© (2008) Swiss Political Science Review 14(3): 413–50

Empirical Modeling of Policy Diffusion in Federal States: The Dyadic Approach

Fabrizio Gilardi and Katharina Füglister
University of Zurich and University of Lausanne

Policy diffusion is a common phenomenon in federal states: indeed, one of the normative justifications of decentralized policy making is that it permits the development and spread of best practices. Following Berry and Berry (1990), event-history analysis has been the method of choice for the quantitative investigation of policy diffusion, but Volden (2006) has recently introduced a dyadic variant of this method in which units of analysis are not states but, instead, pairs of states. This article discusses the dyadic approach with a particular focus on the diffusion of policies in Switzerland. The goal is not to introduce a new method, but rather to provide a practical overview for researchers interested in using it. The article shows how the method has migrated from the international relations literature to the policy-diffusion literature, describes the typical structure of a dyadic dataset in a diffusion context, and discusses several modeling issues. The usefulness of the dyadic approach is illustrated empirically with the example of health-insurance subsidy policies in Swiss cantons.

Keywords: Policy Diffusion • Federalism • Health Care • Event-history Analysis • Dyadic Approach

Introduction[1]

Policy diffusion has been a classic topic in the literature on federalism, which political scientists – especially in the United States – have studied for a long time (McVoy 1940; Walker 1969; Gray 1973; Berry and Berry 1990; Mintrom 1997). The normative starting point is the idea that autono-

[1] Previous versions of this article were presented at the 2006 and 2007 Annual Congress of the Swiss Political Science Association, and at the 2007 Annual Meeting of the Midwest Political Science Association. We thank the participants in these conferences, Isabelle Stadelmann-Steffen, Katerina Linos, and two anonymous referees for helpful feedback. The financial support of the Swiss National Science Foundation (grants no. 100012–105763/1 and PA001–115307/1) is also gratefully acknowledged.

mous subnational units work as policy laboratories in which new policies can be developed, tested, and, if they are successful, spread to the whole country. Following this view, federalism promotes policy learning.

Initially, the literature focused on geographic proximities to explain policy diffusion, but recently, more sophisticated explanations have been developed. Volden (2006), for instance, has shown that successful policies are more likely to spread, which is consistent with learning arguments. Another recent development is the idea to study diffusion not only between units at the same level (such as states), but also across levels (cities, states, country) (Shipan and Volden 2006). In addition, a growing literature has examined how policies diffuse internationally in a wide range of domains, including economic policies (Elkins, Guzman, and Simmons 2006; Simmons and Elkins 2004; Meseguer 2006a; Meseguer 2004; Swank 2006), regulatory policies (Jordana and Levi-Faur 2005; Levi-Faur 2005; Gilardi 2005), and social policies (Brooks 2005; Brooks 2007; Jahn 2006; Franzese and Hays 2006; Weyland 2007; Gilardi, Füglister, and Luyet 2009).[2]

The argument that policies spread across subnational units is obviously highly relevant for Switzerland, although, due to the limited size of the Swiss political science community, no Swiss literature on the subject has developed. A few studies exist, however. Schaltegger (2004), for instance, found that cantonal fiscal policies are influenced by the practices of neighbors, while Kübler and Widmer (2007) concluded that cantonal implementations of a federal drug program diffused both regionally and country wide.

The diffusion literature is essentially quantitative (for an exception, see Weyland 2007), and the method of choice has been event-history analysis, which is a technique specifically focused on the longitudinal study of the occurrence of events (Box-Steffensmeier and Jones 2004). Following Berry and Berry (1990), scholars have conceptualized the dependent variable as the adoption of a policy (the "event") and have used event-history tools to investigate its determinants. In these studies, units of analysis are state-years (or country-years, city-years, etc.): each state is tracked over time until the policy is adopted (or not, in which case the observation is said to be right censored). If the dependent variable is continuous, we have a typical time-series cross-section dataset, in which units of analysis are, of course, also state-years.

[2] For overviews of this literature, see Simmons, Dobbin, and Garrett (2006), Braun and Gilardi (2006), Braun, Gilardi, Füglister, and Luyet (2007), Meseguer and Gilardi (2008).

The goal of this article is to offer a practical guide to a new version of these methods recently put forward by Volden (2006), namely, the dyadic approach. In this approach, units of analysis are not state-years, but are instead dyad-years in which a dyad is a pair of states. While dyadic studies are widespread in the international relations literature (where dependent variables are often relational – war, for instance), Volden (2006) had the insight to use the dyadic approach to study diffusion. As we will see, this approach allows researchers to conveniently model various hypotheses on diffusion processes. Each state is, in turn, allowed to be the potential "receiver" and "sender" of a policy, and independent variables can measure the characteristics of both "receivers" and "senders", as well as their relationships. With this setup, many indicators of diffusion mechanisms can be directly included in the analysis. On the other hand, the dyadic approach requires the redefinition the dependent variable from the simple adoption of a policy to some form of "increased similarity" between states in the dyad.

We strongly emphasize that the goal of this article is *not* to develop a new method or to improve existing ones. More modestly, we aim to present a practical guide to the dyadic approach for researchers interested in using it. As we will see, dyadic methods are not entirely straightforward, and there are many subtleties researchers need to be aware of to use them appropriately. Therefore, this overview will be useful for scholars wishing to study diffusion with quantitative tools. Finally, let us note that policy diffusion is a highly relevant topic for the study of public policies in Switzerland, but the methods presented here can be applied to any setting in which diffusion hypotheses can plausibly be developed. Our main focus here is, however, on Switzerland.

The article is structured in two main parts. In the first, we present the dyadic approach, and we discuss several methodological issues, including the construction of a dyadic dataset and the definition of the dependent variable, standard event-history analysis issues, dependence structure, cross-sectional heterogeneity, and potential "emulation biases". The second part offers an empirical illustration based on an original dataset of health-insurance subsidy policies in Swiss cantons. In the conclusion, we sum up the main points, and we discuss their relevance. Finally, an appendix presents Stata code for the construction of a dyadic dataset.

The Dyadic Approach

From International Relations to Policy Diffusion

The dyadic approach has been widely used in the international relations literature, where often the dependent variable does not measure attributes of countries but rather of pairs of countries. For instance, when scholars investigate the "democratic peace" hypothesis, which states that democracies do not fight each other, the dependent variable is whether two countries are at war in a given year, or whether a given country initiates a conflict against another country (see, for example, Maoz and Russett 1993; Farber and Gowa 1997; Russett, Oneal, and Davis 1998; Hensel, Goertz, and Diehl 2000; Peceny, Beer, and Sanchez-Terry 2002; Leeds 2003; Danilovic and Clare 2007). Similarly, the relevant units of observations are dyads also when researchers study international trade flows and who trades with whom (Morrow, Siverson, and Tabares 1998), or why countries sign bilateral investment treaties (Elkins, Guzman, and Simmons 2006). For these research questions, the dependent variable is inherently dyadic because it measures the relationship between states rather than their characteristics, and taking pairs of countries as observations is a natural choice. Models have the following general form:

$$y_{ijt} = \alpha + X_{ijt}\beta + \epsilon_{ijt}, \tag{1}$$

where i, j, and t indexes are, respectively, "receivers", "senders", and time, y_{ijt} is a vector of relational outcomes, X_{ijt} is a matrix of measures for the characteristics of the dyad, and β is a vector of coefficients to be estimated.

For example, Gartzke (2007) examines the economic aspects of the democratic peace, and argues that democratic pairs are less prone to conflict because they tend to be economically more developed and to have more open financial markets. The dependent variable is coded 1 when a militarized dispute begins between the two countries in the dyads and 0 otherwise, and the analysis includes the following independent variables:

- "Democracy (low)" and "Democracy (high)", which measure (through various indicators) the lower and higher democracy scores in the dyad;
- "Financial openness (low)" and "Trade dependence (low)", which report the lower scores in the dyad for indicators measuring, re-

spectively, restrictions on foreign exchange, current-and capital accounts, and the ratio of trade over GDP;

- "GDP per capita (low)", which measures the lower GDP per capita in the dyad;
- "Affinity", which reports the similarity of interests in the dyad, measured through voting patterns in the UN general assembly; and
- A series of controls, such as geographic contiguity and distance, major power status (coded 1 if at least one state in the dyad is a major power), military alliances (1 if the two states in the dyad are allies), capability ratio, and region.

We see that all of these variables are measured at the level of the dyad, and the statistical model thus corresponds to equation (1). Indeed, Gartzke's (2007) study, like many others, adopts a "nondirected" approach (Bennett and Stam 2000), meaning that each dyad appears in the dataset only once, and no distinction is made between "initiators" and "targets" of conflict. Thus, the United States–Iraq dyad is equivalent to the Iraq–United States dyad. However, we also see that a workaround is employed to include monadic variables (that is, variables that are measured at the level of the state rather than at that of the dyad), namely, the distinction between "low" and "high" values in the dyad. While this distinction does not permit identification of the relevant country in the dyad (we do not know which one has "low" or "high" values), it does incorporate information that is not strictly dyadic.

Some authors have argued that an explicitly "directed" approach is often necessary, as in many cases theoretical hypotheses do not just specify which dyads are more likely to be at conflict, but also reveal which country initiates it (Bennett and Stam 2000; Reiter and Stam 2003). Directed models have the following form:

$$y_{ijt} = \alpha + X_{ijt}\beta + V_{it}\gamma + W_{jt}\delta + \epsilon_{ijt}, \tag{2}$$

where y_{ijt} is a vector of relational outcomes, V_{it} is a matrix of measures for the characteristics of the first state in the dyad, W_{jt} is a matrix of measures for the characteristics of the second state in the dyad, and β, γ, and δ are vectors of coefficients to be estimated. The main difference with respect to the nondirected approach is that variables do not only measure dyadic attributes, but also the monadic features of the two states that are part of the dyad. On the other hand, the dependent variable remains dyadic: it measures an observable relationship between the two countries. An example of directed analysis is Bennett (2006), which investigates how political similarity, rather than joint democracy, promotes peace. The dependent variable measures whether a state initiates a conflict against another state, and the independent variables capture attributes of the potential "initiator" and of the potential "target", notably democracy and power, and characteristics of the dyad such as geographic proximity and alliances. The model thus follows the general specification shown in equation (2), which is adopted also in several other studies (for example Horowitz, McDermott, and Stam 2005; Bussmann and Oneal 2007; Danilovic and Clare 2007).

In sum, dyads are a natural level of analysis for many questions in the international relations literature. By contrast, focusing on dyads when studying policy diffusion is less obvious. In the policy-diffusion literature, there is an explicit interest in relationships between states, but this is estimated rather than observed. While we can measure whether a country initiated an armed conflict against another country in a given year, the influence of a state on another state's policy choices is unobservable and has to be estimated. The advantage of the dyadic approach is that observable relationships of theoretical interest, such as geographic proximity or similarities in socio-economic structures, can be included easily into the analysis. In a directed approach, the first state in the dyad can be identified as the potential "receiver" of a policy and the second as the potential "sender", and their attributes can be taken into account easily. This insight is due to Volden (2006), who used a directed dyadic setup to study how states in the United States influence one another in the development of their Children's Health Insurance Programs. More precisely, Volden investigated the conditions under which a state makes its policy more similar to that of another state, and found that this is more likely when the two states share political, demographic, and budgetary similarities. The main finding is, however, that "successful" states are more likely to be imitated, which suggests that decentralized policy making encourages a quite rational form of policy learning.

The definition of the dependent variable in dyadic diffusion studies is different both from standard state-year diffusion analyses and from dyadic analyses in international relations. In the former, the dependent variable is 1 when a state adopts a policy and 0 otherwise; in the latter, it is 1 when the dyadic relationship of interest, such as militarized conflict or the signature of a bilateral trade agreement, is observed. In dyadic diffusion studies, by contrast, observations are dyad-years, but the observable phenomenon, policy change, is at the state-year level. Therefore, the dependent variable of the analysis must be constructed indirectly. Volden (2006) codes it 1 when the potential "receiver" makes its policy more similar to that of the "sender". As a result, the dependent variable does *not* record policy change, nor the influence of one state over another, but simply increased similarity in the policies of two states. Since the focus is not on the bilateral relations among the two states but on the general diffusion process, the goal is then to detect systematic patterns in increased similarities, and to estimate the influence of various factors on the decisions of states to introduce policy changes that move them closer to other states. This permits one to make inferences about the underlying diffusion process, which, however, remains unobserved.

Dataset Structure and Dependent Variable

The structure of a dyadic dataset is not completely obvious, especially in a diffusion context. In this section we will first show how a dyadic dataset is constructed, and we will then discuss some issues linked to the definition of the dependent variable. Table 1 shows an excerpt of a hypothetical dyadic dataset. Since most of the time the raw data are in a state-year format, its construction is not entirely trivial. Stata code is presented in Appendix 1 and can be easily adapted to other datasets.

Each row lists a dyad-year, which is the unit of analysis. The first column indicates the first canton in the dyad, that is, the potential "receiver" of a policy, while the second column shows the second canton, the potential "sender". Note that the dataset is directed: each dyad appears twice, which allows each canton to be in turn a potential "sender" and "receiver". The third column is simply the year, and columns 4–6 report identifiers for the first and second canton, as well as for the dyad (their use will be explained later). The seventh column lists the dependent variable, which we call "imitation". As discussed earlier, in contrast to the international relations literature, the dependent variable here is unobservable. We cannot observe whether a canton imitates another canton: indeed, estimating this influence is the main goal of the analysis. What we can observe is policy change, but this is measured at the level of cantons, not of dyads. The solution found by Volden (2006) is to code the dependent variable 1 if, at time t, State$_i$ adopts a policy that State$_j$ already had at time t–1. The logic appears clearly if we look a the various dyads in Table 1. For the Geneva–Zurich dyad, the potential "receiver" (Geneva) does not change its policy during the observation period, as shown in the "Policy$_i$" column. The potential

Table 1: Hypothetical directed dyadic dataset (excerpt)

$Canton_i$	$Canton_j$	Year	ID_i	ID_j	ID_{dyad}	Imit.	$Policy_i$	$Policy_j$	Samelang.	Pop_i	Pop_j
Geneva	Zurich	2000	1	2	1	0	2	1	0	409’900	1’222’200
Geneva	Zurich	2001	1	2	1	0	2	1	0	416’400	1’242’000
Geneva	Zurich	2002	1	2	1	0	2	2	0	422’200	1’257’900
Geneva	Bern	2000	1	3	2	0	2	3	0	409’900	950’600
Geneva	Bern	2001	1	3	2	0	2	2	0	416’400	952’900
Geneva	Bern	2002	1	3	2	0	2	2	0	422’200	955’400
Geneva	Schwyz	2000	1	4	3	0	2	1	0	409’900	128’400
Geneva	Schwyz	2001	1	4	3	0	2	1	0	416’400	129’600
Geneva	Schwyz	2002	1	4	3	0	2	1	0	422’200	131’700
Zurich	Geneva	2000	2	1	4	0	1	2	0	1’222’200	409’900
Zurich	Geneva	2001	2	1	4	0	1	2	0	1’242’000	416’400
Zurich	Geneva	2002	2	1	4	1	2	2	0	1’257’900	422’200
Zurich	Bern	2000	2	3	5	0	1	3	1	1’222’200	950’600
Zurich	Bern	2001	2	3	5	0	1	2	1	1’242’000	952’900
Zurich	Bern	2002	2	3	5	1	2	2	1	1’257’900	955’400
Zurich	Schwyz	2000	2	4	6	0	1	1	1	1’222’200	128’400
Zurich	Schwyz	2001	2	4	6	0	1	1	1	1’242’000	129’600
Zurich	Schwyz	2002	2	4	6	0	2	1	1	1’257’900	131’700

"sender" does change its policy (see "Policy$_j$" column), but this is irrelevant. "Imitation" is thus 0 for all years. By contrast, in the Zurich–Geneva dyad, the potential "receiver" (now Zurich) changes its policy from 1 to 2 in 2002. Since the potential "sender" had policy 1 in the previous year, "Imitation" is now coded 1: Zurich has adopted a policy that Geneva already had. The Zurich–Schywz dyad conforms this logic: Zurich changes its policy in 2002, but not in the same direction as Schwyz. Therefore, the dependent variable is coded 0 even though there is a policy change. We see that the dependent variable does not measure policy change nor does it influence, but it only shows the potential influence. The goal of the analysis is to detect systematic patterns in policy changes that move some cantons closer to some other cantons.

Table 1 includes three more variables (columns 10–12). Their theoretical interest is relatively limited, and they are included just as an illustration of the fact that independent variables can measure attributes of the dyad ("Same language"), of the potential "receiver" ("Population$_i$"), and of the potential "sender" ("Population$_j$").

Finally, an important issue is how the dependent variable can be constructed with more complex data structures. In Table 1, the policy of each state is measured by a single variable, and coding the dependent variable is straightforward, but in some cases, policies can have several dimensions. Volden (2006), for example, identifies six different components of Children's Health Insurance Programs in U.S. states, namely, whether the program is part of Medicaid or not, eligibility thresholds, benefits levels, waiting period before enrollment, and presence of copayments. Volden (2006) codes the dependent variable 1 if State$_i$ adopts at time t a policy that State$_j$ already had at time t–1. Unlike in the fictitious example that we have just discussed, however, this coding leads to loss of information, since it is possible that State$_i$ moves closer to State$_j$ on one policy dimension, but at the same time moves further from it on other dimensions. To address this issue, as a robustness check Volden (2006) conducts a factor analysis to reduce the dimensionality of the policy space and then codes the dependent variable 1 if the Euclidean distance between the two states in the dyad decreases. Multidimensional scaling techniques can also be used to compute distances in a multidimensional policy space.

Although throughout the discussion we have assumed that policies are measured through categorical variables (which is the case in many diffusion analyses and most dyadic studies), a continuous measure can also be employed. In this case, one could code the dependent variable 1 if State$_i$

moves closer to State$_j$ (that is, if $|\text{policy}_{At} - \text{policy}_{Bt-1}| < |\text{policy}_{At-1} - \text{policy}_{Bt-1}|$), or simply calculate the (absolute) difference in the two policies, in which case the dependent variable remains continuous.

Methodological Issues

As we have seen, dyadic datasets are commonly employed in the international relations literature, in which dependent variables are often relational. In diffusion studies, the dependent variable is *not* relational because mutual influences are unobservable. Nevertheless, as long as the dependent variable can be defined as some form of "increased similarity" between two states, the dyadic set up can be useful because it permits one to easily include relational independent variables in the analysis, along with variables measuring the characteristics of the two states in the dyad. However, the analysis of dyadic datasets entails a number of complications, to which we now turn.

(1) *Standard Event-history Analysis Issues*.——The first set of methodological issues is not specific to the dyadic approach and concerns all types of event-history analyses. We focus here on the logit approach, which, as Beck, Katz, and Tucker (1998) have shown, is equivalent to the Cox model as long as appropriate corrections are introduced[3].

The first point is time dependence: while standard event-history techniques model the baseline hazard directly (either parametrically – like in the Weibull model – or nonparametrically – like in the Cox model), in the logit approach, one needs an explicit control. Beck, Katz, and Tucker (1998) advocate the use of cubic splines, but recently Carter and Signorino (2007) have suggested that the simple inclusion of t, t^2, and t^3 performs as well as splines. Time dummies are, in principle, another option, but they are problematic for rare-events data because they predict failure perfectly for years when no change happened (that is, when the dependent variable was 0 for all units). In the logit framework, this means that many observations must be dropped (Carter and Signorino 2007).

Second, in case of "multiple failures" – in our context, if more than one "imitation" is possible – we need to relax the default assumption that all imitations events are generated identically, or in other words that subsequent imitations are independent from the previous imitation history. This

[3] For a thorough treatment of event-history techniques, see Box-Steffensmeier and Jones (2004).

issue is not easily solved even in standard state-year framework. Beck, Katz, and Tucker (1998, 1272) suggest the "primitive" solution of simply including a variable that counts the number of previous events. This approach is justified in some instances, namely if its implicit assumption – that "the odds of an event increase by a factor of proportionality with each subsequent event occurrence" (Box-Steffensmeier and Zorn 2002, 1086) – is reasonable. If not, more complex strategies have to be followed in the context of the Cox model[4] (Box-Steffensmeier and Zorn 2002; Box-Steffensmeier and Jones 2004, 155–182). Researchers first have to determine the nature of their events, namely, whether they are best conceptualized as *multiple* or *repeated* events. Multiple events are different events – for example, the introduction of various policies or various dimensions of the same policy – and can be modeled through a multinomial logit. In this case, coefficients are specific for each event. On the other hand, repeated events are events of the same type that can occur more than once – for instance, tax cuts or increases. Repeated events can be analyzed through a conditional Cox model, whose advantage is that the sequence of the events is modeled explicitly (Box-Steffensmeier and Jones 2004, 158-162; Box-Steffensmeier and Zorn 2002). Concretely, the model allows baseline hazards to vary across events. On the other hand, a single set of coefficients is estimated. We should also note that another option has very recently been put forward to model repeated events, namely the conditional frailty model (Box-Steffensmeier, De Boef, and Joyce 2007). Unlike the conditional Cox, this model takes into account not only event dependence but also cross-sectional heterogeneity.

A third complication is that one of the assumptions of event-history models is "that the magnitudes of the effects of covariates on the duration of a state remain proportional across the life of the process" (Box-Steffensmeier, Reiter, and Zorn 2003). This is known as the "proportional hazards" assumption, which is likely to be violated in many cases because it implies that variables have constant effects over time, or in other words, that there is no longitudinal heterogeneity. This is a technical problem but, above all, it is a substantive issue. In effect, there may often be theoretical reasons to think that the nature of the diffusion process changes over time. Gilardi, Füglister, and Luyet (2009), for instance, and that learning gained importance over time in the diffusion of hospital-financing reforms. Mod-

---

[4] We can note that the vast majority of studies in the international relations literature has relied on the inclusion of a count variable.

eling nonproportional hazards is quite straightforward: it requires simply the introduction of an interaction between some measure of time[5] and the relevant variable(s) (Box-Steffensmeier, Reiter, and Zorn 2003; Box-Steffensmeier and Jones 2004). Diagnostics are available to detect the nonproportionality of hazards, but of course, interactions can also be included *a priori* on theoretical grounds.

Finally, King and Zeng (2001b, 2001a) have recently shown that the logit is biased when the event under study is rare, that is, when the dependent variable has many more 0s than 1s, which is likely to be the case in many dyadic diffusion studies (and is almost always the case international conflict research, which is the context of King and Zeng's contribution). King and Zeng (2001b, 2001a) demonstrate that with rare-events data, the logit underestimates $Pr(Y_{it} = 1)$, and they put forward a method, implemented in a package (available for Stata) called relogit (rare-events logit), which corrects for the bias in the estimates of the βs.

(2) *Dependence Structure and Cross-sectional Heterogeneity*.——Dyadic datasets exhibit complex dependence structures, and even more so if they are directed. In the hypothetical dataset shown in Table 1, observations are certainly not independent within dyads: what happens in Geneva-Zurich 2001 depends in part on what happened in the same dyad in 2000. Second, we can see that a change of policy in $State_i$ is a necessary condition for imitation to occur: if $State_i$ keeps its policy constant, it cannot become more similar to $State_j$. Therefore, it means that observations are not independent, not only within the same dyad, but also across all dyads sharing the same $State_i$. Third, one could also question the independence of observations across dyads sharing the same $State_j$: is Geneva–Bern really independent from Zurich–Bern? In this case, however, there does not seem to be such a mechanic link as in the case of same $State_i$ dyads. The problem of the nonindependence of observations in the directed-dyads approach has been noted in the literature, but as Bennett and Stam (2000, 660) conclude, "there is no obvious fix for the problem". A quick fix, however, is to assume is that observations are independent across dyads with different $State_i$ and to adjust standard errors accordingly[6].

Second, time-series cross-section analyses often include "fixed effects" (country dummies) to account for unobserved heterogeneity across units (Beck 2001). Recently, this issue was discussed also in the context of dyad-

[5] Options include t, $ln(t)$, $\sqrt{t}$, and $ln(\sqrt{t})$, among other functional forms.

[6] In Stata, this can be done by clustering observations on ID_i.

ic analyses of interstate conflicts, and Green, Kim, and Yoon (2001) argued that fixed effects for dyads should always be included, at least as a robustness check. Others, however, strongly warned against this practice (Bennett and Stam 2000; Beck and Katz 2001; King 2001). Including fixed effects in this context raises two types of problems. First, as is well known, fixed effects make the estimation of time-invariant (or rarely changing) variables difficult because of high collinearity[7]. Second, and more important, dyads where there is no variation on the dependent variable must be dropped because of the "separation" problem in the logit model, which arises when a variable perfectly predicts the outcome and makes the model impossible to estimate (Carter and Signorino 2007). The issue is serious for both diffusion and international conflict analyses, because in both cases, the event of interest is rare. In the diffusion analysis conducted by Volden (2006), for example, only 10.4% of observations were coded 1 on the dependent variable, while in conflict studies, the proportion can be as low as 0.3% (King and Zeng 2001a, 694). This means that a very large share of the data must be dropped if fixed effects are included, which is obviously problematic. Note that the same problem arises when using time dummies to account for time dependence, as noted earlier (Beck, Katz, and Tucker 1998; Carter and Signorino 2007).

A solution for both problems could come from multilevel modeling (Gelman and Hill 2007). A dyadic diffusion model could be conceptualized as a nonnested structure in which observations are grouped within State$_i$, State$_j$, and years. Such a model could be written as follows (see also Shor, Bafumi, Keele, and Park 2007):

$$y_{ijt} \sim \text{Bern}(\alpha_i + \alpha_j + \alpha_t + X\beta) \quad (3)$$
$$\alpha_i \sim N(0, \sigma^2_{\alpha I}) \quad (4)$$
$$\alpha_j \sim N(0, \sigma^2_{\alpha J}) \quad (5)$$
$$\alpha_t \sim N(0, \sigma^2_{\alpha T}) \quad (6)$$

In this context, the dependent variable is a probability, and so the stochastic component is assumed to follow a Bernoulli distribution (for other dependent variables, other distributions can of course be used). $X\beta$ represents explanatory variables at the State$_i$, State$_j$, and dyadic levels, along with an intercept, and is not indexed for convenience. In addition, the systematic component includes three intercepts at the State$_i$, State$_j$, and year levels,

[7] Plümper and Troeger (2007) have recently put forward a new technique to solve this problem.

which are modeled as being drawn from a normal distribution. Perhaps more intuitively, the model can also be written as:

$$y_{ijt} = (\alpha_i + \alpha_j + \alpha_t) + X\beta + (\epsilon_i + \epsilon_j + \epsilon_t + \epsilon_{ijt}) \tag{7}$$

There are two important points here. First, each State$_i$, State$_j$, and year has its own intercept, which helps account for cross-sectional heterogeneity while at the same time allowing the inclusion of constant or rarely changing variables (Gelman and Hill 2007, 269). Second, each level has its own error and its own estimated variance, which helps address the complex dependencies that arise in dyadic datasets.

This setup seems very promising but would require an in-depth examination that is beyond the scope of this article. We put it forward here as a conjecture and, consequently, we do *not* pursue it in the empirical illustration in Section 3. For an early application, see Gilardi (2008).

(3) *"Imitation Bias"*.——Boehmke (2008) has recently shown that there is a danger of "imitation bias" in dyadic analyses of policy diffusion, which arises especially when the policy under examination is measured by a binary indicator and when there is a clear trend toward the adoption of that policy. This configuration is typical of many country-year event-history analyses of international policy diffusion, for example, pension privatization (Brooks 2005) or independent regulatory agencies (Jordana and Levi-Faur 2005; Gilardi 2005), and corresponds to the classic S-shaped process in which at the beginning, all countries have a given policy and progressively switch to an alternative. If we analyze such data in a dyadic framework, the problem is that $\Pr(y_{ijt} = 1) = 0$ unless State$_j$ has already adopted the new policy. In other words, a necessary condition for the dependent variable to be coded 1 is that State$_j$ does not have the same policy as State$_i$. It also follows that $\Pr(y_{ijt} = 1)$ increases as $\Pr(y_{ijt} = 1)$ increases, that is, as the probability that State$_j$ adopts the new policy (and thus can be emulated) increases. In a dyadic setup, any variable marking a State$_j$ that has adopted the new policy will turn out to be positive and significant simply because it identifies a state that *can* be emulated, in contrast to states that *cannot* be emulated because, like State$_i$, they still have the old policy.

The severity of this problem depends on the nature of the policy. In Volden's (2006) study, there is not a clear trend away from one policy and towards another one, but instead, there is a complex mix of policies that states change over time, which makes them more similar to some states and more different from others. In this context, the problem highlighted by Boehmke (2008) is much less worrisome, although dyad-years where $\Pr(y_{ijt} = 1) = 0$ (that is, where emulation is impossible) may still exist, and those are where State_i and State_j have identical mixes of policies. The likelihood of this combination decreases with the number of policy dimensions and with the level of measurement of each of the dimensions, the best case being where it is interval-ratio, and the worst in which it is nominal or ordinal with just two categories.

As a solution, Boehmke (2008) suggests to condition the estimation on the "opportunity to imitate" – in other words, to exclude from the analysis dyad-years in which $\Pr(y_{ijt} = 1) = 0$. These dyad-years are those where $y_{it} = y_{jt-1}$, which makes it impossible for State_i to become more similar to State_j since they both already have the same policy. Monte Carlo simulations show that with this adjustment, dyadic analyses are much less likely to find diffusion effects where none exist.

Boehmke (2008) studies this problem only within a rather narrow set of assumptions on the data-generating process, and his conclusions should therefore not (yet) be seen as definitive. However, his warning is on-the-spot, and his simple correction can certainly be used at least as a robustness check.

Empirical Illustration: Health-insurance Subsidy Policies in Swiss Cantons

The previous section has shown that a dyadic setup can be used to study policy diffusion in federal states. On the other hand, researchers need to be aware of a series of complications that need to be addressed in empirical analyses. Some of these issues are quite well understood, while others, such as complex dependencies in the data and potential "imitation biases", have just begun to be explored.

In this section we present an empirical application. Our goal is to illustrate how the dyadic approach can be applied in the Swiss context. Although we focus on a specific policy, the main characteristics – loose federal framework and varied cantonal implementations – are typical of Switzerland and can be found in many, if not most, other areas. This illustration, therefore, aims to demonstrate that the dyadic approach is a fruitful way to investigate a highly relevant, yet so far largely neglected aspect of Swiss policy making.

Health insurance subsidies in the LAMal

With the Health Insurance Law of 1994 (LAMal), which came into force in 1996, Switzerland introduced mandatory health insurance with uniform premiums for each person irrespective of his or her financial situation. In order to reduce the social inequalities created by per capita premiums and to ensure solidarity between people with different income levels, the LAMal introduced mechanisms for the reduction of the individual premiums. In 2004, about one third of the Swiss population were beneficiaries of health insurance subsidies. Since cantons are in charge of the implementation of the law, however, outcomes differ greatly across regions.

As a consequence of the significant freedom given to cantons, 26 different subsidy policies have been developed since the LAMal came into force. Cantons have changed parts of their policies several times. We can distinguish five major domains in which the implementation of the federal law varies, namely eligibility and benefits, identification of beneficiaries, up-to-dateness of calculations, modalities of payment, and exhaustion of the federal contribution[8] (Balthasar 2003). Within these five domains, cantons have changed their practices several times during the last ten years. There have been some tendencies towards convergence, for example toward the use of a percentage model to define eligibility, or toward the payment of subsidies directly to the insurers to guarantee the earmarked use of the money. But despite these changes and several attempts to harmonize the system, differences between cantons remain: they still use their freedom in the implementation of the federal law and experiment with different practices.

Hypotheses

In this section, we present the hypotheses that guide the statistical analysis. Their purpose in the context of this article is to illustrate how a dyadic analysis can accommodate variables on Canton$_i$s, Canton$_j$s, and on the relationship between the two.

A first set of simple and quite atheoretical hypotheses relate to characteristics of Canton$_i$. We conjecture that the population of a canton, its language, the level of its insurance premiums, how many people receive subsidies, and the partisan affiliation of the health minister can be related

---

[8] A short description of these dimensions is provided in Appendix 2.

to the probability that a canton adopts a policy that is already in place elsewhere.

A strongly theoretical hypothesis is related to a characteristic of Canton$_j$, namely, the extent to which their subsidy policies are "successful". Defining and measuring the "success" of a policy is of course tricky; we will explain our strategy in detail in the next section. This hypothesis is based on Volden's (2006) finding that more successful policies are more likely to be adopted elsewhere, and is linked to the broader literature on the role of learning in diffusion processes (Meseguer 2004; Meseguer 2006a; Meseguer 2006b; Elkins, Guzman, and Simmons 2006; Gilardi, Füglister, and Luyet 2008).

Third, we have a number of relational hypotheses. We expect cantons to be more likely to take up the policies of other cantons in the same region or that share the same language, which can be either a bounded form of learning in which "available" examples are taken into account, or a form of emulation were conformity within peer groups is pursued. We also expect cantons to look more closely at other cantons with similar levels of insurance premiums, because this signals shared problems and therefore also potentially useful solutions. This would also be a bounded form of learning. Finally, if partisan networks play a role in the diffusion of policies, a canton might be more likely to imitate another canton if the two health ministers are from the same party.

Operationalizations and Data

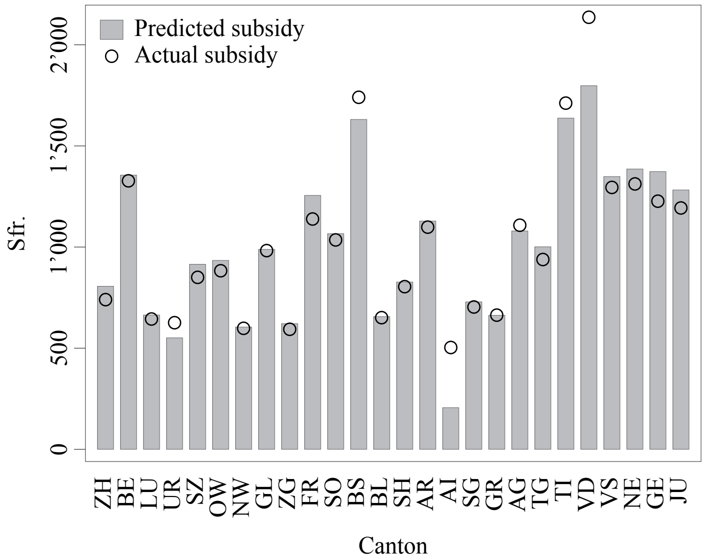
Our dependent variable has six components that refer to four different aspects of cantonal policies. On this basis, the dependent variable records whether Canton$_i$ adopts at least one of the policies of Canton$_j$. We have collected data on changes in the subsidy policies of the 26 Swiss cantons from 1997 (that is, one year after the implementation of the LAMal) to 2005. We have relied on three sources: the Conference of the Cantonal Directors of Public Health, which publishes yearly synoptic tables on the policy instruments concerning the health insurance premium reduction practices of each canton (GDK 2006); monitoring reports published by the Federal Office of Public Health, which examine the effectiveness of the health care subsidies (Balthasar, Bieri, and Müller 2005); and Balthasar's (2003) study of cantonal subsidy policies.

To test the learning hypothesis, we have constructed a variable for measuring the success of cantonal policies. Measuring success is inherently difficult, and what we propose here is just a first cut. We assume that the relevant outcome is the generosity of the subsidy. We measure this simply by dividing the annual budget for insurance subsidies in each canton by the number of beneficiaries. In other words, this is the yearly amount that, on average, a subsidy beneficiary receives in each canton. The assumption is that the higher the amount, the better. While it is probably not accurate to assume that all cantons aim for higher generosity, it is worth reminding that most cantons fail to respect the informal standard set by the Federal Council according to which for no households should health insurance exceed 8% of revenue. In this sense, more generous cantons are closer to this benchmark, and therefore can to some extent be considered more "successful".

In addition, to assess the degree of "success", we take into account a number of factors that are more or less "mechanically" related to cantonal generosity levels, that is, that do not depend on the effectiveness of the specific policy mix. The first is the extent to which federal contributions are used. The federal level attributes to each canton a given sum for subsidies. In principle, cantons should throw in an equal amount, but they can reduce it by up to 50%. The more they make use of this possibility, the lower their budget, all else equal. Second, cantonal variations in generosity are explained by differences in health insurance premiums: the higher the premiums, the higher the subsidies, all else equal. Third, cantons can choose whether to give more money to fewer people or less money to more people. Therefore, the share of the population that receives a subsidy is also a relevant factor that must be taken into account.

The basic idea in our operationalization is that a "good" cantonal policy is one that is more generous than the Swiss average, taking these factors into account. Following Volden (2006), for each year we have regressed generosity on the use of federal contributions, insurance premiums levels, and share of beneficiaries. These three factors are strong predictors of generosity levels: the R^2 is above 0.9. We have then compared predicted and actual generosity levels. The results for year 2000 are shown in Figure 1. In the canton of Geneva (GE), for instance, each beneficiary received less money than he or she should have, taking into account the level of premiums in this canton, the number of beneficiaries, and what share of the federal contribution was used. By contrast, the canton of Vaud (VD) is more successful: the generosity of its subsidy policy is much higher than it should be, given its level of premiums, the number of beneficiaries, and the use of federal contributions. We therefore use the difference between

Figure 1: Measuring the "success" of health insurance policies (year 2000)



actual and predicted values as an indicator of success. While Volden (2006) coded successes in a binary way, we keep the continuous measure that results from this computation. Success can be a matter of degree.

Information on the partisan affiliation of cantonal health ministers comes from the *Jahrbuch Schweizer Politik*, 1995–2005. Shared partisanship for health ministers is coded 1 if $\text{Party}_{A,t} = \text{Party}_{A,t-1}$ and $\text{Party}_{A,t-1} = \text{Party}_{B,t-1}$, that is, if the health minister of Canton_i is from the same party as the minister of Canton_j in the previous year, and was already in place the previous year.

We have also constructed variables for various similarities between cantons. First, two variables record whether the two cantons in the dyad belong to the same region or have the same main language. Second, we look at health insurance premiums, beneficiary rates, and population size. For insurance premiums, we have also computed the absolute difference between Canton_i and Canton_j. For the evolution of the health insurance premiums and the rate of the beneficiaries, we relied on the statistics published by the Federal Office of Public Health, while data on the size of the population come from the Federal Statistical Office.

Results

The results of the analysis are shown in Table 2. The first model is a logit in which all observations are included, while the second, following (Boehmke 2008), excludes dyad-years where $y_{it} = y_{jt-1}$ to correct for potential "imitation bias". Models 3 and 4 repeat this sequence but with a different estimator, namely King and Zeng's (2001b, 2001a) rare-events logit. This may be appropriate because our dependent variable is coded 1 in only 5.4% of observations in the full dataset (6.2% after conditioning on the "opportunity to imitate").

A first look at the results shows that they are quite consistent across specifications and that imitation is influenced by characteristics of both Canton_i and Canton_j, as well as of their relationship. Canton_i is significantly more likely to adopt a policy that Canton_j already has if its population is larger and if its health minister is a member of the liberal FDP party. On the other hand, Canton_j's policies are more likely to be imitated if they are "successful" – in our definition, if the canton manages to be more generous than average, controlling for its specificities. This finding can be seen as supportive of the learning hypothesis. Furthermore, cantons are more likely to imitate cantons in the same region as well as those that have similar levels of health insurance premiums. Finally, an interesting result is that partisan networks seem to matter: cantons in which the health minister is affiliated with the Christian-Democratic party (CVP) are more likely to adopt the policy of another canton if the latter also had a Christian-Democratic health minister in the previous year.

To better understand what these results mean, it is useful to look at how the probability of imitation varies as a function of these variables. Figures 2–5 show predicted probabilities and the associated confidence intervals, computed with Clarify (King, Tomz, and Wittenberg 2000). An important point for the interpretation of the graphs is that the relationships they depict depend on on the values at which variables not shown in the figure (that is, all variables in Table 2 except that represented on the x-axis of the figure) are kept constant, and this for two reasons. First, like in linear models, each independent variable contributes to the level of the dependent variable. Second, and more subtly, nonlinear model such as the logit include implicit interaction effects (Brambor, Clark, and Golder 2006; Kam and Franzese 2007, 112–13; Berry, Esarey, and Rubin 2007). A discussion of this point is beyond the scope of this article, but the intuition is that, in a logit, a variable has greater effects if the predicted probabilities are closer to 0.5

Table 2: Logit analysis of the probability that Canton$_i$ imitates Canton$_j$

	Model 1	Model 2	Model 3	Model 4
Population$_i$	0.002**	0.001**	0.002**	0.001*
	(2.27)	(1.96)	(2.24)	(1.94)
Premiums$_i$	-0.009	-0.011	-0.009	-0.011
	(0.97)	(1.09)	(0.94)	(1.06)
Beneciaries$_i$	-0.031	-0.038	-0.030	-0.037
	(1.10)	(1.29)	(1.09)	(1.29)
Language$_i$ (German)	0.628	0.513	0.623	0.509
	(0.62)	(0.50)	(0.61)	(0.50)
FDP$_i$	1.583***	1.421**	1.559***	1.398**
	(2.85)	(2.43)	(2.82)	(2.40)
SVP$_i$	1.286	1.181	1.266	1.160
	(1.46)	(1.32)	(1.45)	(1.30)
CVP$_i$	1.410*	1.336	1.393*	1.319
	(1.79)	(1.61)	(1.77)	(1.60)
SP$_i$	0.821	0.613	0.804	0.598
	(1.51)	(0.97)	(1.49)	(0.95)
Success$_j$	44.483**	87.130***	46.869**	87.825***
	(2.05)	(5.20)	(2.17)	(5.27)
Same region	0.169***	0.222***	0.172***	0.224***
	(2.63)	(3.05)	(2.69)	(3.10)
Same language	-0.135	-0.169	-0.138	-0.171
	(0.48)	(0.62)	(0.50)	(0.63)
Same party(CVP)	0.586***	0.498***	0.601***	0.513***
	(4.09)	(2.77)	(4.22)	(2.87)
Same party (SVP)	-0.338	-0.385	-0.198	-0.245
	(0.70)	(0.83)	(0.41)	(0.53)
Same party (FDP)	-0.020	0.018	0.010	0.046
	(0.07)	(0.06)	(0.04)	(0.15)
Same party (SP)	-0.512	-0.684*	-0.488	-0.658*
	(1.49)	(1.78)	(1.43)	(1.72)
$\lvert$ Premiums$_A$ – premiums$_B$ $\rvert$	-0.010**	-0.011***	-0.010**	-0.010***
	(2.50)	(2.68)	(2.48)	(2.65)
Wald χ^2	199.81	722.82	–	–
Observations	5’200	4’536	5’200	4’536

Notes: Robust z statistics in parentheses (for clustering on Canton$_i$). * significant at 10%; ** significant at 5%; *** significant at 1%. Constant, `time`, `time`^{2}, and `time`^{3} included but not shown to save space.

(where the slope of the S-shaped curve is steeper). So if x_1 has a strong influence on the outcome and moves the probability close to 1, it forces x_2, x_3, etc. to have smaller effects, since the probability cannot exceed 1. This is sometimes known as the "compression effect".

With these two caveats in mind, we can proceed to the interpretation of the figures.

Figure 2 shows that the partisan affiliation of the health minister influences the probability of imitation: cantons in which the health minister is a radical (FDP) are more likely to take up policies already present elsewhere. More precisely, it seems that an FDP minister is a sort of necessary condition for imitation. If the minister is not FDP, imitation is unlikely: the point estimate is close to 0 with a narrow confidence interval. By contrast, if the minister is FDP, the predicted probability of imitation is higher, but with a much larger confidence interval, meaning that some cantons with FDP ministers tend to imitate, but others do not.

Figure 3 shows how the probability of imitation increases as Canton$_j$'s policy becomes more "successful". This evidence is consistent with learning arguments, which state that policy makers update their beliefs on the effects of policies by looking at the experience of others. Figure 3 indicates that Canton$_i$ is more likely to adopt a policy that Canton$_j$ already has if the latter has proved to be effective. Of course, the specific indicator of "success" can be debated, but these results are consistent with the theoretical expectations.

The next figures show the effects of the relational characteristics of the dyad on the probability of imitation. Figure 4 suggests that partisan networks may matter, especially for the Christian-Democratic party (CVP). The probability that Canton$_i$ imitates Canton$_j$ is greater if in both cantons the health minister is Christian-Democratic, although we see that there is a significant overlap in the confidence intervals. Figure 5 is similar: cantons tend to imitate other cantons in the same region, but the effect is not very significant, neither statistically nor substantively. Finally, Figure 6 shows that similarities of health insurance premiums also matter: cantons tend to adopt the policies of other cantons in which insurance premiums are approximately at the same level. This could be evidence for a bounded form of learning, in which policy makers are influenced by others sharing similar problems when looking for solutions.

These findings teach us something new and important about policy making in Switzerland. Policies do diffuse across cantons, and the pattern is not simply one of geographical proximity, nor of *Röstigraben*. Learning

Figure 2: Predicted probability of imitation as a function of the partisan affiliation of the health minister in $Canton_A$

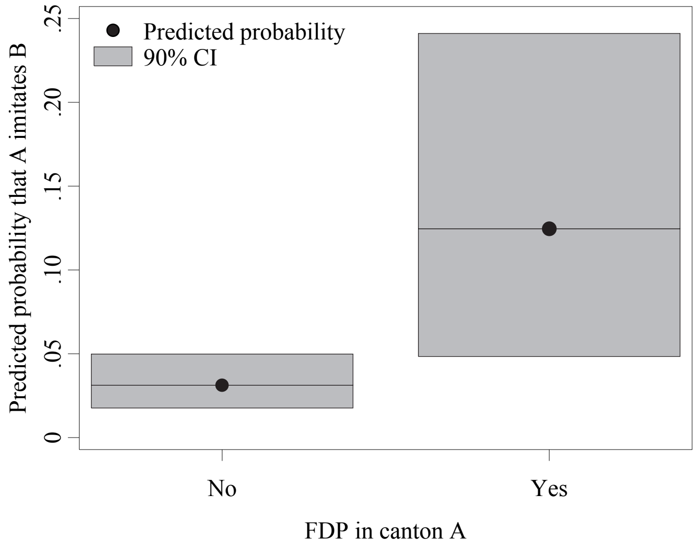


Figure 3: Predicted probability of imitation as a function of the "success" of $Canton_B$'s policy

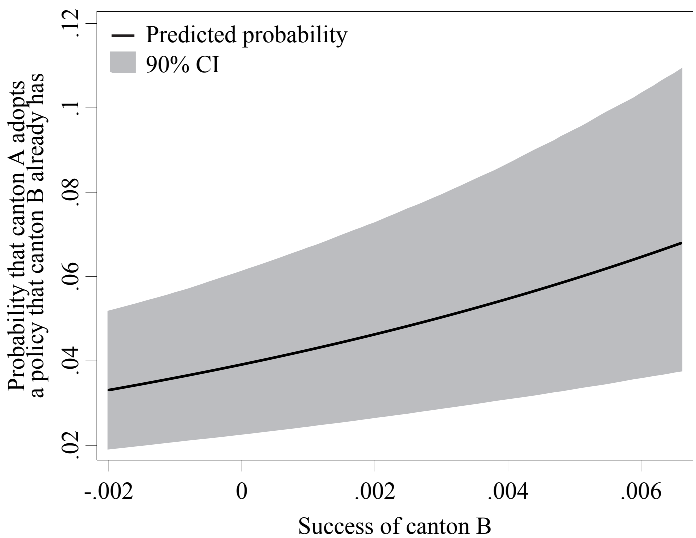


Figure 4: Predicted probability of imitation as a function of shared partisan affiliation of health ministers

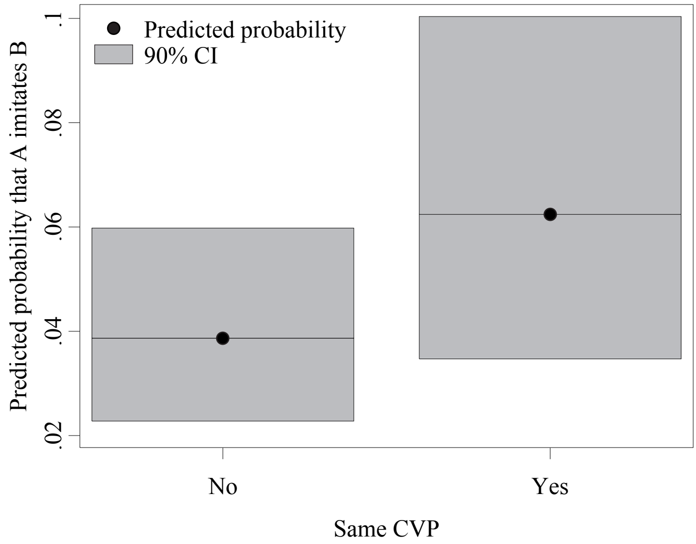


Figure 5: Predicted probability of imitation as a function of geographic proximity

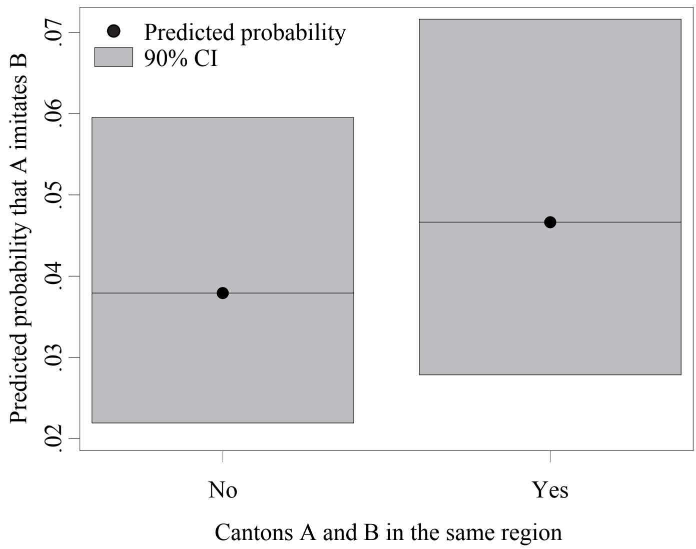
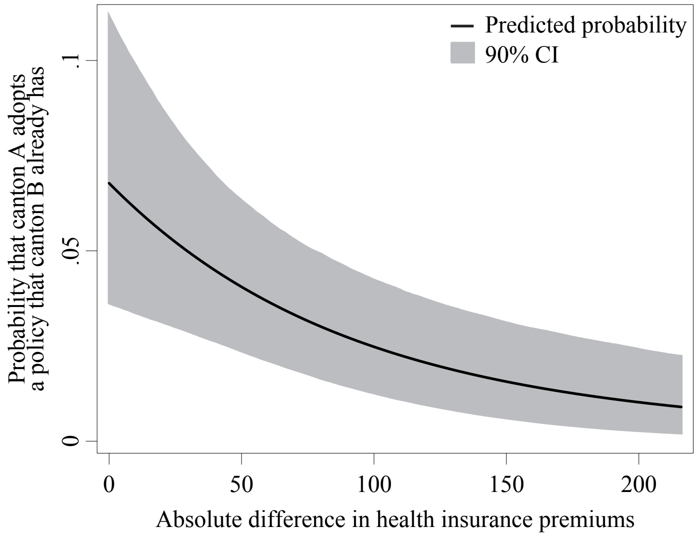


Figure 6: Predicted probability of imitation as a function of the similarity of health insurance premiums



seems to matter, which corroborates the idea that federalism is a "policy laboratory" in which best practices can be developed and spread. In addition, partisan networks seem to play a role, and all cantons are not equally prone to imitation. This is a quite rich and nuanced account of diffusion. While the analysis is still preliminary and reports work in progress, one thing is clear: the dyadic approach is a useful tool to study policy diffusion in federal states.

Conclusion

In this article, we have discussed the use of a dyadic approach to study policy diffusion with a focus on federal states and especially Switzerland. We emphasize again that our aim was *not* to develop a new method, but simply to give a practical overview that can be useful for scholars interested in the empirical analysis of diffusion. The topic seems to have attracted recent interest in the Swiss political science community (Schaltegger 2004; Kübler and Widmer 2007), and we believe that our survey can be a useful tool.

What defines the dyadic approach is that units of analysis are not states, but pairs of states (dyads). This setup has been used extensively in the international relations literature, in which the dependent variable is in many cases relational (e.g., conflict, trade, foreign investment, etc.). Although, in a policy diffusion context, the dependent variable is *not* observably relational, Volden (2006) has convincingly argued that the dyadic approach can nonetheless be useful: defining each state as a potential "sender" and "receiver" of a policy allows one to easily include relational variables, which are at the core of many diffusion hypotheses.

However, the use of a dyadic approach in policy diffusion studies comes with several problems. Since the influence of a state over another state's policy choices is not directly measurable, the dependent variable needs to be constructed indirectly, and special attention has to be paid to its interpretation. Policy change, the observable phenomenon, takes place at the state level, whereas units of analysis are dyads. The dependent variable, therefore, does not record policy change, nor influence (which is unobservable, and whose estimation is the goal of the analysis). It simply measures increased similarity in the observed policies of the two states in a dyad. In our empirical example, the dependent variable records whether Canton_i moves closer to Canton_j in at least one aspect of its health insurance subsidy policy. While similarities can best be measured when looking at pairs, diffusion is not a bilateral phenomenon. The goal, therefore, is to uncover systematic patterns in increased similarities, which then allow one to make inferences about the underlying diffusion process.

Besides the definition of the dependent variable, the analysis of a dyadic data set presents further complications. In addition to standard event-history analysis issues such as time dependence, dyadic diffusion studies exhibit complex dependence structures and are potentially biased toward detecting imitation. Unfortunately, only partial solutions to these problems are currently available. Multilevel methods could be a promising way to deal with the particular structure of these datasets.

Our empirical illustration has shown how these methodological issues can be addressed in practice and how the dyadic approach contributes to a more precise analysis of diffusion patterns in federal states. The empirical analysis of health insurance policies in Swiss cantons has shown that cantons are influenced by the policy choices of their peers, and that diffusion matters in Swiss federalism. Four results highlight the usefulness of the dyadic approach. First, imitation is influenced by the individual characteristics of the cantons, as well as by the relationship between the two cantons in a dyad. Second, "similarity" means more than geographic proximity. Policy makers are not only more likely to imitate cantons in the same region, but they also seem to imitate cantons that are confronted with similar problems. Third, learning matters: cantons with successful policies are more likely to be imitated. And finally, partisan networks seem to matter as well.

Although these results are preliminary, they supply new and interesting insight into the nature of policy diffusion in Switzerland, and they show how the dyadic approach can be fruitfully employed to study how policies spread among federal states. The approach has, however, a much broader scope and could certainly be employed to study diffusion at other levels, such as cities, countries, and in fact any other setting in which diffusion hypotheses can be developed.

Appendix 1: Stata Code for the Construction of a Dyadic Dataset[9]

In most cases, the starting point of a dyadic analysis is a standard dataset in which observations are state-years. Table 3 shows an excerpt of one such dataset.

The first step for the construction of a directed dyadic dataset is to create a copy of the original state-year dataset:

```
copy dataset1.dta dataset2.dta
```

The `joinby` command then permits one to create a new dataset including all pairwise combinations of observations, which is what we need in a dyadic dataset. Before using `joinby`, however, we have to rename the variables so that we know which ones refer to which canton in the dyad. To do this, we can employ the `renvars` routine written by Jeoren Weesie and Nick Cox, which can be located and then downloaded by typing `findit renvars` in Stata. `joinby` requires one to specify the groups within which observations are combined. In our case, the relevant groups are identified by the variable `year`. Therefore, we use the first dataset and append all variables except year (which will be used to join the two datasets) with the postfix `_i`:

[9] The code can be downloaded at www.fabriziogilardi.org.

Table 3: State-year dataset (excerpt) (`dataset1.dta`)

Canton	Year	ID	Policy	Language	Population
Geneva	2000	1	2	1	409’900
Geneva	2001	1	2	1	416’400
Geneva	2002	1	2	1	422’200
Zurich	2000	2	1	2	1’222’200
Zurich	2001	2	1	2	1’242’000
Zurich	2002	2	2	2	1’257’900
Bern	2000	3	3	2	950’600
Bern	2001	3	2	2	952’900
Bern	2002	3	2	2	955’400
Schwyz	2000	4	1	2	128’400
Schwyz	2001	4	1	2	129’600
Schwyz	2002	4	1	2	131’700

```
use dataset1.dta
renvars state-population, postfix(_i)
rename year_i year
save, replace
```

We then do the same for the second dataset, but with the postfix `_j`:

```
use dataset2.dta
renvars state-population, postfix(_j)
rename year_j year
save, replace
```

We can now create the dyadic dataset and rearrange the variables in a more convenient order:

```
use dataset1.dta
joinby year using dataset2.dta
save dataset_dyadic.dta

use dataset_dyadic.dta
order canton_i canton_j year
sort canton_i canton_j year
save, replace
```

The dataset that results is shown in Table 1. We can now construct the dependent variable imitation, following the logic explained in section "Dataset Structure and Dependent Variable". This code loops over all pairwise combinations of cantons and all years, and does the following.[10] First, it captures the policies of Canton$_i$ and Canton$_j$ at time t–1. Second, it codes `imitation 1` if Canton$_i$ has changed it policy at time t and if the new policy is the one that Canton$_j$ had at time t–1.

```
g imitation=0
#delimit ;
forval i=1/26 {;
 forval j=1/26 {;
  forval y=1998/2005 {;
   quietly sum policy_i if year==‘y’-1 & id_i==‘i’;
   local x=r(mean);
   quietly sum policy_j if year==‘y’-1 & id_j==‘j’;
   local z=r(mean);
   quietly replace imitation=1 if year==‘y’ &
   policy_i!=‘x’ & policy_i==‘z’ & id_i ==‘i’ &
   id_j==‘j’;
  };
 };
};
#delimit cr
save, replace
```

We can also create easily independent variables that measure to the relationship between the two states in the dyad – for example, whether they share the same language:

```
gen samelanguage=0
replace samelanguage=1 if language_i==language_j
save, replace
```

We also need to identify the same-canton dyads that were created through the joinby command, because we want to exclude them from the analysis:

[10] The `#delimit ;` line means simply that the end of a command is marked by the semicolon instead of the default carriage return. We do this is just to preserve a correct display of a long command line in the printed page. We then switch back to the default delimiter with `#delimit cr`.

```
gen samecanton=0
replace samecanton=1 if id_i==id_j
save, replace
```

Finally, we also need to create identifiers for dyads. When running the analyses, in some cases we may want to adjust the standard errors for the nonindependence of observations within dyads. To do this, we can employ a simple loop telling Stata to take each pairwise combination of cantons and code the variable id_dyad with a unique identifier. The loop starts with `id_i=1` and `id_j=1` and codes `id_dyad 1`. Then it takes `id_i=1` and `id_j=2` and codes the identifier 2 and so on until the dyad where `id_i=26` and `id_j=26` is coded:

```
gen id_dyad=.
local x=1
forval y=1/26 {
 forval z=1/26 {
  replace dyadid=‘x’ if id_i==‘y’ & id_j==‘z’
  local x=‘x’+1
  }
}
save, replace
```

Assuming that all relevant variables are included, the dataset is now ready for the analysis.

Appendix 2: The Dimensions of Cantonal Subsidy Policies

(1) *Eligibility and Benefits*.——The majority of cantons use a percentage model: if premiums exceed a given share of income (which can vary across cantons), the person is eligible for benefits. Other cantons use a threshold model, in which all households below a certain income are eligible for a fixed subsidy. Cantons also use different bases to calculate income.

(2) *Identification of Beneficiaries*.——Some cantons identify the beneficiaries on the basis of the tax declaration and then pay the subsidies automatically, while others automatically inform eligible people, who, however, still need to fill in an application to receive the subsidy. In a few other cantons, potential beneficiaries need to apply for subsidies without having been informed about their eligibility status. In some cantons, applications

for subsidies can be sent to the cantonal authorities all year through, while in others, the form can be handed in only at a specific date.

(3) *Up-to-dateness of Calculations*.——Some cantons use final taxation decisions, others provisional ones, and still others use salary certificates to estimate revenues.

(4) *Modalities of Payment*.——In some cantons, subsidies are paid directly to the insurers, which ensures the earmarked use of the money, while others emphasize the transparency of the costs and pay subsidies to the beneficiaries, who can thus use them freely and not only to pay their insurance premiums.

(5) *Exhaustion of the Federal Contribution*.——Cantons receive a given sum from the federal government and should, in principle, add an equal amount to the budget for subsidies. However, they can reduce this amount up to 50% provided that the social objectives of the policy are not jeopardized. Note that the definition of the objectives is extremely loose so that virtually every canton can claim that the reduction is justified. If a canton reduces its contribution, then the federal contribution is also cut proportionally.

References

Balthasar, A. (2003). Die Prämienverbilligung im Krankenverischerugsgesetz: Vollzugsföderalismus und sekundäre Harmonisierung. *Swiss Political Science Review* 9(1): 335–54.

Balthasar, A., Bieri, O. and F. Müller (2005). *Monitoring 2004: Die sozialpolitische Wirksamkeit der Prämienverbilligung in den Kantonen*. Bern: Bundesamt für Gesundheit.

Beck, N. (2001). Time-series Cross-section Data: What Have We Learned in the Past Few Years? *Annual Review of Political Science* 4: 271–93.

Beck, N. and J. Katz (2001). Throwing Out the Baby with the Bath Water: A Comment on Green, Kim, and Yoon. *International Organization* 55(2): 487–95.

Beck, N., Katz, J. and D. Tucker (1998). Taking Time Seriously: Time-series Cross-section Analysis with a Binary Dependent Variable. *American Journal of Political Science* 42(4): 1260–88.

Bennett, D. (2006). Toward a Continuous Specification of the Democracy-autocracy Connection. *International Studies Quarterly* 50(2): 313–38.

Bennett, D. and A. Stam (2000). Research Design and Estimator Choices in the Analysis of Interstate Dyads. *Journal of Conflict Resolution* 44(5): 653–85.

Berry, F. and W. Berry (1990). State Lottery Adoptions as Policy Innovations: An Event History Analysis. *American Political Science Review* 84(2): 395–415.

Berry, W., Esarey, J. and J. Rubin (2007). *Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential?* Online at: http://polmeth.wustl.edu/retrieve.php?id=692 [accessed: 30.07.2008].

Boehmke, F. (2008). *Potential Ambiguities in a Directed Dyad Approach to State Policy Emulation*. Paper presented at the Annual Conference of the Midwest Political Science Association, April 3–6, 2008, Chicago, US.

Box-Steffensmeier, J., Boef, S. de and K. Joyce (2007). Event Dependence and Heterogeneity in Duration Models: The Conditional Frailty Model. *Political Analysis* 15(3): 237–56.

Box-Steffensmeier, J. and B. Jones (2004). *Event History Modeling: A Guide for Social Scientists*. Cambridge: Cambridge University Press.

Box-Steffensmeier, J., Reiter, D. and C. Zorn (2003). Nonproportional Hazards and Event-history Analysis in International Relations. *Journal of Conflict Resolution* 47(1): 33–53.

Box-Steffensmeier, J. and C. Zorn (2002). Duration Models for Repeated Events. *Journal of Politics* 64(4): 1069–94.

Brambor, T., Clark, W. and M. Golder (2006). Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis* 14(1): 63–82.

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

Braun, D., Gilardi, F., Füglister, K. and S. Luyet (2007). Ex Pluribus Unum: Integrating the Different Strands of Policy Diffusion Theory. *Politische Vierteljahresschrift* (Special Issue): 39–55.

Brooks, S. (2005). Interdependent and Domestic Foundations of Policy Change: The Diffusion of Pension Privatization around the World. *International Studies Quarterly* 49(2): 273–94.

——— (2007). When Does Diffusion Matter? Explaining the Spread of Structural Pension Reforms across Nations. *Journal of Politics* 69(3): 701–15.

Bussmann, M. and J. Oneal (2007). Do Hegemons Distribute Private Goods? A Test of Power-transition Theory. *Journal of Conflict Resolution* 51(1): 88–111.

Carter, D. and C. Signorino (2007). Back to the Future: Modeling Time Dependence in Binary Data. Online at: http://polmeth.wustl.edu/retrieve.php?id=710 [accessed: 30.07.2008].

Danilovic, V. and J. Clare (2007). The Kantian Liberal Peace (Revisited). *American Journal of Political Science* 51(2): 397–414.

Elkins, Z., Guzman, A. and B. Simmons (2006). Competing for Capital: The Diffusion of Bilateral Investment Treaties, 1960–2000. *International Organization* 60(4): 811–46.

Farber, H. and J. Gowa (1997). Common Interests or Common Politics? Reinterpreting the Democratic Peace. *Journal of Politics* 59(2): 393–417.

Franzese, R. and J. Hays (2006). Strategic Interaction among EU Governments in Active Labor Market Policy-making: Subsidiarity and Policy Coordination under the European Employment Strategy. *European Union Politics* 7(2): 167–89.

Gartzke, E. (2007). The Capitalist Peace. *American Journal of Political Science* 51(1): 166–91.

GDK (2006). *Krankenversicherung: Prämieneverbilligung, synoptische Uebersicht 2006*. Bern: Schweizerische Konferenz der kantonalen Gesundheitsdirektorinnen und -direktoren.

Gelman, A. and J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Gilardi, F. (2005). The Institutional Foundations of Regulatory Capitalism: The Diffusion of Independent Regulatory Agencies in Western Europe. *Annals of the American Academy of Political and Social Sciences* 598: 84–101.

——— (2008). Who Learns from What in Policy Diffusion Processes? Evidence from Unemployment Benefits Retrenchment in OECD Countries. Paper presented at the Annual Meeting of the American Political Science Association, August 28–31, 2008, Boston, US.

Gilardi, F., Füglister, K. and S. Luyet (2009 forthcoming). Learning from Others: The Diffusion of Hospital Financing Reforms in OECD Countries. *Comparative Political Studies* 42(4).

Gray, V. (1973). Innovation in the States: A Diffusion Study. *American Political Science Review* 67(4): 1174–85.

Green, D., Kim, S. and D. Yoon (2001). Dirty Pool. *International Organization* 55(2): 441–68.

Hensel, P., Goertz, G. and P. Diehl (2000). The Democratic Peace and Rivalries. *Journal of Politics* 62(4): 1173–88.

Horowitz, M., McDermott, R. and A. Stam (2005). Leader Age, Regime Type, and Violent International Relations. *Journal of Conflict Resolution* 49(5): 661–85.

Jahn, D. (2006). Globalization as "Galton's Problem": The Missing Link in the Analysis of Diffusion Patterns in Welfare State Development. *International Organization* 60(2): 402–31.

Jordana, J. and D. Levi-Faur (2005). The Diffusion of Regulatory Capitalism in Latin America: Sectoral and National Channels in the Making of New Order. *Annals of the American Academy of Political and Social Sciences* 598: 102–24.

Kam, C. and R. Franzese (2007). *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor: University of Michigan Press.

King, G. (2001). Proper Nouns and Methodological Propriety: Pooling Dyads in International Relations Data. *International Organization* 55(2): 497–507.

King, G., Tomz, M. and J. Wittenberg (2000). Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science* 44(2): 341–55.

King, G. and L. Zeng (2001a). Explaining Rare Events in International Relations. *International Organization* 55(3): 693–715.

——— (2001b). Logistic Regression in Rare Events Data. *Political Analysis* 9(2): 137–63.

Kübler, D. and T. Widmer (2007). Subnationale Unterstützung eines nationalen Drogenprogramms: Eine Panelanalyse mit Ereignisdaten in den Schweizer Kantonen. *Politische Vierteljahresschrift* (Special Issue) 38: 200–20.

Leeds, B. (2003). Do Alliances Deter Aggression? The Influence of Military Alliances on the Initiation of Militarized Interstate Disputes. *American Journal of Political Science* 47(3): 427–39.

Levi-Faur, D. (2005). The Global Diffusion of Regulatory Capitalism. *Annals of the American Academy of Political and Social Sciences* 598: 12–33.

Maoz, Z. and B. Russett (1993). Normative and Structural Causes of Democratic Peace, 1946–1986. *American Political Science Review* 87(3): 624–38.

McVoy, E. (1940). Patterns of Diffusion in the United States. *American Sociological Review* 5(2): 219–27.

Meseguer, C. (2004). What Role for Learning? The Diffusion of Privatisation in OECD and Latin American Countries. *Journal of Public Policy* 24(3): 299–325.

——— (2006a). Learning and Economic Policy Choices. *European Journal of Political Economy* 22(1): 156–78.

——— (2006b). Rational Learning and Bounded Learning in the Diffusion of Policy Innovations. *Rationality and Society* 18(1): 35–66.

Meseguer, C. and F. Gilardi (2008 forthcoming). What is New in the Study of Diffusion? A Critical Review. *Review of International Political Economy.*

Mintrom, M. (1997). Policy Entrepreneurs and the Diffusion of Innovation. *American Journal of Political Science* 41(3): 738–70.

Morrow, J., Siverson, R. and T. Tabares (1998). The Political Determinants of International Trade: The Major Powers, 1907–90. *American Political Science Review* 92(3): 649–61.

Peceny, M., Beer, C. and S. Sanchez-Terry (2002). Dictatorial Peace? *American Political Science Review* 96(1): 15–26.

Plümper, T. and V. Troeger (2007). Efficient Estimation of Time-invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects. *Political Analysis* 15(2): 124–39.

Reiter, D. and A. Stam (2003). Identifying the Culprit: Democracy, Dictatorship, and Dispute Initiation. *American Political Science Review* 97(2): 333–37.

Russett, B., Oneal, J. and D. Davis (1998). The Third Leg of the Kantian tripod for Peace: International Organizations and Militarized Disputes, 1950–85. *International Organization* 52(3): 441–67.

Schaltegger, C. (2004). Finanzpolitik als Nachahmungswettbewerb: Empirische Ergebnisse zu Budgetinterdependenzen unter den Schweizer Kantonen. *Swiss Political Science Review* 10(2): 61–86.

Shipan, C. and C. Volden (2006). Bottom-up Federalism: The Diffusion of Antismoking Policies from U.S. Cities to States. *American Journal of Political Science* 50(4): 825–43.

Shor, B., Bafumi, J., Keele, L. and D. Park (2007). A Bayesian Multilevel Modeling Approach to Time-series Cross-sectional Data. *Political Analysis* 15(2): 165–81.

Simmons, B., Dobbin, F. and G. Garrett (2006). Introduction: The International Diffusion of Liberalism. *International Organization* 60(4): 781–810.

Simmons, B. and Z. Elkins (2004). The Globalization of Liberalization: Policy Diffusion in the International Political Economy. *American Political Science Review* 98(1): 171–89.

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

Volden, C. (2006). States as Policy Laboratories: Emulating Success in the Children's Health Insurance Program. *American Journal of Political Science* 50(2): 294–312.

Walker, J. (1969). The Diffusion of Innovations among the American States. *American Political Science Review* 63(3): 880–99.

Weyland, K. (2007). *Bounded Rationality and Policy Diffusion: Social Sector Reform in Latin America.* Princeton: Princeton University Press.

Empirisches Modellieren von Politikdiffusion in föderalen Staaten: Der dyadische Ansatz

Politikdiffusion ist ein bekanntes Phänomen in föderalen Staaten. Eine der normativen Begründungen eines dezentralen Systems besteht darin, dass diese die Entwicklung und Verbreitung von optimalen Politiken ermöglicht. Gemäss Berry und Berry (1990) ist die Ereignisanalyse die meist gewählte quantitative Methode für die Analyse von Politikdiffusion. Kürzlich hat Volden (2006) jedoch eine dyadische Variante dieser Methode eingeführt, wobei anstelle von Staaten, Paare von Staaten die Untersuchungseinheiten bilden. Dieser Artikel diskutiert den dyadischen Ansatz mit einem besonderen Fokus auf die Diffusion von Politiken in der Schweiz. Dabei soll nicht eine neue Methode vorgestellt, sondern ein praktischer Überblick für Forschende gegeben werden, welche diese Methode anwenden möchten. Der Artikel zeigt auf, wie die Methode von der Literatur der Internationalen Beziehungen zur Politikdiffusionsliteratur gelangt ist, beschreibt die typische Struktur eines Datensatzes im Kontext der Diffusion und diskutiert verschiedene Aspekte der Modellierung. Der Nutzen des dyadischen Ansatzes wird empirisch anhand der Krankenkassenprämienverbilligung in den Schweizer Kantonen illustriert.

Modelage empirique de la diffusion des politiques dans les Etats fédéraux: l'approche dyadique

La diffusion des politiques publiques est un phénomène courant dans les Etats fédéraux: l'une des justifications normatives de la décentralisation est précisément que celle-ci permet de développer et diffuser des solutions novatrices et performantes aux problèmes publics. Suite aux travaux de Berry et Berry (1990), l'analyse de durée (event-history analysis) est devenue la méthode principale pour l'étude quantitative de la diffusion des politiques publiques. Récemment, Volden (2006) a proposé une variante dyadique de cette méthode, dans laquelle les unités d'analyse ne sont pas des Etats, mais des paires d'Etats. En se focalisant sur le cas de la Suisse, cet article montre d'abord comment l'approche dyadique est passée du domaine des relations internationales à celui de la diffusion des politiques publiques, et discute ensuite de manière pratique la structure d'une base de données dyadique ainsi que plusieurs enjeux inhérents à ce type de modélisation. L'utilité de l'approche dyadique est illustrée empiriquement à l'aide des politiques de réduction des primes d'assurance maladie dans les cantons suisses.

Fabrizio Gilardi is Associate Professor of Public Policy at the Institute of Political Science and at the Center for Comparative and International Studies, University of Zurich, Switzerland. His research interests include regulation, comparative political economy, political delegation, methodology, and policy diffusion processes. His work has been published in several international journals, including *Comparative Political Studies*, the *Journal of European Public Policy*, the *Journal of Theoretical Politics*, and the *Review of International Political Economy*, among others. He is also the author of a book on independent regulatory agencies in Europe (Edward Elgar, 2008), and the co-editor of a volume on delegation in contemporary democracies (Routledge, 2006).

Address for correspondence: Institute of Political Science, University of Zurich, Seilergraben 53, CH-8001 Zurich, Switzerland. Phone: +41 (0)44 634 39 87; Fax: +41 (0)44 634 49 25; Email: gilardi@ipz.uzh.ch; Web: www.fabriziogilardi.org.

Katharina Füglister is Research Assistant and PhD Candidate in political science at the University of Lausanne, Switzerland. She is working on the diffusion of health care policies in Swiss cantons. Currently she is a visiting scholar at the Center for Political Studies at the University of Michigan, USA.

Address for correspondence: Institute of Political and International Studies, University of Lausanne, Bâtiment Anthropole, CH-1015 Lausanne, Switzerland. Phone: +41 (0)21 692 31 39; Fax: +41 (0)21 692 31 45; Email: katharina.fuglister@unil.ch.