem”, ESPAnet Conference 2005, University of Fribourg, Switzerland. We wish to thank Birgit Pfau-Effinger and Ian Bruff for their comments. We would also like to thank Christoffer Green-Pedersen and the members of the Working Group on International Relations at Radboud University for helpful criticism and suggestions. Furthermore, we are grateful to two anonymous reviewers for their constructive criticism.

Kees van Kersbergen wishes to thank the University of Konstanz's Centre of Excellence "Cultural Foundations of Integration", and especially its Institute for Advanced Study, for their generous hospitality and support.

References

- Allan, James P. and Scruggs, Lyle, 2004, Political partisanship and welfare state reform in advanced industrial societies. *American Journal of Political Science*, **48**, 496–512.
- Banting, Keith, 2005, Canada: nation-building in a federal welfare state, in: Herbert Obinger, Stephan Leibfried and Francis G. Castles (Eds), *Federalism and the Welfare State. New World and European Experiences* (Cambridge: Cambridge University Press), pp. 89–137.
- Béland, Daniel, 2005, Ideas and social policy: an institutional perspective. *Social Policy & Administration*, **39**, 1–18.
- Bonoli, Giuliano, 1998, Globalization, the welfare state, and recommodification. Round Table "Globalization and social governance in Europe and the US", Brussels.
- Bonoli, Giuliano, 2001, Political institutions, veto points, and the process of welfare state adaptation, in: Paul Pierson (Ed.) *The New Politics of the Welfare State* (Oxford: Oxford University Press), pp. 238–264.
- Breen, Richard, 1997, Risk, recommodification and stratification. *Sociology*, **31**, 473–489.
- Burgoon, Brian and Baxandall, Phineas, 2004, Three worlds of working time: the partisan and welfare politics of work hours in industrialized countries. *Politics and Society*, **32**, 439–474.
- Castles, Francis G., 2004, *The Future of the Welfare State: Crisis, Myths and Crisis Realities* (Oxford: Oxford University Press).
- Castles, Francis G. and Obinger, Herbert, 2007, Social expenditure and the politics of redistribution. *Journal of European Social Policy*, **17**, 206–222.
- Clasen, Jochen, 2005, *Reforming European Welfare States: Germany and the United Kingdom Compared* (Oxford: Oxford University Press).
- Clasen, Jochen and Clegg, Daniel, 2003, Unemployment protection and labour market reform in France and Great Britain in the 1990s: solidarity versus activation? *Journal of Social Policy*, **32**, 361–382.
- Clasen, Jochen and Siegel, Nico A., 2007, *Investigating Welfare State Change. The "Dependent Variable Problem" in Comparative Analysis* (Cheltenham: Edward Elgar).
- Esping-Andersen, Gøsta, 1990, *The Three Worlds of Welfare Capitalism* (Cambridge: Polity Press).
- Esping-Andersen, Gøsta, 1999, *Social Foundations of Postindustrial Societies* (Oxford: Oxford University Press).
- Ferrera, Maurizio and Gualmini, Elisabetta, 2004, *Rescued by Europe: Social and Labour Market Reforms in Italy from Maastricht to Berlusconi* (Amsterdam: Amsterdam University Press).
- Geddes, Mike, 1994 Public services and local economic regeneration in a post-Fordist economy. Towards a post-Fordist welfare state? in: Roger Burrows and Brian Loader (Eds) *Towards a Post-Fordist Welfare State?* (London: Routledge), pp. 151–176.
- Glatzer, Miguel and Rueschemeyer, Dietrich, (Eds) 2005, *Globalization and the Future of the Welfare State* (Pittsburgh: University of Pittsburgh Press).
- Green-Pedersen, Christoffer, 2002, *The Politics of Justification. Party Competition and Welfare-State Retrenchment in Denmark and the Netherlands from 1982 to 1998* (Amsterdam: Amsterdam University Press).
- Green-Pedersen, Christoffer, 2004, The dependent variable problem within the study of welfare-state retrenchment: defining the problem and looking for solutions. *Journal of Comparative Policy Analysis*, **6**, 3–14.
- Holden, Chris, 2003, Decommodification and the workfare state. *Political Studies Review*, **1**, 303–316.
- Huber, Evelyn and Stephens, John D., 2001, *Development and Crisis of the Welfare State: Parties and Policies in Global Markets* (Chicago: The University of Chicago Press).
- Immergut, Ellen M., Anderson, Karen M. and Schulze, Isabelle (Eds) 2007, *The Handbook of West European Pension Politics* (Oxford: Oxford University Press).
- Jaeger, Mads Meier and Kvist, Jon 2003, Pressures on state welfare in post-industrial societies: is more or less better? *Social Policy and Administration*, **37**, 555–572.

- Kangas, Olli, 2004, Institutional development of sickness cash-benefit programmes in 18 OECD countries. *Social Policy and Administration*, **38**, 190–203.
- Kemmerling, Achim, 2005, Tax mixes, welfare states and employment: tracking diverging vulnerabilities. *Journal of European Public Policy*, **12**, 1–22.
- King, Gary, Keohane, Robert O., Verba, Sidney, 1994, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton: Princeton University Press).
- Kitschelt, Herbert, 2001, Partisan competition and welfare state retrenchment: when do politicians choose unpopular policies? in: Paul Pierson (Ed.) *The New Politics of the Welfare State* (Oxford: Oxford University Press), pp. 265–304.
- Kittel, Bernhard and Winner, Hannes, 2005, How reliable is pooled analysis in political economy? The globalization–welfare state nexus revisited. *European Journal of Political Research*, **44**, 269–294.
- Korpi, Walter, 2003, Welfare-state regress in Western Europe: politics, institutions, globalization, and Europeanization. *Annual Review of Sociology*, **29**, 589–609.
- Korpi, Walter and Palme, Joachim, 2003, New politics and class politics in the context of austerity and globalization: welfare state regress in 18 countries, 1975–95. *American Political Science Review*, **97**, 425–447.
- Kühner, Stefan, 2007, Country-level comparisons of welfare state change measures: another facet of the dependent variable problem within the comparative analysis of the welfare state? *Journal of European Social Policy*, **17**, 5–18.
- Kvist, Jon, 2007, Exploring diversity: measuring welfare state change with fuzzy-set methodology in: Jochen Clasen and Nico A. Siegel (Eds) *Investigating Welfare State Change. The ‘Dependent Variable Problem’ in Comparative Analysis* (Cheltenham: Edward Elgar), pp. 198–216.
- Lamping, Wolfram and Rüb, Friedbert W., 2004, From the conservative welfare state to an “uncertain something else”: German pension politics in comparative perspective. *Policy and Politics: Studies of Local Government and its Services*, **32**, 169–192.
- Landman, Todd, 2000, *Issues and Methods in Comparative Politics: An Introduction* (London: Routledge).
- Leibfried, Stephan and Obinger, Herbert, 2003, The state of the welfare state: German social policy between macroeconomic retrenchment and microeconomic recalibration. *West European Politics*, **26**, 199–218.
- Lindsay, Colin and Mailand, Mikkel, 2004, Different routes, common directions? activation policies for young people in Denmark and the UK. *International Journal of Social Welfare*, **13**, 195–207.
- Ludwig-Mayerhofer and Wroblewski, 2004, Eppur si muove? Activation policies in Austria and Germany. *European Societies*, **6**, 485–509.
- Natali, David, 2005, Europeanization, policy arenas, and creative opportunism: the politics of welfare state reforms in Italy. *Journal of European Public Policy*, **11**, 1077–1095.
- Obinger, Herbert, Leibfried, Stephan and Castles, Francis G. (Eds), 2005, *Federalism and the Welfare State. New World and European Experiences* (Cambridge: Cambridge University Press).
- Pierson, Paul, 1994, *Dismantling the Welfare State? Reagan, Thatcher, and the Politics of Retrenchment* (Cambridge: Cambridge University Press).
- Pierson, Paul, 1996, The new politics of the welfare state. *World Politics*, **48**, 143–179.
- Pierson, Paul (Ed.) 2001a, *The New Politics of the Welfare State* (Oxford: Oxford University Press).
- Pierson, Paul, 2001b, Coping with permanent austerity: welfare state restructuring in affluent democracies. in: Paul Pierson (Ed.) *The New Politics of the Welfare State* (Oxford: Oxford University Press), pp. 410–456.
- Przeworski, Adam and Teune, Henry, 1970, *The Logic of Comparative Social Inquiry* (New York: Wiley-Interscience).
- Schmidt, Vivien A., 2000, Values and discourse in the politics of adjustment, in: Fritz W. Scharpf and Vivien A. Schmidt (Eds) *Welfare and Work in the Open Economy, Vol. 1. From Vulnerability to Competitiveness* (Oxford: Oxford University Press), pp. 229–309.
- Schmidt, Vivien A., 2005, *Public Discourse and Welfare State Reform. The Social Democratic Experience* (Amsterdam: Mets & Silt).
- Starke, Peter, 2006, The politics of welfare state retrenchment: a literature review. *Social Policy and Administration*, **40**, 104–120.
- Stiller, Sabina, 2004, Operationalizing and refining ideational leadership: towards explaining institutional change in German pension policy. Paper presented at the NIG Annual Congress, Rotterdam, The Netherlands, October 29, 2004.

- Stiller, Sabina, 2007, Innovative agents versus immovable objects. The role of ideational leadership in German welfare state reforms. PhD thesis, Radboud University Nijmegen.
- Swank, Duane, 2002, *Global Capital, Political Institutions, and Policy Change in Developed Welfare States* (Cambridge: Cambridge University Press).
- Taylor-Gooby, Peter, 2002, The silver age of the welfare state: perspectives on resilience. *Journal of Social Policy*, **31**, 597–622.
- Taylor-Gooby, Peter, Larsen, Trine and Kananen, Johannes, 2004, Market means and welfare ends: the UK welfare state experiment. *Journal of Social Policy*, **33**, 573–592.
- Torring, Jacob, 2001, Path-dependent Danish welfare reforms. The contribution of the new institutionalisms to understanding evolutionary change. *Scandinavian Political Studies*, **24**, 277–310.
- Van Oorschot, Wim, 2004, Balancing work and welfare: activation and flexicurity policies in the Netherlands 1980–2000. *International Journal of Social Welfare*, **13**, 15–27.
- Van Oorschot, Wim and Abrahamson, Peter, 2003, The Dutch and Danish miracles revisited: a critical discussion of activation policies in two small welfare states. *Social Policy and Administration*, **37**, 288–304.
- Vis, Barbara, 2007, States of welfare or states of workfare? Welfare state restructuring in 16 capitalist democracies, 1985–2002. *Policy and Politics*, **35**, 105–122.
- Wilensky, Harold L., 1975, *The Welfare State and Equality: Structural and Ideological Roots of Public Expenditures* (Berkeley: University of California Press).
- Wincott, Daniel, 2003, Slippery concepts, shifting context: (national) states and welfare in the Veit-Wilson-Atherton debate. *Social Policy and Administration*, **37**, 305–315.
- Wright, Sharon, Kopac, Anja and Slater, Gary, 2004, Continuities within paradigmatic change, activation, social policies and citizenship in the context of welfare reform in Slovenia and the UK, *European Societies*, **6**, 511–534.