Four Ways We Can Improve Policy Diffusion Research

Fabrizio Gilardi

Abstract
This article puts forward four strategies to improve policy diffusion research in political science: (1) use existing concepts consistently and improve their measurement, (2) clarify whether the goal is to improve the understanding of diffusion itself or to use diffusion research to explain another phenomenon, (3) pay more attention to the quality of the research design, and (4) discuss explicitly the practical implications of diffusion. Taken together, these recommendations trace a distinctive way forward for policy diffusion research.

Keywords
public policy, policy innovation/diffusion, comparative federalism, federalism, policy process

Introduction
The study of policy diffusion has been a staple of research on American politics for several decades (F. S. Berry and Berry 1990; Boushey 2010; Gray 1973; Karch 2007a; Savage 1985; Walker 1969). A huge number of studies have examined how policies spread among states. Recent developments in areas such as marijuana legalization and same-sex marriage only emphasize the plausibility and relevance of the general argument that policies in one state are influenced by the policies of other states. Meanwhile, parallel literatures have developed in other political science subfields, such as comparative politics and international relations (Gilardi 2012; Graham, Shiman, and Volden 2013; 2014). Despite limited communication, concepts, measures, and findings have been strikingly similar across subfields—and so are the challenges that diffusion research is faced with.

1University of Zurich, Zurich, Switzerland

Corresponding Author:
Fabrizio Gilardi, Department of Political Science, University of Zurich, Affolternstrasse 56, 8050 Zurich, Switzerland.
Email: gilardi@ipz.uzh.ch

DOI: 10.1177/1532440015608761
facing. On one hand, research in many areas has established that diffusion is a real phenomenon. On the other hand, moving beyond this basic finding has proven difficult. Having accomplished the first goal, it is paramount that diffusion research goes beyond showing that “diffusion matters” or “policy X has diffused.” In this article, I discuss four strategies to improve policy diffusion research and make it more relevant for real-world problems: (1) use existing concepts consistently and improve their measurement, (2) clarify whether the goal is to improve the understanding of diffusion itself or to use diffusion research to explain another phenomenon, (3) pay more attention to the quality of the research design, and (4) discuss explicitly the practical implications of diffusion. Although none of the recommendations is completely new, taken together they trace a distinctive way forward for policy diffusion research.

**Tweak Measures, Not Concepts**

I argue that, across political science subfields, we have reached a consensus on what diffusion is and how it should be conceptualized. The definition put forward by Simmons, Dobbin, and Garrett (2006, 787) can be adapted to fit the vast majority of processes that political scientists are interested in:

> [P]olicy diffusion occurs when government policy decisions in a given [jurisdiction] are systematically conditioned by prior policy choices made in other [jurisdictions].

This definition is based on that, even more general, put forward by Strang (1991, 325) in sociology:

> any process where prior adoption of a trait or practice in a population alters the probability of adoption for the remaining non-adopters.

Thus, diffusion is a consequence of interdependence and is not defined exclusively (or even primarily) by the fact that something has spread. An implication is that, when studying diffusion, we are interested more in the process than in the outcome. For instance, convergence—the degree to which the policies of different states become more alike (Holzinger and Knill 2005)—can be a useful complement to a diffusion analysis, or it can motivate the research in the first place, but it is not what we are actually studying. Moreover, although policy adoption is a natural focus for diffusion research, policy abandonment can also be studied from this perspective (Volden 2016).

In addition to the definition of diffusion, there is consensus on three broad classes of diffusion mechanisms: learning, emulation, and competition (Braun and Gilardi 2006; Dobbin, Simmons, and Garrett 2007; Gilardi 2012; Graham, Shipan, and Volden 2013; Karch 2007b; Meseguer and Gilardi 2009; Shipan and Volden 2008; Simmons, Dobbin, and Garrett 2006). Of course, many nuances can be introduced, but these three labels capture the main ideas quite accurately.

First, learning is a process where policies in one unit are influenced by the consequences of similar policies in other units. In other words, policy adoption in one unit
is more likely if the policy has been successful elsewhere (Gilardi 2010; Meseguer 2004; Volden 2006; Volden, Ting, and Carpenter 2008). There are different forms of success. Success can be related to (1) the goals that the policy is designed to achieve, (2) the challenges of its implementation, and (3) its political support. When considering the adoption of a policy, policy makers can learn from others about all these dimensions. For instance, Volden (2006) showed that U.S. states were more likely to imitate health insurance programs targeted to needy children from other states that managed to increase insurance rates while keeping costs low, whereas Gilardi (2010) found that, under some circumstances, policy makers are more likely to imitate a policy if elsewhere it has been shown to be compatible with the reelection of those who enacted it. Nicholson-Crotty and Carley (2016) showed that policies are more likely to diffuse among two states if they share a similar implementation environment, suggesting the importance of this aspect of policy learning.

Second, in contrast to learning, emulation is not related to the objective consequences of a policy. Instead, the symbolic and socially constructed characteristics of policies are crucial here. Inspired by sociological institutionalism, the conceptualization of this mechanism implies that units have to conform to their normative environment. Thus, some policies will enjoy high acceptance, regardless of whether they “work” or not. By contrast, others will be taboo, even though they could possibly be beneficial. Another way to see this mechanism is that the “burden of proof” changes over time as a function of social acceptance. When considering a radical policy innovation, the burden of proof rests on its advocates, but when it becomes widely accepted, it is the opponents of the policy who have to make their case compellingly to prevent its adoption. For instance, F. S. Berry and Berry (1992) argued that, when introducing new taxes, state policy makers can reduce the perception that they are unfair by pointing to the existence of similar taxes in nearby states. At the international level, Greenhill (2010) argued that international governmental organizations (IGOs) enhance the spread of human rights by fostering the development of norms through the socialization of their members. In this view, the material consequences of respecting human rights carry less weight than the pressure for confirming with a norm within a given peer group.

Third, competition occurs when units react to or anticipate one another in the attempt of attracting or retaining resources. Tax competition is the prototypical example (Basinger and Hallerberg 2004; Cao 2010), but competitive dynamics can also be found in many other areas of economic policy, such as capital account and exchange rate policies (Simmons and Elkins 2004), bilateral investment treaties (Elkins, Guzman, and Simmons 2006), and market-oriented infrastructure reforms (Henisz, Zelner, and Guillén 2005). For instance, Simmons and Elkin (2004) found that a country is more likely to liberalize its international economic policies following similar reform among its competitors, defined as countries with which it shares similar trade relationships. Of course, similar dynamics also take place at the subnational level. The notion of “competitive federalism” describes an institutional setting in which “state and local officials determine their own policies in part based on competition with surrounding communities” (Volden 2002, 352). Tax competition is an
obvious case in point (Deskins and Hill 2010; Gilardi and Wasserfallen 2014), but competitive relationships exist in other areas as well. The concept of “welfare magnets” suggests that states cut social benefits to avoid attracting beneficiaries from other states (Schram and Krueger 1994; Volden 2002), whereas the “California effect” refers to the capacity of big states such as California to set stricter (environmental) standards and, by virtue of the size of their market, induce other states and even countries to follow suit (Vogel 1995). Interestingly, although crude theories of competition predict races to the bottom (in taxes, benefits, standards), the evidence is overwhelmingly against uniform downward convergence.

Although policy diffusion is essentially a horizontal phenomenon, in federal states the vertical relationship between states and the federal government plays a considerable role (Karch 2007b, 67–68). National statutes influence policy diffusion by affecting the internal factors that make policy adoption not only at the state level more or less likely (Karch 2006) but also at the agenda-setting stage, by increasing the political relevance of some issues over others. But there is also a bottom-up dimension to vertical diffusion: adoption at lower levels of governments (e.g., cities) can affect the likelihood that the policy is introduced at higher levels (e.g., states). The diffusion of environmental standards is a case in point: stricter rules first introduced in California were later adopted by the federal government (Vogel 1995). Shipan and Volden (2006) developed two hypotheses on bottom-up diffusion. On one hand, widespread adoption by cities can reduce the need for action at the state level (“pressure-valve effect”). On the other hand, it can make it necessary to harmonize or otherwise coordinate the diverse policies that have come into place at the local level (“snowball effect”). Interestingly, Shipan and Volden (2006) found that, in the case of antismoking legislation, the pressure-valve effect dominates when state legislatures have fewer resources, while the snowball effect takes over for more professional legislatures. In this issue, Karch and Rosenthal (2016) found that members of Congress representing a state where a policy is in place are more likely to sponsor or cosponsor bills to introduce it nationwide, although the effect subsides at later stages of the legislative process.

I argue that the diffusion literature in political science is, and should be, coalescing around these definitions. Therefore, conceptually speaking, what we talk about when we talk about diffusion is pretty clear. There is definitely room for some conceptual improvement, but not much. Most new conceptual distinctions are hairsplitting. Where we have real problems is with operationalization. A meta-analysis of over 100 diffusion studies has found that there is considerable confusion on what indicators are appropriate for the different mechanisms (Maggetti and Gilardi 2015). The same indicators are used for different mechanisms, and different indicators are used for the same mechanism. Table 1 shows how frequently six different indicators (geographic proximity, joint membership in organizations, policy success, structural equivalence in networks, the number of previous adopters of a policy, and trade flows) are used to measure three diffusion mechanisms (learning, emulation, and competition). For example, geographic proximity is a frequent operationalization of both learning and emulation, although the two mechanisms are conceptually very different. Similarly, structural equivalence is a common measure for both emulation and competition. In
the over 100 articles examined by Maggetti and Gilardi (2015), each indicator save policy success was used to measure each diffusion mechanism at least once. The measurement of diffusion mechanisms clearly needs stronger harmonization to generate more cumulative knowledge, but this does not necessarily require conceptual advances—just conceptual consistence.

In most cases, indicators based on mere geographic proximity are too blunt. Despite the practices documented in Table 1, there is consensus on this point:

[R]esearchers should be cautious to not simply accept the adoption of similar policies by neighbors as evidence of policy diffusion. (Volden, Ting, and Carpenter 2008, 328)

Although this “neighbors variable” specification of diffusion may successfully identify whether a policy diffuses, it clearly is incapable of determining why policies diffuse. (Baybeck, Berry, and Siegel 2011, 232)

While offering a good starting point, the classic view of policy diffusion as geographic clustering is often overly limiting, sometimes misleading (or even wrong), and increasingly outdated. (Shipan and Volden 2012, 789)

[A]lthough geographic contiguity represents a good first start, ties between neighboring states are not a comprehensive proxy for the policy diffusion network. (Desmarais, Harden, and Boehmke 2015, 397)

The operationalization of learning and competition is relatively straightforward, even though scholars have not been sufficiently consistent. A good indicator of learning must identify relevant outcomes in other jurisdictions: “scholarship providing evidence of a greater likelihood of policy adoption given earlier effectiveness elsewhere would lend convincing support to the concept of learning-based policy diffusion” (Volden, Ting, and Carpenter 2008, 327–28). For competition, the key is to identify which jurisdictions a given jurisdiction is competing with. The task is not trivial and
depends on the specific context, but scholars can rely on and adapt existing strategies (W. D. Berry and Baybeck 2005; Elkins, Guzman, and Simmons 2006; Gilardi and Wasserfallen 2014; Simmons and Elkins 2004; Wasserfallen 2014). By contrast, finding good measures for emulation is challenging because the concept is quite elusive. Without ruling out the possibility of good direct indicators, a more promising strategy for the analysis of emulation is indirect. That is, in many cases, emulation is best analyzed through its observable implications rather than with a direct measure. In particular, one implication of emulation is that the burden of proof shifts over time, which is reflected in political debates. This implication is hard to test with a traditional focus on policy adoption, which underscores the need for better research designs, as discussed in the “Upgrade Research Designs” section. Importantly, indirect strategies need to be pursued not only for emulation but also for other mechanisms when researchers lack adequate data (Volden, Ting, and Carpenter 2008, 328).

Be Clear about Goals: Learn about Diffusion or Use Diffusion to Learn about Something Else?

I argue that there are two types of research questions worth asking. First, we can try to make a contribution to the diffusion literature itself. Second, we can try to use the insights of diffusion research to learn something new about other phenomena. Usually, a project will pursue only one of these questions. Researchers should be clear which one it is.

Making a substantive contribution to the theoretical understanding of policy diffusion has become increasingly hard. The contribution of studies merely showing that a unit is more likely to adopt a policy of its neighbors or competitors is insufficient from this perspective. What is required is better, more focused questions, which themselves require better, more focused theory. As the building blocks of diffusion are fairly clear, theoretical advances should aim to explain more precisely how they operate in different contexts. Volden, Ting, and Carpenter (2008) made the important point that, when our theoretical expectations are too coarse, it is all too easy to mistake independent decision making for diffusion. Indeed, despite the plausibility of the assumption that policy making is interdependent, we should not forget that policy makers attempt to learn from their own experience (Callander 2011). Based on their formal model, Volden, Ting, and Carpenter suggested several hypotheses leveraging expectations on the conditionality of diffusion—for instance, that the longevity of a policy can be a proxy of its success but only in moderate states, which are less biased for or against it. Importantly, Volden, Ting, and Carpenter emphasized that when direct measures of policy success are not available or cannot be constructed, researchers should turn to indirect strategies. We will return to this point below. Baybeck, Berry, and Siegel (2011) carried out a similar theoretical exercise for competition, leading to a distinction between defensive, offensive, and anticipatory competition, giving much more leverage than the simple hypothesis that a state’s policies are influenced by its neighbors’ policies. In general, the characterization of the conditional nature of policy diffusion is a significant direction of the literature. Empirical studies have shown that
diffusion is conditional on the professionalism (Shipan and Volden 2006), expertise (Shipan and Volden 2014), gender (Bouche and Wittmer 2015), or preferences/ideology of policy makers (Butler et al. 2015; Gilardi 2010) as well as on the institutional structures shaping their interactions (Füglister 2012; Gilardi and Wasserfallen 2014) and the characteristics of the policies themselves (Makse and Volden 2011).

Although simple diffusion hypotheses based, in particular, on geographic proximity are usually insufficient for a significant contribution to the diffusion literature, often they can provide an original angle for other questions that previously have not been approached from a diffusion perspective. For instance, the literature on women’s representation has identified many types of spillovers, but it has not looked at cross-unit influences (an exception is Broockman 2014). A diffusion analysis shows that the number of women candidates in one unit increases with the number of women elected in other units, but only in the early stages of women’s representation (Gilardi 2015). This is not only because women’s participation in politics has become increasingly accepted as normal over time but also, more specifically, because cross-unit influences occur only when no female incumbent is running for reelection (Gilardi 2015). Since over time there are fewer and fewer such units, the scope for cross-unit effects becomes smaller. Such an analysis does not improve on our understanding of diffusion processes, but, through the diffusion perspective, it offers new insights on a problem that previously had not been approached from this angle.

A further use for a simple geographic spatial lag could be as an instrumental variable. Even when they are theoretically uninteresting or blunt, geography-based spatial lags are usually effective predictors of policies, which is an important requirement for employing them as instrumental variables. For instance, like many policies, state lotteries are geographically clustered (W. D. Berry and Baybeck 2005; F. S. Berry and Berry 1990). If we are interested in estimating the effect of state lotteries on some outcome of interest (e.g., state finances), endogeneity is a concern because states with poorer finances might be more likely to rely on lotteries to boost revenue. Thus, one could use the share of neighboring states where state lotteries are in place as an instrument. However, for this strategy to be valid, two assumptions need to be met. First, the instrument must affect the outcome only through its correlation with the treatment (“exclusion restriction”). In our example, this assumption is not unproblematic, given that, when lotteries are available in neighboring states, a state will lose some revenue to them if residents do not gamble exclusively in their home state. Second, the instrument must be as if at random (“ignorability assumption”). However, often the internal determinants of policies are themselves geographically clustered. This is a problem for the ignorability assumption, in the same ways as, more generally, it makes it difficult to disentangle diffusion from common pressures. In sum, using spatial lags as instrumental variables seems far from straightforward, but a more systematic discussion would definitely be helpful.

To sum up, diffusion scholars have two options: improving the understanding of diffusion itself or using the diffusion perspective to improve the understanding of some other phenomenon. Researchers should state clearly what their aim is. Both are
valuable. While the latter can be achieved with traditional hypotheses and methods, the former requires significant theoretical or methodological innovation.

**Upgrade Research Designs**

Standard research designs have almost fulfilled their potential. Spatial econometric models where the key variable is a spatial lag, that is, a weighted average of policies in other units (Franzese and Hays 2007; 2008), have been used extensively in comparative politics and international relations, while Americanists have been especially fond of the event-history approach introduced by F. S. Berry and Berry (1990). There are good reasons for these choices: these are good ways to show that something diffuses. But to push things forward, we need to go beyond these standard designs.

First, research designs should be tailored to the specific questions asked. This truism applies to any area, of course, but the problem seems particularly acute in diffusion research. There is a clear template with which we can study almost anything, so applying it to an important new topic (e.g., marijuana liberalization or same-sex marriage) is relatively easy. This strategy can produce interesting insights on the phenomenon that is analyzed, but it is unlikely to make a significant contribution to diffusion theory. Instead, scholars should develop and test more specific and interesting hypotheses, for instance, related to the conditional nature of policy diffusion, as discussed earlier. Moreover, whenever adequate indicators cannot be constructed, instead of using direct but blunt measures such as geographic proximity, researchers should explore ways to test diffusion hypotheses indirectly by focusing on their observable implications, such as the fact that, under the influence of emulation, the ease with which a policy finds acceptance changes over time and as a function of adoption elsewhere. For Americanists, the diffusion network measures computed by Desmarais, Harden, and Boehmke (2015) promise to be a catch-all diffusion indicator superior to geographic proximity.

Second, we need better data. Often this means moving away from cross-national or cross-state analyses, which is a trend in political science in general. Readers of this journal probably need little convincing that, in many cases, subnational or substate units offer data of higher quality, are more comparable, and permit a more reliable coding of key variables (Snyder 2001). However, we should also think creatively about data sources. Most diffusion studies focus on the decision-making stage of the policy cycle, while little is known about how other stages are influenced by diffusion (Shipan and Volden 2012, 793). Karch (2007a) and Pacheco and Boushey (2014) are exceptions, showing that diffusion takes place already at the agenda-setting stage, when bills are first introduced. In this issue, Karch and Rosenthal (2016) found some evidence of bottom-up diffusion at early stages of the legislative process. The problem-definition stage is likely to be equally relevant. It is very plausible that the way a given policy problem and potential solutions are framed and debated is shaped by the experience of other jurisdictions. Automated text analysis is making it increasingly feasible to measure this dimension quantitatively and use it in a diffusion analysis (Gilardi, Shipan, and Wüest 2015). Similarly, diffusion studies could consider other stages, such as the implementation or enforcement of policies (Shipan and Volden...
Third, whenever possible we should carry out “placebo tests,” that is, look for diffusion where there should be none. For instance, Lloyd, Simmons, and Stewart (2012) argued that human trafficking laws spread through policy externalities linked to the diversion of trafficking flow, and they found evidence of diffusion among countries sharing many road connections, which are the main channel through which people are trafficked across international borders. As a placebo test, they then checked whether roads also matter for other closely related policies that, however, do not involve externalities (victim protection), do not involve externalities via surface traffic (money laundering), or do not involve transnational externalities (internal trafficking). The null results for these tests strengthen the original argument. Similarly, in my study of the diffusion of women’s representation, I replicated the analyses using the number of male candidates as the dependent variable instead of that of female candidates (Gilardi 2015). The fact that I find diffusion effects for women but not for men is a strong confirmation of the argument that cross-unit spillovers pick up at least some of the effect of role models and are not due to some generic geographic clustering.

Fourth, we should take causal inference more seriously. This is a big trend (some would say fad) in political science right now, but one need not be an identification Taliban to say that very, very few diffusion studies in political science pay any attention at all to causal inference. Although in our context the problem is even thornier than usual—some say there is little (Angrist 2014) or no hope (Shalizi and Thomas 2011)—the status quo is not satisfactory, and we should do our best to improve on this front. Several options exist. First, experiments are an obvious strategy that has been underapplied to the study of policy diffusion. Butler et al. (2015) is an exception, relying on experiments embedded in a national survey of municipal officials. The social networks literature offers several examples of how experimental evidence of diffusion could be produced (Bond et al. 2012; Coviello et al. 2014; Kramer, Guillory, and Hancock 2014). Sinclair, McConnell, and Green (2012) put forward a multilevel experimental design to test spillover effects that could be adapted to study policy diffusion if a meaningful intervention can be rolled out in several locations, whereas Bowers, Fredrickson, and Panagopoulos (2013) developed a general framework to think about interference between units in experimental settings. Finally, a regression-discontinuity design (RDD) could be used to identify the effect of as-if-at-random decisions on other jurisdictions (Broockman 2014); Keele and Titiunik (2015) developed a geography-based variant of the RDD that has an obvious appeal in our context. Of course, a fruitful application of these strategies is much more difficult than it appears at first sight, but diffusion scholars should give these options careful consideration.

Be Explicit about the Practical Implications of Diffusion

Diffusion scholars are often interested in diffusion as such. But why should others care? Policy diffusion has important practical implications that diffusion research should emphasize more systematically.
First, a well-known tenet of theories of federalism is that decentralization allows local governments to become laboratories of democracies, fostering innovation and the diffusion of best practices (Kollman, Miller, and Page 2000; Oates 1999; Strumpf 2002; Tiebout 1956): “when subnational governments innovate, successful solutions can diffuse to other subnational states with similar preferences and problems” (Bednar 2011, 273). Diffusion research shares the basic premise but offers a more critical and nuanced view. The “states as policy laboratories” view conflates diffusion with learning, but we know that learning is only one possible mechanism through which policies spread. Others, such as emulation, but also competition, have much less clear implications for the desirability of diffusion. Policy diffusion is not always beneficial (Shipan and Volden 2012, 790–91). Work showing how the detrimental effects of diffusion can be limited, and how its positive effects can be stimulated, has an obvious practical interest.

Second, diffusion scholars could leverage the concept of “network interventions”—“the process of using social network data to accelerate behavior change” (Valente 2012, 49)—to highlight the potential of policy diffusion to speed up policy change. Similarly, by analogy with the “social multiplier” (Glaeser, Sacerdote, and Scheinkman 2002), scholars could emphasize the importance of the “diffusion multiplier”: by influencing one jurisdiction directly, many more jurisdictions can be reached indirectly. The methodology put forward by Desmarais, Harden, and Boehmke (2015) to infer diffusion networks from the adoption of many policies over time could be a useful tool to compute a diffusion multiplier. Proponents of marijuana liberalization and advocates of marriage equality are certainly interested in knowing which states should be persuaded to accelerate the spread of these policies, whereas groups opposing them would value insights on how to stop or slow down their diffusion. Diffusion scholars can improve the practical relevance of their work by considering how political actors can use their findings in their campaigns.

**Conclusion**

Diffusion research has diffused, within and across political science subfields, for a good reason: it focuses on a key question in social science—the nature and consequences of interdependence—that has direct implications for political phenomena in many areas, and especially for policy making in federal states. Research has established the existence and relevance of diffusion processes, but it needs improvement to go beyond this general result. I have put forward four strategies. First, use existing concepts consistently and focus on innovative measurement instead of new conceptualizations. Second, decide whether the goal of a specific project is to learn something new about diffusion itself or to use diffusion research to learn something new about another phenomenon. Third, refrain from using existing research templates. Instead, develop a research design tailored to the specific question, use better data, carry out placebo tests whenever possible, and pay more attention to causal inference. Fourth, discuss explicitly the practical implications of diffusion research. None of these suggestions is new. However, taken together, they trace a distinctive way forward for policy diffusion research. I will do my best to follow it, and I hope that others will, too.
Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Because these effects do not operate at the level of the jurisdictions among which the policy is spreading, they do not entirely fit the definitions above.
2. I thank an anonymous reviewer for this suggestion.

References


Author Biography

Fabrizio Gilardi is professor of Public Policy in the Department of Political Science of the University of Zurich, Switzerland and an editor of the Journal of Public Policy. His work on regulatory institutions, policy diffusion, and women’s representation has been published in journals such as the American Journal of Political Science, the British Journal Political Science, and Comparative Political Studies, among others.