Qualitative methods for the study of policy diffusion: Challenges and available solutions

Abstract

This article deals with the question whether and how processes of policy diffusion can be examined with qualitative methods. More specifically, how can qualitative methods address the 'twin challenge of interdependence', namely, the challenge to identify diffusion, on the one hand, and the challenge to discriminate between mechanisms of diffusion, on the other? I argue, *first*, that there are three distinct qualitative techniques that can be used, namely, cross-case analysis (often based on systematic case selection), within-case process tracing, and counterfactual reasoning. I demonstrate how these techniques can be adapted to the study of policy diffusion. *Second*, a combination of these methods is the best practice, since they are largely complementary in terms of the twin challenge of diffusion. The discussion draws on numerous illustrations from recent qualitative policy diffusion studies. The article closes with some suggestions for further methodological development in the study of policy diffusion, including the combination of quantitative and qualitative methods.

Keywords: policy diffusion, Galton's problem, case study methods, qualitative methods, research design

'As the social sciences attempt to grapple with ever more sophisticated forms of complexity, both statistical and qualitative methodologists will need to continue innovating so that our methodologies do not fall behind our ontologies' (Bennett and Elman 2006, 264; see also Hall 2003).

Introduction

The study of policy diffusion is the example of a research program which successfully turned a problem – namely, 'Galton's problem' of the non-independence of cases – into an opportunity – the opportunity to empirically study phenomena of causal interdependence. When studying interdependent phenomena and diffusion empirically, scholars need to address two questions: *First*, are policy changes really the outcome of a process of diffusion? *Second*, if yes, what is the mechanism underpinning this process? This paper shows that qualitative research[1] is methodologically well equipped to deal with the twin challenge of interdependence, especially when several qualitative techniques are combined.

A growing number of books and articles already systematically examine diffusion in a qualitative fashion – and do so quite successfully judged by the number of citations (from both qualitative and quantitative scholars). The list includes both purely qualitative studies of diffusion and policy transfer (Jacoby 2000; Heclo 1974; Orenstein 2008; Mossberger 2000; Resende-Santos 2007; Weyland 2010, 2009, 2006, 2008, 2005; Schmid and Götze 2009) and mixed-methods designs (Karch 2007a; Horowitz 2010). Furthermore, a huge body of qualitative historical studies also trying to go beyond national comparison, particularly works from

transnational history, transfer history and *histoire croisée*, should be mentioned (Rodgers 1998; Espagne 1999; Werner and Zimmermann 2006; Kettunen and Petersen 2011; Middell 2007).

The reasons why authors choose qualitative methods are manifold. Sometimes they simply have to do with the difficulty to measure policy change in a quantitative way, especially policy change that happened in the past or in complex policy areas. In addition, the more sophisticated quantitative methods of diffusion research become, the more demanding are their data requirements. This is why even quantitative scholars such as Berry and Berry recommend 'intensive analysis of a small number of cases, via case studies or small-sample comparative designs' when pooled data needed for up-to-date statistical models are not available (1999: 194). Yet, the use of qualitative methods can also be justified in a stronger manner, not just as a second-best solution. The literature on policy diffusion has increasingly moved beyond the problem of *whether* policies diffuse to the question as to *why* this is the case and through *what causal mechanisms* diffusion occurs (Shipan and Volden 2008; Karch 2007b). In the literature on policy diffusion, the most prominent mechanisms discussed are coercion, learning, competition, migration and emulation. It can be argued that qualitative research is better equipped to study causal mechanisms (see also George and Bennett 2005). This argument will be discussed in more depth below. Suffice to say that, as scholars of policy diffusion become more and more interested in the mechanisms of diffusion, systematic process-based qualitative methods should, if anything, grow in importance.

At the same time, there is no denying that both quantitative and qualitative methods have important strengths as well as weaknesses. This is why a combination of qualitative and quantitative approaches (Lieberman 2005) is a promising way forward, which I will briefly discuss in the conclusion. However, the main focus of this contribution is to better understand

the qualitative side. How can existing qualitative techniques be harnessed for diffusion studies? How does qualitative scholarship approach empirical cases and how are inferences drawn about diffusion processes?

The particularities of qualitative diffusion research have so far not been acknowledged. Especially compared to the growing literature on spatial econometrics (e.g. Beck, Gleditsch, and Beardsley 2006; Franzese and Hays 2008; Plümper and Neumayer 2010), the existing methodological literature, as will be shown, offers little guidance for qualitative researchers wishing to study policy diffusion. When it comes to interdependence and diffusion, the literature can be divided into four groups. The *first* group completely ignores the issue (e.g. Geddes 2003; Ragin 2008). The *second* does acknowledge it as a methodological problem but fails to offer qualitative solutions or generally doubts that qualitative methods can successfully address it (Goldthorpe 1997; Przeworski and Teune 1970; Gerring 2001). The *third* claims to offer solutions which consist in factoring out interdependence, for example, by selecting geographically or temporally distant or unconnected cases (Anckar 2007), or discounting for its effect by trying 'not [to] treat these data as providing as many observations as we would have obtained from independent observations' (King, Keohane, and Verba 1994, 222). By definition, these solutions are not suitable to studying interdependence as an *object* of research. Finally, the *fourth* group of authors claims that qualitative researchers have ways to deal with interdependence of cases and even study processes of diffusion. Unfortunately, these ways often remain mysterious. Consider the following quote from Rueschemeyer and Stephens (1997, 62):

> The two most critical advantages of comparative historical research […], the ability to study the local context of critical characteristics and the knowledge of relevant historical sequences, are very efficient tools for settling the questions raised by Galton's problem.

What exactly this means in practice remains unclear.[2] George and Bennett are somewhat more precise in that they advocate process tracing for the study of interdependent phenomena (2005: 33). But still, while their overall treatment of process tracing methods has been rightfully influential, their account of policy diffusion is brief and incomplete. Like Rueschemeyer and Stephens, they claim that process tracing is particularly well-suited for diffusion analysis, but do not explain in more detail how it can be applied to this task.

The present article builds on the last group of authors but aims to make the particular methods used much more explicit. I focus on the most important methods for diffusion research and discuss their specific advantages and blind spots. To illustrate the methodological points I am making, examples are drawn from a selection of four recent book-length studies of policy diffusion, (see below on how the studies were selected). However, the article is not intended to be a lengthy book review, but rather an applied methods paper with specific practical lessons.

I argue that qualitative researchers have found ways to successfully address phenomena of interdependence and that these ways are clearly different from quantitative approaches. The main lesson from this paper is that a *combination of cross-case and within-case qualitative methods* – which tend to address different aspects of interdependence – works particularly well. More specifically, in order to draw inferences, qualitative diffusion studies should use a combination of cross-case comparison based on systematic case selection, within-case process tracing and, in some cases, counterfactual analysis. I will show how these different methods complement each other in terms of their strengths and weaknesses, which is why their combination represents a powerful research strategy for analyzing the existence as well as the mechanisms of policy diffusion.

The paper is structured as follows: I first define a number of central concepts and describe in more detail what the specific methodological challenge of the interdependence of cases – the basic condition of phenomena of diffusion – consists in and briefly sketch how quantitative research typically addresses this challenge. Then, I critically discuss three qualitative techniques used in diffusion research – cross-case comparison, process tracing and counterfactual analysis. The final section sums up their strengths and weaknesses and concludes.

The twin challenge of interdependence

Interdependence generally refers to a fundamental causal relationship in which the outcome or a change in the outcome of a case depends (at least to some degree) on the outcomes of one or several other cases of the same type – and *vice versa* (Elkins and Simmons 2005). It must be differentiated from the somewhat narrower concept of diffusion. The diffusion of policies across jurisdictions (especially nation states and subnational entities) is the most prominent application of diffusion theory in political science (pioneer studies include Walker 1969; Berry and Berry 1990; for reviews, see Berry and Berry 1999; Dobbin, Simmons, and Garrett 2007).

Diffusion can be broadly defined as 'one government's choices being influenced by the choices of other governments' (Shipan and Volden forthcoming: 1). Note that policy diffusion and interdependence are not the same, in that diffusion should rather be understood as a *consequence* of interdependence. Interdependence, due, for example, to economic interconnectedness, can thus be understood as a necessary, but insufficient condition for policy diffusion. Across disciplines, the concept of diffusion has been applied to various subjects, including the technological innovation and social movements (Rogers 1995; Givan, Soule, and

Roberts 2010). This article focuses on the diffusion of public policies, with national and subnational governments (e.g. states, local governments) being the typical units of analysis. Its methodological conclusions, however, should be equally relevant for other substantive fields.

This paper focuses on what has be called 'horizontal diffusion' (Shipan and Volden 2008), that is, mutual influence of units located at the same level of analysis. Cases may also respond to a 'common shock' while staying mutually independent units. While such 'vertical' processes – e.g. the impact of an international organization or 'world society' (Meyer et al. 1997) on nation states – are often discussed in the context of policy diffusion, they must be analytically distinguished and call for distinct research designs and methods. Furthermore, it is important to note that diffusion is not the same as policy convergence. Convergence may or may not be the end result. It is, for example, possible that some cases (e.g. countries) react in a positive way to a policy change in another country – by copying a policy – and some in a negative way – by countering innovation or through negative learning, that is, learning from 'bad examples', and draw negative lessons (for whatever reason). In this case policy learning is universal, yet policies may still diverge. That is why the phenomena of policy convergence and diffusion must be kept analytically distinct. Diffusion is a causal process, while convergence and divergence are possible descriptions of an outcome pattern.

The methodological challenge of interdependence – for both qualitative and quantitative research – is twofold: on the one hand, scholars need to *identify a diffusion effect* and, on the other, they need to *discriminate between different mechanisms* of diffusion. In essence, both aspects have to do with assessing specific alternative explanations for the empirical pattern observed. Regarding the first challenge, one needs to empirically demonstrate a diffusion effect against the alternative explanations of a) independent causation and b) a common shock.

Regarding the second challenge, finding ways to empirically distinguishing between the various alternative mechanisms of diffusion suggested by theory is the key task. I follow Gerring's theoretically broad definition of causal mechanism as 'the pathway or process by which an effect is produced or a purpose is accomplished' (Gerring 2007: 178).

The *first* challenge for diffusion studies is thus to judge whether what looks like an effect of diffusion across cases – e.g. a 'wave' of policy changes – did not simply come about either due to independent domestic or internal causes, a common external shock or pure chance. In other words, it is important to judge whether the empirical spatio-temporal clustering is not one of 'spurious diffusion' (Braun and Gilardi 2006, 299, 305-6).

Quantitative scholars commonly respond to this challenge by introducing the outcomes of other cases – the 'spatial lag' for each observation – on the right-hand side of a regression model in addition to a range of variables that control for domestic conditions and common exogenous shocks (Simmons, Dobbin, and Garrett 2007; Franzese and Hays 2008).[3] How exactly this is specified in the model varies (see below). At a minimum, studies of this type can demonstrate whether or not there is temporal and/or spatial clustering of outcomes, even when controlling for internal conditions and external influences. Yet, while such clustering strongly suggests the presence of diffusion processes, it alone cannot explain how exactly diffusion occurs.

The *second* challenge, therefore, is to go beyond the description of empirical clustering and demonstrate how – through which mechanisms – diffusion takes place. Over the years a variety of different causal mechanisms of diffusion have been formulated and spelled out in great detail, notably mechanisms of coercion, learning, competition, migration and emulation (Gilardi 2012; Braun and Gilardi 2006; Elkins and Simmons 2005; Franzese and Hays 2008). These are

the processes or pathways that unfold during processes of diffusion, linking cause (i.e. policy change in another case) and effect (i.e. policy change in the case under review).

Quantitative diffusion research usually studies the theoretical mechanisms of diffusion in an indirect way, through specific conditional effects stemming from the type and extent of interlinkages between cases. This can be done through so-called 'weighted spatial lags' in a regression model, that is, by introducing a variable measuring the outcomes of other cases – the above-mentioned 'spatial lag' for each observation – weighted by the degree and type of connectivity between cases (Franzese and Hays 2008). Finding empirical indicators of connectivity that reflect a particular channel of diffusion (and *only* this channel) is far from straightforward, however. For example, in order to tap into the theoretical mechanism of competition, Swank (2006) weighs spatial lags with bilateral flows of trade and foreign direct investment (FDI) between countries as well as the correlation of their overall patterns in FDI flows – which reflects the similarity in the profile of trade and investment between different national economies and therefore the likelihood that countries are direct competitors. Shipan and Volden, in an article on local antismoking legislation in the United States (2008), create an alternative competition variable based on the potential negative economic effects from adopting legislation. The variable includes information about whether there are neighboring or proximate cities (within a 10 mile radius) *without* antismoking legislation and additionally weighs this information by the size of these cities. Their competition variable is specifically designed to capture the potential economic 'outflow' to other cities if antismoking legislation is introduced.

Quantitative measurements are usually based on the strength of *potential* channels of certain types of diffusion rather than the actual processes of learning, emulation etc. In other words, they involve numerous assumptions. In the examples above, completion is inferred on the

basis of geographic proximity in one case (Shipan and Volden) and on similarity in trade patterns on the other (Swank). Inferring the mechanism of competition from such conditional effects therefore depends on the plausibility of these assumptions.[4] This is not to say that quantitative diffusion research is ineffective, I am rather arguing that there are specific limitations in the analysis of causal mechanisms of policy diffusion inherent in the methods currently in use.

Qualitative studies, by and large, follow an alternative approach to studying the causal mechanisms of diffusion. Here, the empirical strategy, as will be shown in more detail below, is often to look not at conditional effects but at empirical processes as evidence for the unfolding of causal mechanisms over time. Arguably, notwithstanding well-known problems of generalizability of small-N methods, by studying empirical processes rather than effects, the qualitative strategy comes closer to what is meant by causal mechanisms in the methodological literature.

To sum up, there are two main challenges empirical work on diffusion needs to address, namely, identifying diffusion and discriminating between mechanisms of diffusion. Quantitative research attempts to deal with the twin challenge by explicitly modeling spatial lags in combination with statistical controls and by choosing weighting matrices derived from theoretical mechanisms of diffusion. Unfortunately, neither solution can be applied to qualitative research, mainly because qualitative methods are not based on a large number of cases and a logic of simultaneous control of many explanatory factors. Yet, as will be shown in the next section in detail, qualitative scholars have found genuine solutions for this issue. Generally, most qualitative diffusion studies combine cross-case analysis of covariation on the level of independent and dependent variable ($X \rightarrow Y$) with detailed within-case analysis, often focusing on the mechanisms linking X and Y.

Qualitative solutions

The illustrations in the following sections come from four qualitative studies of policy diffusion (of which two also employ statistical methods): Michael C. Horowitz's book *The Diffusion of Military Power* (2010), Andrew Karch's *Democratic Laboratories* (2007a), Karen Mossberger's *The Politics of Ideas and the Spread of Enterprise Zones*, and Kurt Weyland's *Bounded Rationality and Policy Diffusion* (2006). These examples were chosen on the basis of several considerations. First, they are all monographs, not journal articles or edited volumes. This allows a sufficiently detailed analysis of diffusion processes and outcomes. Qualitative journal articles, by contrast, are often simplified versions of larger studies and, arguably, constrained by the word limits required by most journals. Second, these books are 'best practice examples' which, admittedly, is a rather subjective category. In any case, they are all well-cited and/or award-winning works.[5] Third, they cover a wide range of theoretical approaches and policy areas in jurisdictions at different political levels. Horowitz's book is an international study of the determinants of military innovation from the invention of battlefleet warfare in the early twentieth century until today's suicide terrorism. The topic of Karch's study is policy diffusion at the U.S. state level in various social policy program areas (e.g. prescription drug subsidies, welfare programs). Mossberger examines the spread of the policy innovation of 'enterprise zones' – that is, the application of special tax and regulation policies in depressed inner city areas in order to revitalize private investment – across U.S. states. Weyland's study, by contrast, compares national policies in Latin America, more specifically, pension and health reforms, and the way in which they were influenced both by reforms in other countries and prescriptions from

international actors such as the World Bank and the IMF. Both Karch and Horowitz also employ quantitative techniques. Yet the main weight of their arguments is carried by qualitative methods.

While the wide variety of topics, theories and cases was an important reason for choosing these four books as illustrations, all four authors make similar methodological choices. As I will show, all, in one way or another, choose a combination of cross-case comparison, process tracing at the within-case level and, to a more limited extent, counterfactual thought experiments. Of course, these qualitative techniques are well established outside the realm of diffusion research (Gerring 2006; George and Bennett 2005; Lebow 2010). Yet, the examples show that, when combined, they can be used to meet the specific challenges of studying diffusion. In the following sections, I will develop this argument in more detail.

Cross-case analysis

The first way in which qualitative studies analyze policy diffusion is through cross-case analysis, that is, the systematic investigation of qualitative similarities and differences of values on theoretically relevant variables across several cases. Cross-case analysis can either be applied to the whole universe of cases or a carefully selected subset of this group. Often, broad associations are first described for the whole group before zooming in on a smaller selection of cases. In contrast to mainstream quantitative methods (Brady and Collier 2010), however, cases must be selected in a non-random fashion. Cross-case comparisons – whether among the whole group or just a subset of cases – are made through what has been described as ‘intuitive regression’ (Collier, Mahoney, and Seawright 2004, 94), in other words, non-standardized nominal or

ordinal comparison of cases on a few central theoretical dimensions in order to assess competing explanatory claims (Mahoney 1999). Associations between theoretically important characteristics are noted in a non-numerical fashion. Inferences are often quite basic, relying mostly on a series of interconnected bivariate correlations, which is why cross-case comparison is usually not a stand-alone technique.

When policy diffusion is analyzed for the whole universe of cases, the first step is to derive implications about the empirical *patterns of spatio-temporal clustering* of policy changes that one can expect to find on the basis of a given theory of diffusion. The empirical patterns can then be described in verbal accounts or on the basis of basic quantitative measures. The most common way to graphically depict cross-national diffusion patterns is with a diffusion curve, a simple graph of the cumulative number of positive cases along a time axis. Another way is a map that focuses on the geographical aspects of clustering.

All four books selected for this article take such broad comparisons as a starting point. In his discussion of the spread of pension privatization in Latin America, Kurt Weyland uses the diffusion curve, combined with a verbal account of the geographical pattern of diffusion, to draw inferences. His analysis reveals that the overall pattern of diffusion of Chilean-style pension privatization in Latin America is characterized by three traits that are typical for policy diffusion processes: first, the temporal S-shaped pattern of adoption (see Rogers 1995); second, geographical clustering (i.e. contagion among neighboring countries); third, convergence of outcomes over time. On the basis of this description, he then goes on to evaluate three diffusion hypotheses, namely external pressure, normative appeal and boundedly rational learning. Weyland states that the ‘external pressure argument [i.e. pressure stemming from international organizations such as the World Bank] cannot easily account for the geographic clustering of

diffusion, especially its pronounced neighborhood effects' (2006, 37). Moreover, 'given modern means of instant communication, the normative appeal framework [i.e. policy emulation] would expect an even faster diffusion of innovations than is captured in the slow initial upswing of the S-shaped pattern. If the quest for legitimacy drives policy choice, the appearance of a novelty should immediately trigger emulation' (Weyland 2006, 40-1). From this cross-case exercise Weyland therefore infers that it supports neither the external pressure argument nor the policy emulation hypothesis. Moreover, the different features of the overall pattern suggest that boundedly rational learning may be driving the diffusion of pension reform in Latin America since they closely correspond to the empirical implications of central 'cognitive heuristics' of bounded rationality (i.e. 'availability', 'representativeness' and 'anchoring'). This, then, becomes the main hypothesis for his process tracing analysis (see below).

In her book on the diffusion of enterprise zones, Karen Mossberger uses not only a bar chart tracking the number of states that newly adopt the enterprise zone policy each year but also draws a map depicting the geographical distribution of policy adopters across the fifty states (2000, 83-4). The temporal pattern of state-level adoptions suggests that '[f]ederal incentives clearly stimulated states' interest in the program, for adoptions usually peaked in years when the possibilities for federal legislation seemed brightest' (2000, 82). Cross-case examination is thus, again, used as a first test of a potential causal explanation which is then made subject to more intensive scrutiny.

Qualitative diffusion scholars do not stop here, however. In addition to inferences from the basic spatio-temporal pattern of policy change, cross-case comparison draws inferences from controlled contrasts between a small number of carefully chosen cases (often a subset of the cases included in the basic diffusion curve or map). The aim is to make up for the restricted

degrees of freedom by choosing particularly relevant cases, from the point of view of the theory. The way in which cross-case inferences for this smaller group are made, however, is not essentially different from what has been described for the whole group of relevant cases. Inferences are also based on simple nominal and ordinal comparisons between case characteristics in the manner of 'intuitive regression', but for a more restricted range of cases.

As would be expected, the qualitative studies surveyed in this paper all select cases in a non-random fashion (for an overview, see Seawright and Gerring 2008).[6] It is important to note, however, that qualitative diffusion scholars cannot use Mill's methods of agreement and difference (also known as the 'most similar cases design') as these designs do not allow for interdependence of cases. The older 'small-N' literature in the social sciences, associated with the names of Mill, Przeworski and Teune, Lijphart and Skocpol, is ill-equipped to deal with interdependence. It focused heavily on cross-case inference based on rigid case selection procedures as a method of causal explanation (Lijphart 1971; Mill 1974 [1872]; Przeworski and Teune 1970; Skocpol 1979). Some of the methods advocated by these authors, such as Mill's 'method of difference' (aka the most similar cases design) and 'method of agreement' have become subject to severe methodological criticism over the course of the years (e.g. Lieberson 1991). In theory, Mill's methods could be applied to policy diffusion in a very limited way. Following the method of difference, for example, one could select cases with a positive outcome – e.g. the adoption/transfer of particular policy reform – as well as negative cases – non-adopters, in this example – and try to see whether positive and negative cases differ only on one relevant independent variable – ideally, a conditional variable for diffusion – while being similar on all other (potentially relevant) dimensions. The independent variables – such as cultural similarity, trade openness – would have to be based on different theoretical diffusion mechanisms.

Conversely, when using the method of agreement in diffusion research, two or several adopters would have to be found that differ on all potentially relevant dimensions bar one, which in turn must be a variable that can be linked to one particular channel of diffusion in order to explain the outcome. It is easy to see that these research designs are exceedingly restrictive and unrealistic. One of the most important restriction is the problem of limited empirical diversity. Especially when countries are the units of selection, it is unlikely that fitting cases can be found – not to mention the various other problems associated with Mill's methods.

Instead, the much less restrictive 'diverse cases' design can be used as a basis of cross-case comparisons (Seawright and Gerring 2008). All qualitative diffusion studies reviewed here apply this design. In contrast to Mill's 'method of difference' (Mill 1974 [1872]) or Przeworski and Teune's 'most similar cases design' (1970), the 'diverse cases design' allows several – not just one – potentially relevant independent factors to vary across the selected cases. Nor does the approach follow a strict experimental logic but it is more concerned with having enough variation on several independent variables in order to test or develop several alternative propositions.

One of the key strengths of the 'diverse cases' strategy is that it can be systematically linked to several alternative theoretical mechanisms of diffusion. It allows to contrast cases with high or low values on crucial variables and thereby assess different hypotheses by comparing theoretical predictions with empirical reality under different constellations of theoretical factors or explore new hypotheses by unearthing contrasting outcomes in different settings. Moreover, the design seeks to avoid studying a biased set of cases but rather aims at picking cases that represent the universe of cases, at least approximately (Seawright and Gerring 2008).[7] In diffusion studies, cases are typically chosen that exhibit a large degree of divergence on central

conditioning factors of diffusion such as geographical location or factors that influence the propensity of policy learning such as administrative professionalism.

In her book on the spread of enterprise zones, Karen Mossberger (2000) selects five U.S. states – Indiana, Massachusetts, Michigan, New York and Virginia – on the basis of the diverse cases approach. She states (2000, 214) that

> [t]he number of cases and particular states were selected to provide a sufficient chance for variation on some factors that had a relationship to relevant decision making or diffusion theories. These variables are timing of adoption, region, professionalization, political factors, and program design.

In a similar manner, in his study of state-level diffusion of policy innovation, Andrew Karch selects three diverse cases for in-depth analysis: Massachusetts, Oregon and Virginia, that is, states that 'are located in diverse regions of the country, and […] vary along the demographic and socioeconomic dimensions highlighted by previous diffusion research. Their scores along such dimensions as innovativeness, income, partisanship, and ideology diverge significantly' (Karch 2007a, 23). While he claims that this selection resembles the 'most different systems design' of Przeworski and Teune, he also points out the shortcomings of a mechanical application of that method. Therefore, as Karch puts it, 'it is necessary to illustrate how specific causal mechanisms contribute to the outcomes under review' (2007a, 24), in other words, to employ process tracing in the cases chosen (see below).

Michael Horowitz also selects his cases in a highly systematic and explicit way to reflect the possible range on the central independent variables of his theory of the diffusion of military innovation. In contrast to the other studies mentioned, however, 'the unit of analysis is the innovation itself' (2010, 60), not a country or subnational entity. After listing the universe of cases of all major military innovations since the nineteenth century that are relatively

uncontroversial in the literature, he picks four: early twentieth century battlefleet warfare, carrier warfare, nuclear weapons and suicide terrorism (2010, 61). These innovations reflect diversity in terms of the two crucial independent variables of his 'adoption-capacity theory', that is, the required *financial intensity* of a military innovation – high in the case of carrier and nuclear warfare, medium or low in the other two cases – and the level of *organizational capital* required for adoption – high in the case of carrier warfare and suicide terrorism and medium or low in the other two cases (see table in Horowitz 2010, 63). While none of the innovations chosen has low requirements in terms of finances or organizational capacities, battlefleet warfare is described as having 'medium' requirements on both counts (see Horowitz 2010, 63-4, for a detailed explanation of this coding decision). Two of the cases – battlefleet and carrier warfare – are examined with qualitative methods, while the other two are made subject to statistical analysis.[8]

Obviously, cross-case methods based on systematic case selection is suitable to address the second challenge of interdependence – the challenge to discriminate between mechanisms of diffusion – as long as the selection criteria are strictly derived from diffusion theories. They are much less powerful in terms of the first challenge of judging whether the outcome is really one of policy diffusion and did not simply come about independently. It is virtually impossible in qualitative cross-case analysis that use a kind of 'intuitive regression' to control for confounding factors the way statistical multiple regression does. In the case of diffusion, looking at the spatio-temporal pattern of adoptions alone cannot determine whether and to what extent domestic factors also shape the pattern of change and whether what looks like a process of diffusion is, in fact, an illusion stemming from similar but independent causation. Even the 'diverse cases' method only very loosely 'controls' for domestic causes in the sense that evidence of similar policies in very dissimilar cases – e.g. in terms of socio-economic profile or size – suggests that

domestic differences did not matter much for policy changes. However, in several analyses, both positive and negative cases of adoption are studied.[9] In this situation, the cross-case differences in domestic circumstances must be addressed more directly, which is difficult, if not impossible in the manner of 'intuitive regression'. In any case, once several domestic factors and common shocks enter the picture, qualitative covariational methods reach their limits.

To sum up, cross-case analysis can serve different aims in the study of diffusion, from a simple match of primary outcomes with expectations to sophisticated tests of various sub-hypotheses and implications with different kinds of data and different means of presenting that data. In diffusion research, a simple diffusion curve is often the starting point of more in-depth analyses and the overall pattern of diffusion in terms of timing or geography can be an important guide as to the value of various hypotheses, though not a decisive test. It comes as no surprise, then, that these methods are usually combined with more in-depth process tracing.

Process tracing

All four diffusion studies reviewed rely not only on cross-case analysis but at least as much on within-case methods, namely process tracing. Process tracing is understood by George and Bennett as an attempt to 'identify the intervening causal process – the causal chain and causal mechanism – between an independent variable (or variables) and the outcome of the dependent variable' (2005: 206). This is done by looking at the empirical within-case implications of (alternative) causal mechanisms linking one or several independent variables with an outcome (see also Bennett 2008; Collier et al. 2010; Mahoney 2010), which implies close attention to issues of timing, sequencing and causal conjunction (Grzymala-Busse 2011). As Bennett has

pointed out, in contrast to qualitative methods of cross-case comparison, process tracing studies weigh their observations in a 'folk-Bayesian' manner (Bennett 2008). Not all pieces of evidence are equally valuable for causal inference, they need to be judged against the researcher's (or the literature's) prior expectations and their potential to change these theoretical expectations.

How can process tracing help to address the twin challenge of diffusion research? Regarding the first challenge, instead of trying to identify processes of diffusion through strategies of variation and control, notably spatial regression, qualitative researchers trace the spread of policies and ideas through historical reconstruction of decision-making at the actor level. They often try to find at 'smoking gun' evidence (Van Evera 1997, 30-4) of transnational encounters, migration and exchange of influential actors and learning across jurisdictions. 'Smoking gun' evidence can provide strong positive corroboration of a theory (see also Bennett 2008, 2010). An excellent example is Sharman's evidence showing that policymakers sometimes even copy not just the legislative texts but even the errors contained in those texts due to copy-and-paste 'learning' or 'policy plagiarism' (Sharman 2010, 625). Finding such a 'smoking gun' is not always easy, though. The origins and spread of policies and ideas are frequently murky, policies are adapted and renamed as they travel to other countries, influences are deliberately hidden by policymakers who wish to retain the full credit for reforms and so on. Yet, when it is found, it provides strong evidence that policy diffusion is non-spurious.

Kurt Weyland extensively uses process tracing in order to demonstrate that Latin American countries took clues from other countries in designing social security reforms (see Weyland 2006, 12-14, 226-8 on why he uses process tracing). In chapter 4, he traces how the model of Chile's pension privatization spread to other countries through various channels – among others, through personal contacts between policymakers. He reconstructs these contacts with data from a

large number of expert interviews. For example, he reports how the Chilean-style model entered the agenda in Bolivia (Weyland 2006, 101):

> [T]he Finance Ministry's budget director, Helga Salinas, attended the annual conference of the business peak association, Confederación de Empresarios Privados de Bolivia (CEPB). The keynote speaker was the architect of Chile's reform, José Piñera (interview with Cuevas 2002). Widely credited with great persuasive powers, Piñera advertised his innovation with missionary zeal. Salinas was so captivated that she took Piñera aside and kept talking to him over lunch. She depicts this encounter as an eye-opening experience, a crucial turning point in her ongoing efforts to combat pension problems, and the starting point of the privatization project, which drew on very extensive advice and consultations with Chilean experts during the following years (interviews with Salinas 2002 and Bonadona 2002).

Although similar exchanges at the level of policymakers and ministry officials can be found for Peru and El Salvador, the evidence is seldom as clear-cut as in this 'smoking-gun' example. The process tracing analysis nonetheless shows how the reform process in these countries evolved after these initial exchanges and led to a copying of the Chilean blueprint to the exclusion of alternative models (Weyland 2006, chapter 4). In Brazil and Costa Rica, by contrast, the Chilean reform was also studied but had to compete with less radical alternative reform proposals which were eventually implemented. Moreover, Weyland demonstrates that the World Bank and the International Monetary Fund had surprisingly little influence on policy changes in these countries. Again, process tracing is used, for example, by examining the sequence of events. In most cases, these organizations 'did not initiate this wave of change but only supported it once it was already under way' (2006, 95). By tracing the evolution of the reform debates in five countries, Weyland therefore shows that diffusion of ideas had real, i.e. non-spurious, effects in some countries and no effect in others. He shows that the similarity in policy solutions was neither a coincidence nor simply the result of similar domestic conditions or international pressure but can be clearly traced to the Chilean reform.

What about the second challenge, then, the explanation of diffusion? In the methods literature, process tracing is portrayed as *the* method that deals with explanatory causal mechanisms. Checkel, for example, argues that 'if one is going to invoke the philosophy-of-science language of mechanisms, then process tracing is the logically necessary method for exploring them' (2006, 366). In consequence, process tracing should be particularly well-suited to deal with the second challenge of interdependence. George and Bennett indeed claim that process tracing can be used to examine 'the kinds of detailed sequences in learning and diffusion processes that can create relationships between cases, allowing researchers to gauge more accurately how much of the variance in outcomes is explained by learning or diffusion and how much is explained by other variables' (2005, 33-4). Yet, there are reasons for pessimism. Lee and Strang (2006, 889), after discussing a qualitative case study on the transfer of British and American public sector reform to Australia and South Korea (Halligan 1996):

> While providing evidence that policy diffusion is a concrete reality and not merely an academic conceit, process tracing studies have important limitations. They tell one little about the underlying network structure of influence. For example, Australia might have attended to the United Kingdom because of its colonial heritage, close economic and political ties, economic competition, or a host of other linkages; we cannot discriminate between these alternatives by knowing more about the Raynor scrutinies [i.e. a British policy program established in 1979].

In other words, process tracing performs well in terms of the first challenge, but may fare poorly when it comes to the second challenge. Qualitative researchers do not have privileged access to causal mechanisms. In many situations mechanisms may even be unobservable. This is, for instance, true for the cognitive processes so prominent in policy learning theories. Yet process tracing works on the basis of the assumption that, while the mechanisms themselves may be unobservable, their operation leaves behind empirical 'traces' in the form of – often scattered

and unsystematic – empirical data. Sometimes, and contrary to what Lee and Strang suggest, finding traces of theoretical mechanisms can be successful. Yet the success of process tracing in discriminating between various mechanisms of diffusion crucially depends on two things: the specification of the mechanism and the availability of the necessary evidence – preferably 'smoking gun' evidence. In the manner of a detective – to cite an often-used metaphor – scholars combine diverse pieces of evidence and weigh their importance with respect to different hypotheses and the underlying causal mechanisms (Bennett 2008). And like detectives, case study researchers may not directly observe the crime but still find the person who committed it.

In chapter 3 of *The Diffusion of Military Innovation*, for example, Michael Horowitz traces the spread of aircraft carrier warfare – not just the spread of the technology itself but of the use of carrier warfare, including support ships, changing doctrine etc. – over time. Only three countries – Britain, the U.S., and Japan – have used carrier warfare to date (although more navies do possess aircraft carriers). On the basis of an initial cross-case analysis of theoretical predictions (Horowitz 2010, 76-84), he reviews the historical record of the decision-making processes since the first aircraft carriers were used in World War I. Horowitz demonstrates that the empirical decision-making 'stories' about the adoption and non-adoption closely follow sequences that could be expected according to his 'adoption-capacity theory' of military innovation. Three countries are studied in depth: Great Britain, the Soviet Union/Russia and Italy. Whereas the former, at least initially, had the financial and organizational means to fully adopt carrier warfare, the Soviet Union and Italy never managed to implement the innovation (although they built some carrier ships). The Soviet Union simply did not have the financial resources required to build both the carriers and the fleet of supporting ships needed to engage in carrier warfare. Nor did its military organization have the capacity to learn how to organize large

carrier operations (2010, 92). The case study of the Soviet Union also confirms another prediction made by Horowitz's theory, namely that even the countries that are not able to fully adopt a military innovation will react in other ways, for example, by 'countering' the innovation. The Soviet Union did exactly this 'by investing heavily in land-based naval aviation, antiship missiles, and submarines' (2010, 92).

One clear danger of focusing too much on tracing the mechanisms of diffusion is to neglect the first challenge of interdependence, that is, demonstrating that the change in outcome under consideration was really caused by outcomes in other cases, and did not simply come about independently. Horowitz's study is a good example for this tendency. He spends little time on the interlinkages between countries and more on the domestic organizational and financial resources needed for adopting innovation. He does *not* show how the technology transfer and learning process across countries really came about. How did countries acquire the knowledge to even try to innovate? Did knowledge diffuse via migration of experts, common defense networks, military intelligence, experience in battle or other channels? In this sense, the study appears more like a study of parallel adoption than of diffusion.

Counterfactual reasoning

Counterfactual thought experiments – that is, theory-guided reasoning about what might have been – are an example of a more controversial qualitative technique. This is perhaps the reason why it is also the most marginal for the diffusion studies under consideration. There is no space for a comprehensive discussion of the question whether counterfactuals are a good or a bad thing

or perhaps even an 'unscientific' method.[10] They may simply be to some extent 'unavoidable' (Levy 2008, 628) in the social sciences (see also Tetlock and Belkin 1996b; Lebow 2010). In any case, the current discussion is much more about the 'how' of counterfactual reasoning. What makes a counterfactual credible and useful? Authors have emphasized several attributes of high-quality counterfactuals, including 'clarity of the antecedents and consequents, the plausibility of the antecedent, and the conditional probability of the consequent given the antecedent' (Levy 2008, 633) or Tetlock and Belkin's various criteria of good counterfactual reasoning (Tetlock and Belkin 1996a, 16-31).

What have counterfactuals to offer for diffusion research? Qualitative diffusion research could use what-if scenarios to strengthen inferences about the non-spuriousness of diffusion. ('If the neighboring country x had not introduced policy y three years earlier, policy y would probably not have entered the agenda in country z.') At least potentially, counterfactuals could help to address both challenges, the challenge of identifying diffusion – by specifying what would have happened without the stimulus from other cases – as well as the challenge of explaining diffusion – by trying to specify what the empirical pattern or sequence of events would have looked like in the absence of a causal mechanism or causal condition. ('If rational policy learning were the causal mechanism at work, policy innovation would have been much more carefully introduced, with attention to local context.')

Surprisingly, the qualitative diffusion studies reviewed make only limited use of explicit counterfactual reasoning. One of the few examples in the four books studied is from Horowitz's study on military innovation. He employs counterfactual reasoning when highlighting the importance of the Soviet Union's lack of financial resources for the failure to adopt carrier warfare as opposed to the end of the Cold War. He argues that while the Soviet Union had again

started to construct carriers in the 1980s, it was not just due to the collapse in 1990 that the Soviet, and later Russian, navy did not adopt the innovation. It did not have the resources required. Even in the absence of the collapse in 1990, therefore, the Soviet Union would not have fully adopted carrier warfare (2010, 92):

> Even if all of its developmental projects had succeeded and the cold war had not ended until the late 1990s, in the mid-1990s the Soviet Navy would have fielded two carriers capable of launching fixed-wing aircraft. While this would have been an incredible engineering and systems integration accomplishment for the Soviet Navy, the resources required to complete the *Kuznetsov* and *Varyag* would have made it difficult to purchase the destroyers, cruisers, and logistical ships necessary to conduct carrier task force operations.

Although counterfactual reasoning is not extensively used in an explicit manner in the four books which form the basis of this article, it figures more *implicitly* in some of the case studies. For example, in Chapter 3 of his book, Karch considers the role of national policy activity on the diffusion of policy ideas at the level of U.S. states, notably at the agenda-setting stage. He looks in depth at various episodes in three states (Massachusetts, Oregon and Virginia). Despite large differences between these three states (in terms of demographics, partisanship etc.), similar legislative bills in five policy fields were introduced almost simultaneously, 'an overlap [which] is puzzling in light of current theorizing about policy diffusion' (Karch 2007a: 69). Introduction of these bills came after discussions at the federal level had become more salient. In other words, the implicit counterfactual here is that the three states would not have introduced such innovations, at least not simultaneously. Empirically, Karch first analyzes national debates and media coverage of a program and, in a second step 'compare[s] them to trends in state policy-making', going back several years, to see whether federal attention preceded state-level agenda-setting (2007a: 73). In other words, the counterfactual is that the long-term state-level

policy trends preceding national agenda-setting would have more or less continued in the absence of heightened national attention. Even though the reasoning is not couched in the 'what-if' terminology of counterfactuals, the logic is by and large the same. Inferences are drawn not primarily from the cross-case policy variation between the three states – they are all positive cases – but rather from the difference to a counterfactual situation – i.e. a situation without national attention to these policy issues. However, it may have been advantageous to make the counterfactual structure of this empirical strategy more explicit. Nonetheless, the example of Karch's analysis of top-down influences on state-level agenda setting highlights the usefulness of counterfactual reasoning for qualitative diffusion research. The counterfactual design allowed him to study the diffusion effect in a relatively simple manner, even in a situation where all relevant cases (i.e. the 50 U.S. states) are subject to the same potential cause (i.e. national activity), that is, where there is no useful cross-case variation on the central independent variable.

Conclusion

Policy diffusion has been one of the growing fields of both theoretical and empirical work in political science and public policy in recent years. In this paper, I have argued that there is little reason to expect that, in a situation presumably marked by increasing interdependence between cases, qualitative research is bound to become irrelevant. There is, in my view, no reason that qualitative methods should fall behind the new ontologies of the social sciences, to paraphrase the quote from the beginning. On the contrary, there is an emerging set of powerful applications of qualitative techniques for studying interdependent phenomena. Yet, the most powerful

research designs are those that do not rely on just one tool but those that combine different qualitative techniques. All four books reviewed here support this conclusion. The reason for this combinatory strategy clearly lies in the specific strengths and weaknesses of the different qualitative tools. It turns out that they are highly complementary, especially when it comes to the twin challenge of interdependence (see Table 1).

Table 1 about here

The typical strengths and weaknesses of qualitative techniques for diffusion research can be summed up as follows. *Cross-case analysis* based on systematic case selection is well-suited to address the theoretical mechanisms of diffusions. Yet, it does not directly address the problem of clearly identifying diffusion, i.e. the first challenge. And while it is useful for judging the plausibility of alternative explanatory mechanisms, it does not help in sorting out the problem of spurious diffusion as qualitative cross-case methods cannot simultaneously control for confounding domestic or international factors.[11] *Process tracing* is potentially better suited to address the twin challenge of interdependence, especially if the mechanisms of diffusion are well specified and if 'smoking gun' evidence can be found. These conditions, however, are not always easily fulfilled, especially with respect to theoretical diffusion mechanisms, i.e. the second challenge. This is the reason why there are relatively few successful empirical applications of this. However, I believe that there is a huge potential here, given that cross-case methods (including spatial regression) struggle with this problem of identifying the real channels of diffusion. For example, the distinction between the mechanism of imitation/emulation and policy learning is not well understood and hard to measure empirically. Qualitative research

should try to dig into this distinction, for instance, by tracing a change in beliefs and preferences of central decision-makers over time due to learning. It is, however, not necessary that policy-making actors are aware of the theoretical mechanisms themselves. The specific empirical implications of different causal mechanisms can be at the level of routine behavior of actors rather than explicit reasoning, and actors do not have to be aware of processes of learning or competition themselves when in an interview situation. One important general disadvantage of process-tracing is, of course, that it is usually only possible to trace diffusion mechanisms for a very small number of cases. Finally, while *counterfactual analysis* has been portrayed as a method which addresses questions of causality head-on, and may therefore be useful for both identifying non-spurious diffusion and explaining it, qualitative diffusion researchers have not made extensive use of this technique, at least not in an explicit manner. The example of Karch's analysis has shown, however, that counterfactual reasoning can be employed in a systematic manner to analyze diffusion effects. There may be great potential to analyze the mechanisms of diffusion, too.

The main conclusion from this discussion is that it is important to continue combining cross-case analysis, within-case process tracing and counterfactuals as these methods are highly complementary for studying policy diffusion.

In addition to combining different qualitative procedures, however, a combination of quantitative and qualitative methods in one research design holds great promise (see Lieberman 2005). While this was not the main question of the present paper, there are a number of characteristics of qualitative diffusion studies that distinguish them from most quantitative research designs. There is, for example, an inherent complementarity in approaching causal mechanisms of diffusion. While quantitative studies look at conditional effects of various proxy

variables for the different diffusion mechanisms, qualitative researchers use process tracing to study the unfolding of diffusion processes within and between different units of analysis. The strength of the quantitative approach is the ability to map very specific patterns of diffusion – e.g. distinctive geographical patterns – across a large number of cases. However, the empirical models are usually based on various untested assumptions. And while the strength of process tracing lies in the temporal reconstruction of diffusion processes, one important downside is the limited generalizability due to the small number of cases. There is a huge potential here for mixed-methods designs in policy diffusion research. Continuing the dialogue across methodological and disciplinary divides will certainly accelerate progress in this fascinating research program.

Table 1: Strengths and weaknesses of different qualitative techniques for diffusion analysis

	Diffusion effects	Mechanisms of diffusion
Cross-case analysis	-	+
Process tracing	+	(+)
Counterfactual analysis	(+)	(+)

Note: Own depiction. + = well suited; - = ill-suited; in parentheses = few empirical applications.

References

Anckar, Dag. 2007. "Selecting Cases in Cross-National Political Research." *International Journal of Social Research Methodology* 10 (1): 49-61.

Beck, Nathaniel, Kristian Skrede Gleditsch, and Kyle Beardsley. 2006. "Space Is More than Geography: Using Spatial Econometrics in the Study of Political Economy." *International Studies Quarterly* 50 (1): 27-44.

Bennett, Andrew. 2008. "Process Tracing: A Bayesian Perspective." In *The Oxford Handbook of Political Methodology*, ed. J. M. Box-Steffensmeier, H. E. Brady and D. Collier. Oxford: Oxford University Press, 702-21.

———. 2010. "Process Tracing and Causal Inference." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards, Second Edition*, ed. H. E. Brady and D. Collier. Lanham, MD: Rowman & Littlefield, 207-19.

Bennett, Andrew, and Colin Elman. 2006. "Complex Causal Relations and Case Study Methods: The Example of Path Dependence." *Political Analysis* 14 (3): 250-67.

Berry, Frances Stokes, and William D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84 (2): 395-415.

———. 1999. "Innovation and Diffusion Models in Policy Research." In *Theories of the Policy Process*, ed. P. A. Sabatier. Boulder,CO: Westview Press, 169-200.

Brady, Henry E., and David Collier, eds. 2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards, Second Edition*. Lanham, MD: Rowman & Littlefield.

Braun, Dietmar, and Fabrizio Gilardi. 2006. "Taking 'Galton's Problem' Seriously: Towards a Theory of Policy Diffusion." *Journal of Theoretical Politics* 18 (3): 298-322.

Checkel, Jeffrey T. 2006. "Tracing Causal Mechanisms." *International Studies Review* 8 (2): 362-70.

Collier, David, Henry E. Brady, and Jason Seawright. 2010. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology." In *Rethinking Social Inquiry, Second Edition*, ed. H. E. Brady and D. Collier. Lanham, MD: Rowman & Littlefield, 161-99.

Collier, David, James Mahoney, and Jason Seawright. 2004. "Claiming Too Much: Warnings about Selection Bias." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, ed. H. E. Brady and D. Collier. Lanham, MD: Rowman & Littlefield, 85-102.

Dobbin, F., B. Simmons, and G. Garrett. 2007. "The global diffusion of public policies: Social construction, coercion, competition, or learning?" *Annual Review of Sociology* 33: 449-72.

Elkins, Z., and B. Simmons. 2005. "On Waves, Clusters, and Diffusion: A Conceptual Framework." *Annals of the American Academy of Political and Social Science* (598): 33-51.

Espagne, Michel. 1999. *Les transferts culturels franco-allemands*. Paris: Presses Universitaires de France.

Franzese, Robert J., and Jude C. Hays. 2008. "Interdependence in Comparative Politics: Substance, Theory, Empirics, Substance." *Comparative Political Studies* 41 (4-5): 742-80.

Geddes, Barbara. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor, MI: University of Michigan Press.

George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Gerring, John. 2001. *Social Science Methodology: A Criterial Framework*. Cambridge: Cambridge University Press.

———. 2006. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.

———. 2007. "The Mechanismic Worldview: Thinking Inside the Box." *British Journal of Political Science* 38 (1): 161-79.

Gilardi, Fabrizio. 2012. "Transnational Diffusion: Norms, Ideas, and Policies." In *Handbook of International Relations*, ed. W. Carlsnaes, T. Risse and B. A. Simmons. London: Sage.

Givan, Rebecca Kolins, Sarah A. Soule, and Kenneth M. Roberts, eds. 2010. *The Diffusion of Social Movements: Actors, Mechanisms and Politicals Effects*. Cambridge: Cambridge University Press.

Goldthorpe, John H. 1997. "Current Issues in Comparative Macrosociology: A Debate on Methodological Issues." *Comparative Social Research* 16 (1): 1-26.

Grzymala-Busse, Anna. 2011. "Time Will Tell? Temporality and the Analysis of Causal Mechanisms and Processes." *Comparative Political Studies* 44 (9): 1267-97.

Hall, Peter A. 2003. "Aligning Ontology and Methodology in Comparative Research." In *Comparative Historical Analysis in the Social Sciences*, ed. J. Mahoney and D. Rueschemeyer. Cambridge: Cambridge University Press, 373-404.

Halligan, John. 1996. "The Diffusion of Civil Service Reform." In *Civil Service Systems in Comparative Perspective*, ed. H. A. G. M. Bekke, J. L. Perry and T. A. J. Toonen. Bloomington: Indiana University Press, 288-317.

Hanson, Stephen E., and Jeffrey S. Kopstein. 2005. "Regime Type and Diffusion in Comparative Politics Methodology." *Canadian Journal of Political Science* 38 (1): 69-99.

Heclo, Hugh. 1974. *Modern Social Politics in Britain and Sweden*. New Haven: Yale University Press.

Horowitz, Michael C. 2010. *The Diffusion of Military Power: Causes and Consequences for International Politics*. Princeton, NJ: Princeton University Press.

Jacoby, Wade. 2000. *Imitation and Politics: Redesigning Modern Germany*. Ithaca, NY: Cornell University Press.

Karch, Andrew. 2007a. *Democratic Laboratories: Policy Diffusion among the American States*. Ann Arbor, MI: University of Michigan Press.

———. 2007b. "Emerging Issues and Future Directions in State Policy Diffusion Research." *State Politics & Policy Quarterly* 7 (1): 54-80.

Kettunen, Pauli, and Klaus Petersen, eds. 2011. *Beyond Welfare State Models: Transnational Historical Perspectives on Social Policy*. Cheltenham: Edward Elgar.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Lebow, Richard N. 2010. *Forbidden Fruit: Counterfactuals and International Relations*. Princeton, NJ: Princeton University Press.

Lee, Chang Kil, and David Strang. 2006. "The International Diffusion of Public-Sector Downsizing: Network Emulation and Theory-Driven Learning." *International Organization* 60 (4): 883-909.

Levy, Jack S. 2008. "Counterfactuals and Case Studies." In *The Oxford Handbook of Political Methodology*, ed. J. M. Box-Steffensmeier, H. E. Brady and D. Collier. Oxford: Oxford University Press, 627-44.

Lieberman, Evan S. 2005. "Nested analysis as a mixed-method strategy for comparative research." *American Political Science Review* 99 (3): 435-52.

Lieberson, Stanley. 1991. "Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." *Social Forces* 70 (2): 307-20.

Lijphart, Arend. 1971. "Comparative Politics and the Comparative Method." *American Political Science Review* 65 (3): 682-93.

———. 1975. "The Comparable-Cases Strategy in Comparative Research." *Comparative Political Studies* 8 (2): 158-77.

Mahoney, James. 1999. "Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis." *American Journal of Sociology* 104 (4): 1154-96.

———. 2010. "After KKV: The New Methodology of Qualitative Research." *World politics* 62 (1): 120-47.

Marsh, David, and Jason C. Sharman. 2009. "Policy Diffusion and Policy Transfer." *Policy Studies* 30 (3): 269-88.

Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103 (1): 144-81.

Middell, Matthias, ed. 2007. *Dimensionen der Kultur- und Gesellschaftsgeschichte: Festschrift für Hannes Siegrist zum 60. Geburtstag*. Leipzig: Leipziger Universitätsverlag.

Mill, John S. 1974 [1872]. *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. 8th ed. Toronto/London University of Toronto Press and Routledge & Kegan Paul.

Mossberger, Karen. 2000. *The Politics of Ideas and the Spread of Enterprise Zones*. Washington, DC: Georgetown University Press.

Orenstein, Mitchell A. 2008. *Privatizing Pensions: The Transnational Campaign for Social Security Reform*. Princeton, NJ: Princeton University Press.

Plümper, Thomas, and Eric Neumayer. 2010. "Model Specification in the Analysis of Spatial Dependence." *European Journal of Political Research* 49 (3): 418-42.

Plümper, Thomas, Vera Troeger, and Eric Neumayer. 2010. "Case Selection and Causal Inference in Qualitative Research." unpublished manuscript.

Przeworski, Adam, and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: Wiley-Interscience.

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.

Resende-Santos, João. 2007. *Neorealism, States, and the Modern Mass Army*. Cambridge: Cambridge University Press.

Rodgers, Daniel T. 1998. *Atlantic Crossings: Social Politics in a Progressive Age*. Cambridge, MA: Harvard University Press.

Rogers, Everett M. 1995. *Diffusion of Innovations, 4th edition*. New York: Free Press.

Rueschemeyer, Dietrich, and John D. Stephens. 1997. "Comparing Historical Sequences -- A Powerful Tool for Causal Analysis." *Comparative Social Research* 16 (1): 55-72.

Schmid, Achim, and Ralf Götze. 2009. "Cross-National Policy Learning in Health System Reform: The Case of Diagnosis Related Groups." *International Social Security Review* 62 (4): 21-40.

Seawright, Jason, and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61 (2): 294-308.

Sharman, J. C. 2010. "Dysfunctional Policy Transfer in National Tax Blacklists." *Governance* 23 (4): 623-39.

Shipan, Charles R., and Craig Volden. 2008. "The Mechanisms of Policy Diffusion." *American Journal of Political Science* 52 (4): 840-57.

———. forthcoming. "Policy Diffusion: Seven Lessons for Scholars and Practitioners." *Public Administration Review*.

Simmons, Beth A., Frank Dobbin, and Geoffrey Garrett, eds. 2007. *The Global Diffusion of Markets and Democracy*. Cambridge: Cambridge University Press.

Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.

Swank, Duane. 2006. "Tax Policy in an Era of Internationalization: Explaining the Spread of Neoliberalism." *International Organization* 60 (4): 847-82.

Tetlock, Philip, and Aaron Belkin. 1996a. "Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives." In *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*, ed. P. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press, 3-38.

———, eds. 1996b. *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*. Princeton, NJ: Princeton University Press.

Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press.

Walker, Jack L. 1969. "The Diffusion of Innovations among the American States." *American Political Science Review* 63 (3): 880-99.

Werner, Michael, and Bénédicte Zimmermann. 2006. "Beyond Comparison: Histoire Croisée and the Challenges of Reflexivity." *History and Theory* 45 (1): 30-50.

Weyland, Kurt. 2005. "Theories of Policy Diffusion: Lessons from Latin American Pension Reform." *World politics* 57 (2): 262-95.

———. 2006. *Bounded Rationality and Policy Diffusion: Social Sector Reform in Latin America*. Princeton, NJ: Princeton University Press.

———. 2008. "Bounded Rationality in Latin-American Pension Reform." In *Lessons from Pension Reform in the Americas*, ed. S. J. Kay and T. Sinha. Oxford Oxford University Press.

———. 2009. "The Diffusion of Revolution: '1848' in Europe and Latin America." *International Organization* 63 (3): 391-423.

———. 2010. "The Diffusion of Regime Contention in European Democratization, 1830-1940." *Comparative Political Studies* 43 (8-9): 1148-76.

Notes

[1] Qualitative research is understood as the intensive study of a relatively small number of cases which does not employ formal statistical techniques (on the multiple dimensions of the qualitative-quantitative distinction, see Collier, Brady, and Seawright 2010: 177-182).

[2] Similar optimistic statements can be found elsewhere (Lijphart 1975, 171; Anckar 2007, 52; Hanson and Kopstein 2005, 93).

[3] The goal here is not to give a review of the various available quantitative methods of diffusion research. I limit myself to a relatively brief treatment of some particularities of what is currently the most prominent approach, namely spatial regression models, in order to be able to explain what is special about qualitative methods of diffusion research.

[4] Shipan and Volden acknowledge this difficulty when they write: '[A]lthough our main independent variables are designed to isolate key aspects of different diffusion mechanisms, they […] may capture aspects of other mechanisms as well' (2008: 851).

[5] In 2011, Horowitz received both the Lasswell Prize by the Society of Policy Scientists and the Best Book Award by the International Security Studies Section of the International Studies Association for *The Diffusion of Military Power*. Weyland's book was runner-up in the Twelfth Annual Robert W. Hamilton Book Awards in 2008. By November 2012, according to Google Scholar, Horowitz had already 24 citations, Karch 71, Mossberger 78, and Weyland 142.

[6] Here, I focus on those procedures that are the basis for particular cross-case comparisons, in contrast to within-case analyses of a single or a small number of cases that are also often based on considerations of systematic case selection – e.g. the deviant/typical case or the most-likely/least-likely case study.

[7] The diverse cases method seems to do so quite successfully, according to statistical simulations (Plümper, Troeger, and Neumayer 2010).

[8] The country selection Horowitz makes for the case studies is based on both positive and negative cases of adoption, but, interestingly, the selection criteria for the countries are much less explicit than those for the objects of diffusion.

[9] The possibility of picking negative cases is, in fact, a clear strength of the diverse case method since many studies of transfer and diffusion may suffer from inferential biases due to an overemphasis of positive cases (Marsh and Sharman 2009, 281).

[10] I am also sidestepping the question whether all causal reasoning is based on a counterfactual logic and focus on counterfactual reasoning as a methodological tool, not an epistemological foundation of causal analysis.

[11] Set-theoretic methods such as fuzzy-set Qualitative Comparative Analysis (fsQCA) could, in theory at least, be adapted to include a spatial lag, similar to quantitative methods. However, neither are the methodological consequences spelled out nor are there any empirical applications of QCA to policy diffusion.