Implementing and complying with EU governance outputs

Oliver Treib
University of Münster,
Department of Political Science,
Scharnhorststr. 100,
48151 Münster, Germany
email: oliver.treib@uni-muenster.de

Accepted: 25 July 2014
Published: 28 August 2014
(Update of lreg-2008-5)

Abstract

This essay takes stock of the literature on how European Union policies are being put into practice by the member states. It first provides an overview of the historical evolution of the field. After a relatively late start in the mid-1980s, the field has meanwhile developed into one of the growth industries within EU research. The paper identifies four waves of EU implementation scholarship, each with its own theoretical, empirical and methodological focus. In the second part, the review discusses the most important theoretical, empirical and methodological lessons to be drawn from existing studies. Four conclusions emanate from the analysis of existing EU implementation research. First, the literature has focused heavily on the transposition of EU directives, while comparatively little is known about issues of enforcement and application of both directives and regulations or about member states’ reactions to negative integration. Second, scholars studying the transposition of directives seem to agree that we need to address factors that influence member states’ willingness and capacities to comply. The main task to be accomplished by future research is to establish under what conditions which configurations of factors prevail, especially with regard to sectoral differences. Third, more energy needs to be devoted to systematic research on the phase of practical implementation, and this research should make more use of theoretical insights from domestic implementation research as well as from management and enforcement approaches. Fourth, quantitative transposition research will have to improve the data it uses to measure the dependent variable. Scholars should explore better data sources and invest more energy in collecting their own data on transposition timing and correctness. Research on application and enforcement, on the other hand, needs to go beyond case studies and instead search for or produce data with which the practical phase of implementation can be analysed on a broader, more comparative scale.

Keywords: implementation, European law, directives, regulations, policy analysis, Europeanisation, methodological issues
Imprint / Terms of Use

*Living Reviews in European Governance* is a peer reviewed open access journal published by the Institute for European Integration Research (EIF), Strohgasse 45/DG, 1030 Vienna, Austria. ISSN 1813-856X.

This review is licensed under a Creative Commons Attribution-Non-Commercial 3.0 Austria License: [http://creativecommons.org/licenses/by-nc/3.0/at/](http://creativecommons.org/licenses/by-nc/3.0/at/) Figures that have been previously published elsewhere may not be reproduced without consent of the original copyright holders.

Because a *Living Reviews* article can evolve over time, we recommend to cite the article as follows:

```
Oliver Treib,
“Implementing and complying with EU governance outputs”,
*Living Reviews in European Governance*, Vol. 9, (2014), No. 1:
http://www.livingreviews.org/lreg-2014-1
(accessed <date>).
```

The date given as <date> then uniquely identifies the version of the article you are referring to.
Article Revisions

*Living Reviews in European Governance* supports two methods of keeping its articles up-to-date:

**Revisions** consist of minor editorial changes and allow the author the opportunity to update the article according to current research findings, new publications or significant developments. These include, for example, the addition of new references or URLs. The revised article is reviewed by the Section Editor and an overview of changes will be listed through a ‘change log’ at the beginning of the article.

**Major Updates** involve extensive changes that require a substantial revision of the article’s content, or the addition of new ideas that were not in the original text. Major updates are subject to full external reviewing and are published with a new publication number. Access to the original text is available only through the updated version of the article.

For detailed documentation of an article’s evolution, please refer to the history document of the article’s online version at [http://www.livingreviews.org/lreg-2014-1](http://www.livingreviews.org/lreg-2014-1).

**28 August 2014:** Major revision, updated and expanded. In particular, a new Section 2.4 was added and Section 3 substantially revised and extended. The number of references increased from 143 to 223.
Contents

1 Introduction: What is implementation and why should we care about it? .......................... 5

2 The emergence and evolution of EU-related implementation research .......................... 7
2.1 The first wave: Implementation and institutional efficiency ........................................ 7
2.2 The second wave: Misfit and more .............................................................................. 8
2.3 The third wave: Coming to grips with diversity ......................................................... 10
2.4 The fourth wave: Overcoming the focus on positive integration and the divide between EU and domestic policy-making .......................................................... 13

3 Discussion: What have we learnt? ............................................................................. 15
3.1 Instruments, implementation phases, countries and policy areas studied ..................... 15
3.2 Measuring compliance ................................................................................................. 17
3.3 Theoretical insights I: Transposition ......................................................................... 20
3.3.1 EU-level factors ..................................................................................................... 20
3.3.2 Domestic factors related to member states’ willingness to comply .......................... 22
3.3.3 Domestic factors related to member states’ capacity to comply ............................... 25
3.3.4 Sectoral patterns .................................................................................................... 27
3.3.5 Country patterns .................................................................................................... 27
3.4 Theoretical insights II: Application and enforcement ............................................... 29

4 Conclusion ................................................................................................................... 31

References ..................................................................................................................... 33
Implementing and complying with EU governance outputs

1 Introduction: What is implementation and why should we care about it?

This essay takes stock of the literature on how European Union policies are implemented by the member states. The notion of ‘policy implementation’ is tied to what has been called the “textbook conception of the policy process” (Nakamura 1987: 147). This conception assumes the policy cycle may be divided into several clearly distinguishable phases, ranging from problem definition and agenda-setting to policy formulation, policy implementation, evaluation and finally to policy termination or re-formulation. Policy implementation thus refers to “what happens after a bill becomes a law” (Bardach 1977) or, as one scholar aptly put it, to the process of “translating policy into action” (Barrett 2004: 251).

A similar but slightly different concept is the notion of ‘compliance’. It has been prominent in international relations research among scholars studying the domestic fulfilment of international agreements (for an overview, see Raustiala and Slaughter 2002). Compliance refers to “a state of conformity or identity between an actor’s behavior and a specified rule” (Raustiala and Slaughter 2002: 539). Thus, the compliance perspective also starts from a given norm and asks whether the addressees of the norm actually conform to it. It thus focuses less on the process than on the outcome of implementation. Moreover, compliance can occur without implementation. For example, the current practice may already conform to the status quo required by a norm. Conversely, implementation does not necessarily have to result in compliance but may be incomplete or contrary to the prescribed goals (Raustiala 2000: 391–399). Irrespective of these semantic differences, most compliance and implementation research is interested in both the process of how a given norm is being put into practice and in the outcome in terms of rule conformity. In this sense, implementation and compliance are two sides of the same coin.

A third concept that is often used in the context of EU implementation research is the notion of ‘Europeanisation’, which points to the effects of European integration in the member states. While both concepts and the respective bodies of literature associated with them have a considerable degree of overlap, it is important to keep in mind that there are also important analytical differences. The implementation of EU legislation usually entails certain policy or institutional changes at the domestic level. In this sense, implementation is one important mechanism of Europeanisation. However, the domestic effects of European integration are not confined to processes of policy transfer and the institutional adaptations associated with these. Instead, European integration may have many other, direct or indirect, intended or unintended, effects at the domestic level (Töller 2004). In this sense, Europeanisation research encompasses, but is broader than, EU implementation studies.

The following discussion will not focus on the wider field of Europeanisation research, as there are separate Living Reviews on “The Europeanisation of national political systems” (Goetz and Meyer-Sahling 2008), on “Europeanization of political parties” (Ladrech 2009) and on “Europeanization beyond Europe” (Schimmelfennig 2012). In contrast, the current Living Review will deal with insights gained from implementation studies, focusing on both ‘old’ and ‘new’ member states. In this sense, it partly overlaps Ulrich Sedelmeier’s Living Review on “Europeanisation in new member and candidate states” (Sedelmeier 2011).

Why should social scientists care about implementation? The answer is straightforward: putting a piece of legislation or a government programme into practice does not happen automatically, nor is it a purely technical or apolitical affair. Instead, long delays and attempts at shirking seem to be a matter of everyday business in the field of implementation. In other words, if we are interested in the extent to which a particular polity is able to solve the problems with which it is confronted, we need to study not only the way it reaches decisions and the character of the resulting legal output, but also the way in which the law is executed in practice.

This is particularly true in a large and complex polity like the European Union. Due to the
high number of veto players involved in policy formation, EU legislation often contains fuzzy concepts and ‘rhetorical compromises’ in order to facilitate agreement. Moreover, EU legislation regularly leaves certain issues to the discretion of member states in order to take account of specific regional or local circumstances. In other words, crucial decisions that may decide on the success or failure of a particular policy are regularly taken at the implementation stage. What is more, it is far from self-evident that implementers will behave dutifully. The EU is marked by a highly decentralised implementation structure. It does not have its own administrative machinery to implement its legislation, but has to rely on the member states to fulfil this task. If we focus on the implementation of EU directives, one of the major legal instruments of the EU, even parts of the decision-making process are delegated to the domestic level. The provisions of directives have to be incorporated into national law by member states within a certain period of time. It is only after transposition has been completed that the rules may be applied by administrations and societal target groups and enforced by administrations and the legal system at the domestic level (see Figure 1).

![Figure 1: Stages of the Implementation Process](http://www.livingreviews.org/lreg-2014-1)

Directives thus share important characteristics of traditional international agreements which need to be ratified by member states before they become effective. Unlike international agreements, however, the implementation of EU law by member states is supervised by the European Commission and the European Court of Justice (ECJ). If the Commission detects violations of EU law, it can instigate legal proceedings against the respective member state. These proceedings may ultimately lead to a judgement by the ECJ and, in the event of continuing non-compliance, to follow-up proceedings which may result in financial fines imposed by the Court. However, there is no European police force that could compel member states to obey the rules. As a result, the level of legal obligation and the extent of actual enforceability place EU law somewhere between traditional domestic law on the one hand and traditional international law on the other.

This makes the European Union a particularly interesting object of study for implementation researchers. After a relatively late start, the literature on how EU policies are put into practice domestically has in fact proliferated over the last three decades. This Living Review offers a
structured overview of EU implementation studies. Further synopses of the growing literature are provided by Mastenbroek (2005); Knill (2005); Sverdrup (2007); Versluis et al. (2011); Angelova et al. (2012) as well as two comprehensive databases that seek to take stock of quantitative and qualitative studies in the field (Toshkov 2014; Toshkov et al. 2014). The next Section 2 provides an overview of the emergence and evolution of the field. It identifies four distinct waves of research each with its own theoretical, empirical, and methodological focus. Afterwards, the paper discusses the most important theoretical, empirical and methodological lessons to be drawn from existing studies and highlights promising avenues for future research. The final section concludes by summarising the main findings.

2 The emergence and evolution of EU-related implementation research

Scholars studying European integration, like their colleagues interested in domestic politics, have long been preoccupied with issues of policy formation and decision-making, thus neglecting the question of how policies are being put into practice. At both levels, it was ambitious legislative reform initiatives that spurred interest in policy execution. ‘Classical’ domestic implementation research had its starting point mainly in two countries: the United States and Germany. In the US, Lyndon B. Johnson’s ‘Great Society’ project of the 1960s, a package of federal initiatives aimed at combating poverty and racial discrimination, fuelled a set of research projects on the implementation of federal programmes (see, e.g. Derthick 1972; Pressman and Wildavsky 1973; Bardach 1977). In Germany, the same effect was brought about by the bold reform initiatives of the grand coalition and the ensuing social-liberal government in the late 1960s and 1970s Mayntz (1977, 1979, 1980, 1983); Scharpf (1978). Starting from these pioneer studies, domestic implementation research has produced a raft of mainly case study based contributions. Most of this research revolved around the cleavage between two schools of thought: the top-down approach, which conceived of implementation as hierarchical execution of centrally-defined policy intentions, and the bottom-up camp, which emphasised instead that policies were decisively shaped by the everyday problem-solving strategies of the actors involved in policy delivery. A third group of scholars tried to bridge the gap between these opposing approaches by combining insights from both sides (for an overview, see Pütlz and Treib 2007).

2.1 The first wave: Implementation and institutional efficiency

It was the Single Market Programme that acted as a stepping stone to implementation studies in the EU context. The programme involved a raft of legislative measures whose consistent implementation was seen as a precondition for the completion and smooth functioning of a Europe-wide market. In the mid-1980s, these concerns gave rise to the first wave of EU-related implementation research. In theoretical terms, the main inspiration came from domestic implementation studies, most importantly from the top-down school (Pressman and Wildavsky 1973; Bardach 1977; van Meter and van Horn 1975; Sabatier and Mazmanian 1981; Mazmanian and Sabatier 1983). First-wave studies thus portrayed the domestic implementation of European law as a rather apolitical process whose success primarily depended on clearly worded provisions, effective administrative

1 It should be noted that the waves I describe in the following are to some extent characterised by what has become known in historical research as “the simultaneity of the nonsimultaneous” (Koselleck 2002: 168), i.e., the fact that phenomena may occur earlier or later in time than the broader societal processes to which they belong. In our context, this means the waves describe typical clusters of studies that emerged and became prominent during a certain period of time. Often, however, individual studies with a very similar theoretical or methodological orientation were published earlier. Likewise, we sometimes find studies that systematically belong to an earlier wave published in a later wave of research.
organisation and streamlined legislative procedures at the domestic level. At the same time, they also absorbed some of the insights of the bottom-up camp (Lipsky 1980; Hjern and Porter 1981; Elmore 1982), stressing the need for involving all relevant domestic actors (such as parliaments, important interest groups, or subnational entities) in the preparation of the countries’ European negotiating position and for co-ordinating the negotiation and implementation tasks within domestic administrations, ideally by attaching responsibility for both phases of the policy cycle to one person (Ciavarini Azzi 1985; Krislov et al. 1986; Siedentopf and Ziller 1988; Schwarze et al. 1990, 1991, 1993; From and Stava 1993). The absence of a ‘political’ conceptualisation of the implementation process among first-wave studies to some extent may be explained by the disciplinary background of the authors, who mainly came from legal studies and administrative science. Scholars from these fields continued to publish contributions with a first-wave focus even after the research mainstream had moved on to other theoretical shores (Demmke 1994, 1998, 2001; Pappas 1995; van den Bossche 1996; Ciavarini Azzi 2000; Bursens 2002).

Most of the first-wave studies covered transposition as well as application and enforcement. However, the authors did not draw a sharp distinction between legal incorporation and the later stages of the implementation process. Instead, the main explanatory variables for all stages were clearly stated policy objectives and the availability of a well-organised state apparatus. With regard to enforcement and application, the main conclusion was that “Community law, once it has been incorporated, is applied neither better nor worse than national law” (Ciavarini Azzi 1988: 199) since “street-level bureaucrats” (Lipsky 1980) and target actors are usually unaware of the European origins of a particular transposition law. However, the analysis of the domestic implementation of regulations revealed that specific problems occurred since the one-size-fits-all rules enshrined in EU regulations could not be adapted to specific domestic circumstances and traditions (Ciavarini Azzi 1988: 199).

2.2 The second wave: Misfit and more

In the late 1990s, a second wave of studies began to analyse the “Europeanisation” of domestic political systems. This broader perspective has hitherto produced a host of contributions dealing with the impact of membership in the European Union on such phenomena as national parliaments, party systems, state-society relationships, territorial state structures, or democratic structures of government (for an overview, see the Living Reviews by Goetz and Meyer-Sahling 2008; Ladrech 2009, and Schimmelfennig 2012). In this context, scholars have also returned to the narrower question of the domestic impact of European policies, as witnessed by the national implementation of European policy measures.

Focusing mainly on environmental policy, many of the second-wave scholars pointed to the degree of fit or misfit between European rules and existing institutional and regulatory traditions as one of the central factors determining implementation performance (Duina 1997, 1999; Duina and Blithe 1999; Knill and Lenschow 1998, 2000a; Börzel 2000, 2003a). The focus thus moved from administrative and procedural efficiency to the degree of compatibility between EU policies and domestic structures. This view ultimately rests on historical institutionalist assumptions about the ‘stickiness’ of deeply entrenched national policy traditions and administrative routines, which pose great obstacles to reforms aiming to alter these arrangements (see, e.g., Thelen and Steinmo 1992; Immergut 1998; Pierson 2000).

The basic rationale behind the misfit argument was to reduce the complexity of analysing implementation processes by exploring how far the “institutional filter” (Knill and Lenschow 1998: 610) provided by the compatibility between EU demands and domestic policy traditions alone could explain the implementation of particular pieces of EU legislation. The assumption was that further actor-based factors needed to be taken into account only if the institutional context was not able to explain the outcomes (Knill and Lenschow 1998: 610–611, Knill and Lenschow 2001).
The main problem with this approach was that only a few cases could actually be explained by an exclusive focus on the ‘goodness of fit’. In the empirical analysis by Knill and Lenschow (1998), only three out of eight cases conformed to the expectations gained by looking at the degree of misfit. The remainder of the cases needed to be explained by additional actor-based factors. Later studies confirmed the limited explanatory power of the ‘goodness of fit’ (see, e.g., Haverland 2000; Héritier et al. 2001; Falkner et al. 2002, 2004, 2005; Mastenbroek and Kaeding 2006). In the end, therefore, it turned out that not much analytical leverage could be gained from using this ‘institutional filter’. Instead, most cases required “a lower level of abstraction, namely the independent analysis of the given interest constellations and the strategic interaction of domestic actors” (Knill and Lenschow 2001: 126).

One of the main weaknesses of the second wave of research was that the logic of these interactions, and the preferences of domestic actors, remained seriously under-theorised. Instead, authors often offered individual accounts of the ‘deviant cases’ without offering systematic theoretical arguments. What is more, different contributions implicitly operated on the basis of quite divergent views. The misfit argument in principle implied that domestic governments and administrations are motivated by the desire to protect their existing policy legacies and are thus expected to drag their heels on fulfilling EU policies that require fundamental changes to the domestic status quo. When having to implement European policies, national governments, administrations, and parliaments are thus seen to act as “guardians of the status quo, as the shield protecting national legal-administrative traditions” (Duina 1997: 157). This view was based on the insights of earlier research on EU decision-making, which demonstrated that domestic governments try to export their own policy models to the European level (Héritier 1996; Héritier et al. 1996). As a result, it was argued that governments who failed to ‘upload’ their own policies to the EU level would try to resist during the ‘downloading’ process, when the agreed-upon measures were to be implemented (Börzel 2002). Therefore, the implementation of policies with significant misfit was either doomed to fail altogether due to reluctant domestic governments and/or administrations (Duina 1997; 1999; Duina and Blithe 1999; Knill and Lenschow 1998, 2000a), or the unwilling state machinery needed to be forced by societal actors to comply with mismatching EU policies, probably combined with outside pressure from the Commission (Börzel 2002, 2003a, 2006). Some of the misfit-centred contributions, however, also argued that a low number of veto players or, alternatively, a consensual political culture, could help overcome resistance against EU policies that implied significant adjustment costs (Risse, Cowles, and Caporaso 2001; Héritier 2001; Héritier and Knill 2001). While the veto player argument was usually presented in opposition to the misfit approach (Haverland 2000), this view tried to combine both factors. It still subscribed to the basic idea that the degree of misfit was an important determinant of implementation outcomes, with mismatching policies provoking fierce domestic opposition. In contrast to scholars like Börzel, Duina, or Knill and Lenschow, however, this approach implicitly assumed that resistance did not stem primarily from governments and administrations, but from negatively affected societal interests. The number of veto players then determined whether it was likely that these reluctant societal actors would be able to impede implementation or not. This approach represented a big step away from the mechanistic conception of the basic misfit argument, laying much more emphasis on the political contestation between reform promoters and opponents at the domestic level.

Like the first wave of research, many of the second-wave contributions analysed not only legal but also practical implementation (see, e.g. Knill and Lenschow 1998, 2000a; Duina 1999; Knill 2001; Börzel 2003a; Bailey 2002). Just as their predecessors, however, the second-wave authors did not systematically distinguish between factors that influence transposition and causal conditions that have an impact on enforcement and application. Typically, these contributions tended to treat the whole process of implementation as following a single theoretical logic in which the ‘goodness of fit’ played an important role. This also meant they ignored the different actor constellations in
the different phases. Thus, scholars frequently referred to opposition to mismatching policies by “public administration” (Börzel 2000: 142) or to “administrative resistance” (Knill and Lenschow 2000a: 161) as the main reason for problems in the implementation process as a whole. At the transposition stage, however, administrations are certainly not the only crucial actors. Instead, government representatives and political parties seem to be at least as important in a process that differs from regular domestic law-making only in that it is substantively constrained by EU framework legislation. This either means these authors had a rather bureaucratic conception of the transposition process, that they simply failed to acknowledge that the different stages involve different actors and thus also require different explanatory models, or that they considered the main problems to occur not during transposition but at the enforcement and application stage, where administrations play a crucial role. At any rate, little could be learnt from this literature about the specific problems associated with transposition, enforcement and application.

2.3 The third wave: Coming to grips with diversity

It is conceptual shortcomings of this type in conjunction with a growing uneasiness about the relatively narrow theoretical and empirical focus of earlier research that gave rise to the third wave of EU implementation studies. It is marked by a plurality of theoretical and methodological approaches, and by highly divergent empirical findings, ranging from highly politicised to largely bureaucratic conceptions of the domestic transposition process. In order to come to grips with this diversity, third-wave scholars began to explore factors at the sectoral as well as the country level that could explain the rather incongruous empirical findings.

A first new development was that qualitative researchers in particular began to discover the importance of domestic politics in determining the speed and correctness of legal adaptation to European directives. Along these lines, Treib (2003a,b, 2004) showed that party political preferences of governments may have a decisive impact on transposition outcomes. In particular, he took issue with the behavioural assumptions underlying the mainstream misfit argument; notably, that domestic governments will always try to defend their existing policy traditions. Instead, he showed how governments may well accept wide-ranging deviations from the status quo if the direction of the required reforms is in line with their party political preferences. Conversely, government parties may also drag their heels on the realisation of rather minor adaptations if these modifications go against the grain of their party political goals. Based on an advocacy-coalition approach to transposition, Bähr (2006) advanced a similar argument, pointing out that the preferences of crucial domestic actors and their institutional positions in the decision-making process, rather than the degree of compatibility between European and domestic policies, are key to understanding the legal incorporation of EU legislation. Similarly, Mastenbroek and Kaeding (2006) argued that EU implementation research needs to go “beyond the goodness of fit” and instead focus more thoroughly on the preferences of crucial players in the domestic political arena (see also Mastenbroek and van Keulen 2006). The second remarkable development in the third wave was the growing popularity of quantitative studies. Thus, more and more scholars started to analyse the data on the Commission’s infringement proceedings against member states (Mbaye 2001; Börzel 2001, 2003b; Tallberg 2002; Sverdrup 2004; Beach 2005; Perkins and Neumayer 2007) and on the transposition measures member states officially notified to the Commission (Lampinen and Uusikylä 1998; Bergman 2000; Giulian 2003; Borghetto et al. 2006; Berglund et al. 2006; Kaeding 2006, 2007; Mbaye 2001; 2003b; Haverland and Romeijn 2007; Lino 2007; Toshkov 2007a). The second remarkable development in the third wave was the growing popularity of quantitative studies. Thus, more and more scholars started to analyse the data on the Commission’s infringement proceedings against member states (Mbaye 2001; Börzel 2001, 2003b; Tallberg 2002; Sverdrup 2004; Beach 2005; Perkins and Neumayer 2007) and on the transposition measures member states officially notified to the Commission (Lampinen and Uusikylä 1998; Bergman 2000; Giulian 2003; Borghetto et al. 2006; Berglund et al. 2006; Kaeding 2006, 2007; Mbaye 2001; 2003b; Haverland and Romeijn 2007; Lino 2007; Toshkov 2007a).

2 For a critical discussion of this article from the perspective of the goodness of fit approach, see Duina (2007). See also the response by Mastenbroek and Kaeding (2007).

3 One of the earliest quantitative studies using official transposition data was presented by Mastenbroek (2003). Although her study on the Netherlands did not use official EU data but domestic legal databases, her methodological

Living Reviews in European Governance
http://www.livingreviews.org/lreg-2014-1
Implementing and complying with EU governance outputs

In theoretical terms, many of these quantitative contributions were informed by compliance approaches developed in the international relations literature. These approaches revolve around the dichotomy between voluntary and involuntary non-compliance. Scholars stressing problems of voluntary non-compliance argue that the willingness of states to comply with international commitments depends on the domestic costs and benefits of adaptation and on the costs of defiance. Where the costs outweigh the benefits, states will try to evade these burdens by non-compliance. Therefore, effective monitoring and sanctioning by international supervisory authorities is required to force unwilling states into compliance. This approach is known as the enforcement approach (Downs et al. 1996). The management approach, in contrast, considers lacking administrative and financial capabilities at the domestic level or ambiguous norms as the main sources of non-compliance. International organisations thus need to assist their members by organising training programmes and by providing assistance such as financial aid (Chayes and Chayes 1993).

The theoretical insights of quantitative studies in the third wave were rather diverse. Among the few factors that seemed to find support in several quantitative analyses in the third wave are the various aspects of administrative capabilities such as administrative capacity (Mbaye 2001, Berglund et al. 2006, Hille and Knill 2006, Kaeding 2006, Haverland and Romeijn 2007, Linos 2007, Perkins and Neumayer 2007, Berglund 2009, Knill and Tosun 2009, Borzel et al. 2010) or administrative experience with transposing EU law (Berglund et al. 2006, Kaeding 2006, Berglund 2009, Steunenberg and Rhinard 2010, Haverland et al. 2011).

As a result, two divergent developments could be observed in the third wave: while qualitative scholars increasingly stressed the political character of transposition, supporting the insights of the enforcement approach in compliance research, the results of quantitative research pointed towards the management approach, highlighting the importance of efficient and well-co-ordinated administrations with skilled personnel. Against the background of these divergent findings, scholars started searching for factors that could explain why transposition in some cases seemed to be a highly contested political affair while in other cases it was conducted in a depoliticised, bureaucratic mode.

The first of this type of explanation was offered by Bernard Steunenberg and his colleagues. They presented a policy-specific procedural model of transposition, which is based on the idea that the actor constellation relevant for transposition varies between individual cases. These may range from a politicised legislative mode to a more technical bureaucratic mode (Steunenberg 2006, 2007, see also Dimitrova and Steunenberg 2000). The crucial factor in this model is the type of instrument that needs to be adopted to transpose a directive. The type of legal instrument (parliamentary legislation, decree, ministerial order, etc.) determines whether the actor constellation comprises the broad set of ministries, political parties and interest groups usually involved in enacting a piece of legislation or whether the process is determined by a smaller set of actors, or even by a single ministry, as is the case if a ministerial order is enough to transpose a directive.

This policy-specific model has found empirical support in an analysis of a large dataset comprising more than 1100 cases of transposition. These include four sectors, five member states and a time span of more than 20 years (Steunenberg and Rhinard 2010, see also Berglund et al. 2006, Kaeding 2006, 2007, 2008, Haverland and Romeijn 2007). In this view, the varying transposition performance across cases is heavily influenced by different legal requirements and national or sectoral traditions concerning the type of legal instrument to be chosen for transposition. Since many directives have a low political salience and are transposed by ministerial decrees with modest or no substantive involvement from parliaments or cabinets without the involvement of parliaments or cabinets, most transposition processes mainly depend on administrative capabilities, co-ordination

4 It should be noted, however, that there are also qualitative studies that use this kind of framework (see, e.g., Zürn and Joerges 2005, Hartlapp 2005a, 2007).
and experience. It is only in exceptional cases that transposition becomes politicised and political parties, interest groups and other political actors enter the scene. The main variation found in these studies is between different sectors, which have different typical profiles of more technical or more politicised directives and more autonomous ministerial or more politicised cabinet or parliamentary transposition procedures (Haverland et al. 2011).

An alternative argument about the circumstances under which more bureaucratic or more politicised transposition processes occur was offered by Gerda Falkner, Oliver Treib, Miriam Hartlapp and Simone Leiber. Based on a large qualitative study on the implementation of six directives from the field of EU social policy in the fifteen ‘old’ member states prior to Eastern enlargement, Falkner et al. (2005) presented a typology of three “worlds of compliance”, which rested on the idea that groups of countries are marked by the varying importance of a culture of compliance in member states’ political and administrative systems (see also Falkner et al. 2007c). In the “world of law observance”, which consists of the Nordic countries, the presence of a culture of respect for the rule of law among political and administrative actors usually ensures fast and correct transposition (Falkner et al. 2005: 317–341; see also Leiber 2005). In the “world of neglect”, the absence of such a culture in both the political and administrative systems typically leads to long phases of bureaucratic inertia and rather apolitical transposition processes. Greece, France or Portugal conform to this pattern (Falkner et al. 2005: 317–341; see also Hartlapp 2005a; Hartlapp and Leiber 2010). Lastly, in the “world of domestic politics”, administrations usually work dutifully, but since a culture of compliance is absent in the political realm, transposing EU law typically depends on the fit with the political preferences of government parties and other powerful players in the domestic arena. This is the largest cluster, involving countries like Belgium, Germany, the Netherlands, Spain, and the UK (Falkner et al. 2005: 317–341; see also Treib 2003a,b, 2004; Hartlapp and Leiber 2010).

The typology was later extended to the new member states from Central and Eastern Europe (CEE). Based on a qualitative study of compliance with three social policy directives in the Czech Republic, Slovakia, Hungary and Slovenia, Falkner and Treib (2008) identified a fourth world of compliance characterised by politicised transposition processes which, due to pressure arising from accession conditionality, were concluded rather swiftly and by serious shortcomings in actual policy delivery in terms of law enforcement and practical application. The authors dubbed this country cluster the “world of dead letters”.

The worlds of compliance typology inspired several of the quantitative compliance scholars to test the argument against some of the original ‘Complying with Europe’ data and against other data on transposition. These studies did not unearth much confirmatory evidence, however (Toshkov 2007a; Thomson 2007, 2009, 2010). Although the validity of the data used in these tests and the causal relevance of some of the statistical findings were called into question (Falkner 2007; Falkner et al. 2007a,b; see also Toshkov 2007b), these quantitative tests raised doubts about the explanatory leverage of the typology.

The considerable proliferation of studies dealing with transposition should not conceal the fact that third-wave research has also looked at the later stages of the implementation process. Compared to earlier research, however, studies covering not only transposition but also enforcement and application became a very small minority in the third wave. Among the few exceptions is a study by Versluis (2003, 2004, 2007), whose explicit focus was on the enforcement of two directives from the field of chemical safety in four countries. She discovered major enforcement problems in some of her cases and argued that issue salience is crucial in determining whether domestic inspectors take a particular directive seriously or whether they ignore it.

The study by Falkner et al. (2005) also included not only transposition but also enforcement and application. They presented a set of institutional conditions that determine the effectiveness of domestic enforcement systems and distinguished between different types of enforcement for different types of norms (Falkner et al. 2005: 33–40). Applied to the fifteen member states included...
in their study, they discovered major shortcomings in the enforcement systems of several countries (Falkner et al. 2005: 271–276). The follow-up study by Falkner et al. (2008) [see also Falkner and Treib 2008; Falkner 2010] on the implementation of social policy directives in four Central and Eastern European countries demonstrated how wide the gulf between legal compliance and practical application can be. Looking only at transposition would have led to the conclusion that the four new member states were significantly more compliant than their counterparts in the EU-15. Studying enforcement and practical application, however, revealed significant shortcomings in all four countries.

2.4 The fourth wave: Overcoming the focus on positive integration and the divide between EU and domestic policy-making

In the fourth wave of research, quantitative and qualitative research camps continued to move in different directions. Qualitative studies started to explore uncharted territory by addressing member states’ reactions to European Court of Justice preliminary rulings. Quantitative research, in contrast, continued to focus on the transposition of directives, but started to explore the impact of EU decision-making on domestic transposition.

It is surprising that it took so long for Europeanisation and implementation scholars to address the interesting issue of member states’ reactions to ECJ rulings, especially since there was already a substantial amount of empirical research on the quantitative significance of the preliminary ruling procedure and on the role of domestic courts in submitting references to the ECJ (for an overview, see Stone Sweet 2010). Research on the actual reactions of member states to ECJ case law was spurred by highly visible Court rulings on issues such as patient mobility or trade union rights, which seemed to have major repercussions on member states’ policy legacies in domains that had hitherto been sheltered from EU influence.

So far, two major insights emerge from this research. First, there seems to be a general tendency for member states to contain the effects of ECJ case law, as Lisa Conant (2002) suggested in what was probably the first systematic study on the implementation of ECJ preliminary rulings in the member states. These efforts by member states to limit the impact of the Court’s judicial activism often go along with a gradual process of “finetuning” (Obermaier 2009: ch. 9). After an initial landmark ruling, additional cases are referred to the Court. These proceedings provide a forum for reluctant member state governments to voice their concerns, either as parties to the proceedings or by submitting written observations to other proceedings. As a result, the Court often qualifies earlier, more radical doctrines by accepting certain exemptions, which then helps smooth out domestic adaptation processes (Obermaier 2009: ch. 9; Schmidt 2014).

Second, the high degree of legal uncertainty of ECJ case law for member states plays a major role in domestic reactions. While the legal implications of a preliminary ruling is straightforward for the member state from which the reference originated, it is much more open to interpretation whether and to what extent other member states are affected as well. Uncertainty about the exact legal consequences of ECJ rulings often allows member states to drag their heels on unwanted judicial activism. At the same time, legal uncertainty may also spur domestic adaptation processes. Michael Blauberger (2014) has called this mechanism “anticipatory obedience”: governments that do not comply with ECJ rulings are vulnerable to additional legal proceedings initiated by domestic actors who have an interest in reforms in line with the Court’s doctrines. To avoid such additional case law, governments sometimes decide to obey with ECJ case law even though legally speaking they might have also gotten away with remaining passive (see also Schmidt 2008; Obermaier 2008, 2009).

Quantitative research on the implementation of EU legislation took a different turn during the fourth wave. While most of the research in the first three waves focused on explaining implementation processes at the domestic level without much interest in the impact of EU decision-making
on the behaviour of political and administrative actors at the national level, more and more quantitative scholars now tried to overcome the divide between EU decision-making and domestic implementation. It is true that some earlier studies did look at the impact of EU decision-making on domestic implementation. However, these studies either had a rather narrow focus on administrative co-ordination [Siedentopf and Ziller 1988; Schwarze et al. 1990, 1991, 1993], or they concluded that features of the EU decision-making process had little value in explaining transposition problems (Falkner et al. 2004; Falkner et al. 2005: 277–280).

Quantitative studies in the fourth wave tried to tackle the problem in a more fundamental way, trying to establish the impact of EU decision-making on domestic transposition for a broader set of cases. The underlying idea was that transposition is a continuation of EU-level policy formation by other means so that knowledge about the EU decision-making process would allow for major insights into member states’ behaviour at the transposition stage. In other words, member states that oppose a directive in the Council are expected to be less compliant than member states that voted in favour of a piece of legislation (Thomson, Torenvlied, and Arregui 2007). Subsequently, directives whose adoption was highly contested, or directives adopted on the basis of qualified majority voting, should be fraught with more transposition problems than non-controversial directives or directives that were adopted on the basis of unanimity ( König and Luetgert 2009), and directives with more discretion should be easier to transpose since they leave member states more room for adapting EU policies to domestic circumstances than rather rigid directives (Thomson 2007, 2009; Thomson, Torenvlied, and Arregui 2007; Steunenberg and Toshkov 2009; Zhelyazkova and Torenvlied 2009).

The results of these studies have been mixed so far, with a lot of inconsistent and often even contradictory findings (see Section 3.3 below for more details). This may partly be due to the wide variety of data sources, operationalisation decisions, and empirical scopes of different studies. However, there are also two more theoretical reasons that may account for the inconsistent empirical picture. The first is that some of the effects of EU decision-making on domestic transposition could easily be construed in an opposite direction. For example, if the role of the Commission’s enforcement policy is taken into account, directives that were highly contested during policy formation could be expected to be monitored and enforced more tightly by the Commission, suggesting that transposition becomes more compliant (Zhelyazkova and Torenvlied 2009). The second reason lies in the fact that transposition may not be as tightly coupled to EU decision-making as assumed by this line of research. It is quite possible that not all governments are equally aware of potential problems a given directive may create at the domestic level because they do not involve all societal actors who might be negatively affected into the formation of their negotiation position or because they fail to discern that a given directive actually involves policy changes that may cause trouble domestically. Moreover, there may be a change of government between the adoption and transposition of a given directive, or transposition becomes controversial not because of the directive to be incorporated into domestic law but because of some other issues that are regulated in the same reform (Falkner et al. 2004).

As was already the case in earlier research on the implementation of EU legislation, fourth-wave scholarship scarcely tackled the extent to which member states not only transpose EU law but also how they enforce and apply it in practice. Notable exceptions to this general trend can be found in the field of environmental policy, where the especially wide-ranging procedural and organisational reforms required by the Water Framework Directive has spurred scholarly interest in issues of practical implementation (Uitenboogaart et al. 2009; Liefferink et al. 2011; Bourblanc et al. 2013), but there are also studies on the application and enforcement of other pieces of EU legislation in the field of environmental policy (Nimmo Smith et al. 2007; Buckley 2012). It should be noted, however, that especially the latter studies remain at a rather descriptive level, whereas research on the Water Framework Directive was so far only able to present first glimpses of practical implementation since the directive involves several stages of reform so that “[d]rawing
final conclusions is only possible after 2027” (Bourblanc et al. 2013: 1450). At any rate, Bourblanc et al.’s (2013) comparative analysis of how four member states implemented the Water Framework Directive suggested that the availability of strong implementation agencies that are supportive of the policy goals to be put into practice plays a key role in successful policy execution – a role that is more important than high ambitions at the transposition stage that are then not followed up by well-co-ordinated implementation. This finding echoes key insights gained by the top-down school in domestic implementation research (see, e.g., Mazmanian and Sabatier 1983).

An innovative line of research started to address the role of the ever-growing number of EU-level agencies in ensuring compliance. While some of them are primarily involved in policy development, agencies such as the European Maritime Safety Agency, the European Aviation Safety Agency and the European Railway Agency have been demonstrated to take an active part in improving domestic enforcement and application, which not only involves monitoring but also training and exchange of administrative experience (Groenleer et al. 2010; Gulbrandsen 2011; Versluis and Tarr 2013). The Solvit network, a network of national centres trying to resolve complaints lodged by citizens or companies about the possible misapplication of Internal Market rules, has been shown to serve similar purposes, with learning processes improving the effectiveness of individual network participants over time (Hobolth and Martinsen 2013). Research along these lines is highly welcome as it not only provides additional information on the severely under-researched issues of enforcement and application, but also links EU implementation studies to research on the emergent European administrative space (Curtin and Egeberg 2009; Rittberger and Wonka 2012; Trondal and Peters 2013).

A second innovative approach in the study of enforcement and application was presented by Gelderman et al. (2010), who analyse the reasons for practical non-compliance with EU public procurement legislation in the Netherlands. To gather information on practical compliance with the rules enshrined in EU public procurement directives, they conducted a survey among ‘users’ of the legal provisions, i.e., purchasing professionals in the Dutch Ministry of Defence, asking them about the extent to which they adhered to several key rules laid down in EU directives, and about a couple of items from which they constructed their explanatory variables. Although the overall outcome of this analysis seems rather modest – rule compliance seems to be influenced by the extent to which target actors expect gains or losses from compliance and by the extent to which the respective organisation exerts pressure to comply – employing a survey among users of the law seems a promising method for collecting systematic empirical information on the application and enforcement of European legislation.

3 Discussion: What have we learnt?

After this historical overview of how the field of EU implementation and compliance research has evolved over the past decades, this next section will discuss the main insights we have gained so far and will point to a number of problems and shortcomings that need to be addressed by future research.

3.1 Instruments, implementation phases, countries and policy areas studied

Most of the research hitherto has focused on the transposition of EU directives. Despite a few notable exceptions, the tendency to neglect issues of enforcement and application has even increased in the third and fourth waves of research. One reason for this seems to be a methodological one: as more and more scholars have turned to quantitative approaches, enforcement and application issues have taken a back seat since there is simply no appropriate quantitative data for analysing
the practical aspects of implementation. This is rather unfortunate, since all qualitative studies that did include issues of practical implementation have demonstrated that the “law in the books” is not necessarily the same as the “law in action” (Versluis 2004: 13). Along the same lines, it is remarkable that the implementation of EU regulations, which do not require domestic transposition, have so far attracted little interest (but see Siedentopf and Ziller 1988). From the standpoint of EU policy-making, for example, one very interesting question to be addressed would be whether the instrument of regulations works better than directives, since regulations offer less opportunities for policy drift and shirking or whether it is even the other way round, since regulations are too inflexible to adjust to the diverse institutional and societal conditions in a polity that comprises so many different societies, as the Siedentopf and Ziller project suggested (Ciavarini Azzi 1988: 199).

A second important aspect of EU compliance research so far is that it has predominantly dealt with measures of positive integration, where the EU defines certain policy goals or standards that member states are required to implement. Much less attention has been paid to member states’ reactions to negative integration, where the EU lays down rules that require member states to discontinue certain policies or to abstain from certain activities such as subsidising individual branches of the economy or to award certain public contracts without Europe-wide tenders. Although several authors have begun to study the implementation of ECJ rulings, which often fall in the realm of negative integration, there is still some way to go until we have understood properly the logic of complying with negative as opposed to positive integration. As negative integration requires member states to continually abstain from certain activities that would distort competition or restrict free trade, researchers usually cannot identify implementation dates or transposition rates (Schmidt 2008; Schmidt, Blauberger, and van den Nouland 2008; Blauberger 2009). This makes the study of compliance with negative integration more akin to studies on application and enforcement. Moreover, negative integration is often related to a considerable amount of legal uncertainty about which activities are and are not in line with relatively broad EU rules, often merely laid down in the treaties and interpreted by the Commission and the Court. Scholars have only just begun to have a systematic look at this feature and its implications for analysing and theorising compliance with rules of negative integration (Schmidt 2008; Schmidt, Blauberger, and van den Nouland 2008; Blauberger 2009).

A third interesting feature of EU compliance research is that qualitative studies especially have tended to overrepresent certain member states to the detriment of others that are only scarcely covered by case studies. While it is too soon to draw meaningful conclusions about the new member states from Central and Eastern Europe (Angelova et al. 2012: 14), there are significant differences in the coverage of the fifteen old EU member states. Apart from the study by Falkner et al. (2005), qualitative research has very rarely studied the Nordic countries, Luxembourg, Portugal, Austria or Greece, whereas Germany, the Netherlands, and the UK have been frequently analysed (Angelova et al. 2012). This uneven coverage of member states is much less true for quantitative studies, which can usually rely on data covering all member states. Yet, there are also quantitative studies that only look at individual countries (Mastenbroek 2003; Borghetto et al. 2006), and since important research consortia such as the one established in the Netherlands also analysed a selection of countries, the resulting pattern reveals countries like Germany, the Netherlands, and the UK are overrepresented in quantitative studies as well, whereas relatively few studies covered Austria or the Nordic countries (Angelova et al. 2012). Future research should thus try to cover as many countries as possible. If a selection has to be made, scholars should consider selecting countries that cover a wide range of possible country differences – be that size, wealth, political and administrative structures, legal traditions, length of EU membership or membership in different ‘worlds of compliance’ – thus ensuring that they take a deeper look not only at the usual suspects but also at smaller, less frequently covered countries.

A fourth remarkable aspect of existing EU implementation research is that studies with an explicit focus on cross-sectoral policy comparison are in short supply, so there is still a dearth...
Implementing and complying with EU governance outputs

of empirical insights on the systematic sectoral differences in implementation. Most qualitative studies to date have concentrated on one policy area. Particularly popular have been studies on environmental policy (Demmke 1998; Jordan 1999a; Knill and Lenschow 2000; Haverland 2000; Börzel 2003a; Bugdahn 2005; Slepcevic 2009; Bourblanc et al. 2013) and social policy, including labour law and gender equality (Hoskyns 1996; Duina 1997; Caporaso and Jupille 2001; Falkner et al. 2005; Dimitrova and Rhinard 2005; Falkner et al. 2008; Sedelmeier 2009). This was also shown by Angelova et al. (2012) research synthesis. The few qualitative studies that spanned two or more policy areas (see, e.g., Siedentopf and Ziller 1988; Duina 1999) did not follow a systematic comparative approach in the sense that they were looking for systematic sectoral patterns in implementation.

The same is also partly true for quantitative studies. First of all, there are also numerous studies that focus on only one sector, again overrepresenting social policy and environmental policy (Angelova et al. 2012). At the same time, the field has seen a growing number of studies looking at multiple sectors (Steunenberg and Rhinard 2010; Haverland et al. 2011) or with a truly cross-sectoral scope (Perkins and Neumayer 2007; Toshkov 2008; König and Luethert 2009; Börzel et al. 2010). However, rather few of these studies have sought to test whether there are systematic cross-sectoral differences in compliance patterns (but see Steunenberg and Rhinard 2010; Haverland et al. 2011).

Therefore, EU implementation research is in need of cross-sectoral studies that could dig deeper into the specific sectoral logics of EU legislation and its domestic implementation. Following the seminal article of Lowi (1972), domestic implementation researchers such as Mayntz (1977), Windhoff-Heritier (1980), or Ripley and Franklin (1982) have argued that different policy types imply different conflicts and problems, and therefore lead to different types of implementation processes. Moreover, EU compliance research will need to broaden its perspective in order to cover policies that have hitherto attracted only scant attention, such as agricultural policy or issues belonging to the expanding area of justice and home affairs.

3.2 Measuring compliance

One of the most problematic areas of previous research relates to measuring the dependent variable: compliance. If we focus on the implementation of EU directives, the most widely-studied phenomenon in the existing literature, compliance encompasses three individual aspects: (1) incorporating the policy provisions of EU directives into domestic law within the specified deadline and in a correct manner, i.e., so that domestic law conforms to the standards laid down in the respective EU directive (transposition); (2) guaranteeing that the norm addressees actually behave in a way that is in line with the legal norms laid down in the directives in question (application); and (3) providing for judicial and administrative mechanisms to ensure that non-compliant behaviour by addressees may be detected and non-compliers can be forced to change their behaviour with a view to respecting the norms (enforcement). How have these aspects of compliance been measured so far?

Measuring transposition performance of member states has been approached rather differently by qualitative and quantitative research. Qualitative scholars have offered direct measures of the timeliness and correctness of transposition, usually based on information gathered from expert interviews, legal documents, NGO reports and media coverage. This in-depth approach ensures that the measures are likely to be valid, but it of course implies that only a rather small number of directives and/or countries can be studied.

Quantitative studies are not bound to such limitations, at least as long as they continue to rely on readily available official data. Thus, most quantitative studies span several or even all policy sectors, cover longer periods of time, and include many if not all member states. This large empirical scope, however, is counterbalanced by a reliance on rather indirect measures of
transposition performance, which raises serious issues of measurement validity (for an overview, see Hartlapp and Falkner 2009).

Many studies have drawn on official information describing member states’ notification of their transposition measures. The literature has been based on three main approaches to using notification data.

The first approach relies on aggregate transposition rates published annually by the European Commission (Lampinen and Uusikylä 1998; Bergman 2000; Giuliani 2003; Toshkov 2007a). This data indicates the percentage of applicable directives a particular member state had notified the Commission of as transposed in a particular year. It does not allow any insights into the exact dates of notification, nor does it indicate which directives were notified in which year. Therefore, one cannot analyse whether directive-level properties have an impact on notification. Moreover, there is no information on when the transposition laws themselves were adopted or entered into force.

The second and most widely-used approach rests on information about domestic transposition measures drawn from the official Celex/EUR-Lex databases, which is often complemented with data from domestic legal databases. It sheds light on the individual laws a member state notified to the Commission as measures transposing a given directive. This may include laws that were already in place before the adoption of the respective directive as well as any number of laws that were enacted to fulfill the legal requirements, each with a date indicating its adoption, publication or entry into force. If there is more than one law, scholars have to decide which of the laws they choose to determine the date of transposition. The largest project that has used this type of data decided to take the first law enacted after the adoption of the directive as the key law representing the date of transposition (Berglund et al. 2006; Berglund 2009; Steunenberg and Rhinard 2010; Haverland et al. 2011). This approach assumes the first law is the most important one, ignoring all laws that are adopted after this initial reaction. Moreover, the number, sequence, and relative importance of transposition laws are likely to be influenced by national and/or sectoral regulatory styles, which means that choosing the first adopted instrument may bias the measurement of transposition. In sum, what this type of operationalisation measures is not transposition performance but the speed with which member states take a first legislative step towards incorporating EU directives into their legal systems. To overcome this limitation, others have decided to take the last measure reported in Celex/EUR-Lex as the date at which the transposition process was completed (Toshkov 2007c, 2008; Spendzharova and Versluis 2013).

The third approach was proposed by König and Luetgert (2009), who also decided to use all transposition measures contained in the databases, but classify member state reactions in an ordinal scale with the main categories ranging from situations where all notified measures were adopted (1) before the adoption of the directive, (2) between the adoption of the directive and the transposition deadline, and (3) after the deadline. Moreover, there are intermediate categories covering mixed situations and a final category where no measures were notified to the Commission. This operationalisation seems to go a long way towards covering not only all measures but also extracting information on different member state reactions in terms of transposition timing.

No matter what specification of the dependent variable is chosen, however, all of these operationalisations on the basis of notification data fail to grasp the completeness and substantive correctness of transposition. In other words, what these studies analyse is the temporal reaction of member states to EU directives rather than compliance.

This is different for the second major data source on compliance: official data on infringement proceedings initiated against member states that, in the eyes of the Commission, failed to comply with EU law (Mbaye 2001; Börzel 2001, 2003b; Taliberg 2002; Sverdrup 2004; Beach 2005; Perkins and Neumayer 2007; Börzel et al. 2010, 2012). The advantage of this data is that it covers at least some of the actual cases of incorrect or insufficient transposition. However, in-depth empirical case studies have demonstrated that this data is also far from perfect. Due to a lack of resources,
Implementing and complying with EU governance outputs

the Commission is only able to systematically detect and pursue cases of late notification, while many cases of inaccurate transposition slip past its attention. Thus, Falkner et al. (2005: 204–205) conclude that the Commission’s infringement data only represents the “tip of the iceberg”, which does “not necessarily say much about the size or the shape of those parts that remain below the waterline” (see also Hartlapp and Falkner 2009; König and Mäder 2013, 2014).

Moreover, quantitative research has found clear indications that the Commission’s enforcement policy is not neutral. Thomson, Torenvlied, and Arregui (2007) showed that directives that ran counter to the Commission’s policy preferences at the decision-making state are less likely to receive infringement proceedings than directives which were close to the Commission’s preferences. König and Mäder (2014) demonstrated that the salience the Commission attaches to a directive and the probability of enforcement success influences the Commission’s decisions to take action against non-compliant member states. These indications of strategic behaviour on the part of the Commission should caution against treating infringement data as unbiased indicators of compliance. Instead, they represent violations of EU law the Commission was able to detect and was willing to enforce.

In sum, the different types of quantitative data used to measure transposition performance are all fraught with major problems of validity, which suggests scholars using this data should be very careful about what it is they are actually analysing – member states’ initial reactions to EU directives, the Commission’s activities against allegedly non-compliant member states, etc. – instead of safely assuming they are examining indicators of compliance. Moreover, scholars using infringement data should try to incorporate measures of the Commission’s preferences into their analyses in order to disentangle strategic Commission activities from member state non-compliance.

Since it is highly undesirable to merely look at things that are somehow related to compliance but overlook major aspects of the concept, scholars are well-advised to invest more energy in projects that use qualitative techniques to measure compliance and at the same time try to go beyond the small-n settings to which qualitative research is often confined. This could be done by forming larger collaborative projects such as the ones organised by Siedentopf and Ziller (1988) or Falkner et al. (2005), which would be able to analyse more than just a few cases and would thus have a chance of improving the external validity of their findings. Quantitative scholars could also invest more energy in finding better data sources on compliance, especially on the completeness and substantive correctness of transposition. For example, Zhelyazkova and Torenvlied (2009, 2011) extracted information on compliance with individual provisions of directives from transposition reports prepared by the Commission and NGOs. Hille and Knill (2006) analysed the Commission’s progress reports in order to assess candidate countries’ alignment with the acquis, and König and Mäder (2013, 2014) asked law school graduate students to assess the legal correctness of domestic transposition measures. Although data sources such as these might not always be available, involve cumbersome coding procedures and, in the case of Commission monitoring reports, may be ridden with similar problems of Commission bias as infringement data, they could nevertheless be used more systematically in order to shed more light on the substantive rather than the temporal aspect of legal compliance.

Compared to transposition performance, measuring enforcement and application has been almost exclusively restricted to qualitative studies, which usually rely on interviews with enforcement actors, independent experts or representatives of affected societal groups as well as on analysing media reports and official documents to learn about the extent to which EU rules are actually complied with by norm addressees and the ways in which member states ensure that violations are sanctioned and redressed. Siedentopf and Ziller (1988), Knill and Lenschow (1998), Duina (1999), Versluis (2003, 2004, 2007), Falkner et al. (2005, 2008). This may lead to findings with a relatively high degree of internal validity, although qualitative techniques have more problems establishing the extent to which certain rules are actually complied with in practice than with finding out whether a given legal provision is fulfilled by the laws of a country. At the same time, collecting the wealth of data necessary to assess application and enforcement on a qualitative basis is of course

http://www.livingreviews.org/lreg-2014-1
a cumbersome task which reduces the number of cases that can be studied and thus, again, has a negative effect on the external validity of findings generated through such qualitative procedures.

While this problem could partly be solved by forming larger collaborative projects, it may also be a promising strategy to search for alternative data sources that could allow for a broader, more quantitative assessment of application and enforcement. One strategy, which has already been employed by individual studies, is to extract information from reports that assess certain aspects of member states’ enforcement systems, for instance a report prepared by a European-level committee of labour inspectors reporting on domestic enforcement systems in the field of monitoring compliance with labour law (Falkner et al. 2005; Jensen 2007; see also Hartlapp 2014 for a comparison with more recent data). There could well be other monitoring reports prepared by the EU Commission, NGOs or other international agencies which could shed light on the effectiveness of enforcement or the quality of practical compliance with EU legislation (for a promising first step in this direction, see Toshkov 2012). Moreover, scholars should invest more time into producing new or uncovering existing quantitative data on norm application and rule compliance. More quantifiable evidence on rule application could be gathered by surveys among ‘users’ of the norms or among people affected by certain rules (Gelderman et al. 2010).

Finally, scholars could also use existing secondary data sources to measure outcome variables and try to learn something about the actual effectiveness of EU law. This could be accident statistics, survey data, economic indicators, or data on energy use or greenhouse gas emissions. While it is true that such outcomes may be affected by many other variables besides legislative measures, this type of research could nevertheless shed more light on the degree to which EU policies are actually able to solve the problems they are intended to solve.

3.3 Theoretical insights I: Transposition

After these remarks on the problems of measuring the dependent variable, this section addresses the insights gained so far about the political, legal, administrative and cultural parameters influencing the implementation of EU law. This overview will start with the phenomenon that has received most attention so far, the transposition of EU directives, and will then move on to application and enforcement. The discussion of factors influencing transposition performance starts with EU-level variables, moves on to two groups of domestic factors, one influencing member states’ willingness to comply, as highlighted by the enforcement approach, the other shaping their capacity to comply, as stressed by the management approach, and finally addresses sectoral and country-related differences.

3.3.1 EU-level factors

To what extent does EU decision-making shape domestic transposition? Is the incorporation of EU legislation into domestic law a direct continuation of the battles fought in Brussels? Research addressing this question has yielded mixed results so far. Some studies found that a member states’ opposition to a proposal during negotiations in Brussels has a negative impact on its transposition performance (Thomson, Torenvlied, and Arregui 2007; Thomson 2010; König and Mader 2013; 2014; Zhelyazkova 2013), others found that effect only in a few exceptional cases (Falkner et al. 2004; Falkner et al. 2005: 277–280), and still others found no significant effect (Linos 2007; Zhelyazkova and Torenvlied 2009).

Similarly, the level of conflict during the EU decision-making process sometimes turned out to be a significant factor, with more conflict increasing the likelihood of transposition problems (Kaeding 2008a; König and Luettgert 2009; Luettgert and Dannwolf 2009; König and Mader 2013), while other studies, using data of a more limited scope, found no significant relationship between conflict in the Council and transposition performance (Zhelyazkova and Torenvlied 2011; Zhelyazkova and Torenvlied 2013; 2014; 2015; 2016; 2017).
Implementing and complying with EU governance outputs 21

or even the opposite effect – that more contested directives tended to be transposed faster (Zhelyazkova and Torenvlied 2009).

Findings on the impact of the decision-making rule in the Council on transposition performance have so far been rather negative. Theoretically, most actors expect directives adopted on the basis of qualified majority voting to be fraught with more transposition problems as it is more likely that such decisions disregard vital domestic interests, which could then strain transposition. So far, however, only one study could establish this effect (König and Luetgert 2009), while several others found no significant impact of the decision-making rule (Mbaye 2001; Kaeding 2006; Haverland and Romeijn 2007; Luetgert and Dannwolf 2009) even found qualified majority voting facilitated rather than hampered transposition.

Another EU-level factor whose effect has been scrutinised so far is the power of member state governments at the European level. Here, the results are again rather inconclusive, with some studies showing that more powerful governments tend to have less compliance problems (Giuliani 2003; Jensen 2007; Perkins and Neumayer 2007), while others revealed the opposite effect (Mbaye 2001; Börzel et al. 2010; 2012). Both findings are plausible theoretically. The problem is that they all result from analyses of infringement data and could thus be influenced by uneven enforcement on the part of the Commission. The best thing to shed more light into the impact of relative member state power would be to test for the effect of this factor in a study that either controls for the Commission’s preferences or uses transposition data. The latter has been done by Spendzharova and Versluis (2013), but their analysis of notification data from the field of environmental policy did not reveal a significant effect of the power of member state governments on transposition outcomes.

Several authors have started addressing the strategic nature of Commission enforcement and its potential impact on member states’ transposition behaviour explicitly. However, the results as yet have been mixed as well. It seems the Commission’s policy preferences do not have a direct impact on the timing of member states’ transposition (Thomson, Torenvlied, and Arregui 2007; Zhelyazkova and Torenvlied 2009; Thomson 2010). However, show that the Commission’s policy preferences influence the number of infringement proceedings initiated: directives which are far away from the Commission’s policy preferences at the policy formation stage are less likely to receive infringement proceedings than directives which were close to the Commission’s preferences (see also Steunenberg 2010 for a theoretical argument along the same lines). This ties in with case-study evidence showing that the Commission strategically uses infringement proceedings against member states that block liberalisation directives with a view to forcing them to open up their markets and, as a consequence, give up their opposition in the Council (Schmidt 2000). König and Mader (2014), however, using similar data as Thomson, Torenvlied, and Arregui (2007) but different statistical models, could not establish a significant effect of Commission preferences on infringement action. What they found instead was that the Commission prioritises cases to which it attaches high salience and in which successful enforcement appears likely.

In contrast to these rather inconclusive findings, many studies have demonstrated that the level of discretion granted to member states has a significant effect on transposition, with directives granting more discretion being easier to transpose since they give member states more opportunities to adopt EU policies to domestic conditions (Thomson 2007; 2009; Thomson, Torenvlied, and Arregui 2007; Steunenberg and Toshkov 2009; Zhelyazkova and Torenvlied 2009; König and Mader 2013; 2014) for individual provisions granting discretion, see Zhelyazkova and Torenvlied (2011) and Zhelyazkova (2013). The opposite effect was found in only one study focusing on the transport sector (Kaeding 2007; 2008a,b). Other EU-related factors that were high on the agenda in the first wave of research, such as the involvement of parliaments in the preparation of government negotiation positions so as to ensure that they would not obstruct the transposition of policies adopted Brussels, have not found support in more recent research (Falkner et al. 2005: 283–284; Sprungk 2011). This can hardly come as a
surprise to scholars studying the logic of parliament-government relations in parliamentary systems, where parliamentary majorities and governments are usually tied together through parties or party coalitions, which makes it unlikely that parliaments would block proposals tabled by governments, no matter whether they were involved in the adoption of the underlying policy decision or not.

In sum, research on the impact of EU-level factors on domestic transposition has yielded few robust findings so far. Most supportive evidence has been presented on a directive-level factor, the facilitating effect of higher levels of discretion, while many of the results on the other parameters are still rather inconclusive and unstable if other indicators or other samples of countries, directives and periods of time are analysed. In general, most research on the impact of EU-level factors on transposition operates on the basis of quantitative data that measure only certain aspects of legal compliance, which could be an additional reason for the contradictory findings of different studies. As things stand now, knowledge about the EU decision-making process does not help us much in understanding transposition processes. Instead, domestic factors still seem to carry the main weight in explaining member states’ varying transposition performance.

3.3.2 Domestic factors related to member states’ willingness to comply

Among the domestic factors influencing transposition, the literature has dealt with four main indicators influencing the willingness to comply: party politics, misfit, public opinion, and interest groups.

The impact of party politics was predominantly demonstrated in qualitative studies. Research on the implementation of EU social policy and equality directives revealed a relatively strong impact of government parties’ left-right positions on transposition outcomes in several countries [Treib 2003a,b, 2004, 2008; Falkner et al. 2005, 2008; Sedelmeier 2009, 2012]. Other studies also uncovered individual cases where the ideologies of government parties or changes of government had a crucial impact on transposition [Bähr 2006; Mastenbroek and Kaeding 2006; Mastenbroek and van Keulen 2006].

This partisan effect could also be discerned in some quantitative studies. Toshkov (2008) demonstrated that the new member states from Central and Eastern Europe were more likely to transpose the acquis on time if they were governed during the accession period by more pro-European and more right-wing parties. The effect of government positions towards European integration is corroborated by a study on infringement proceedings in the field of social policy (Jensen 2007) and by an analysis of transposition notification data of social policy directives in Central and Eastern Europe (Toshkov 2007c). Jensen and Spoon (2011) showed the partisanship of governments had a crucial influence on member states’ compliance with the greenhouse gas reduction targets agreed among EU member states to implement the Kyoto Protocol. Additional evidence on the relevance of party politics was presented by studies on member states’ reactions to ECJ case law in the fields of healthcare, trade union rights, and citizenship [Obermaier 2009; Sack 2012; Blauberger 2012; Schmidt 2014].

The problem with the partisan argument is that it is hard to operationalise properly in cross-sectoral and longitudinal studies. Since left-wing governments are expected to be more favourable towards policies in some sectors while more reluctant towards policies in other sectors, simply controlling for the left-right position of government parties will often not yield meaningful results (but see Toshkov 2008). Therefore, the effect of government left-right positions is mainly tested in sectoral studies, for example on social policy, where it has, however, often turned out to be insignificant [Jensen 2007; Linos 2007; Toshkov 2007a,b]. In a quantitative study on the transposition of environmental policy directives, in contrast, Spendzharova and Vershuis (2013) could demonstrate that governments attaching high salience to environmental protection, and governments comprising Green parties, tend to transpose environmental directives faster.

König and Luetgert (2009) have decided to go for a different operationalisation, which does not
measure the partisanship of government but the diversity of (sector-specific) partisan preferences in the domestic arena. They calculate the maximum ideological distance of all parties represented in the respective national parliament and find that this variable has a significant impact on transposition performance: the higher the heterogeneity of preferences, the more problematic the transposition. It is debatable whether this effect makes sense theoretically since transposition measures are unlikely to require supermajorities in parliament (Steunenberg and Toshkov 2009: 956). Apparently, König and Luetgert (2009: 189) also tested the effect of preference heterogeneity among government parties only, but this covariate did not turn out to be statistically significant.

These findings suggest party political effects may be present, but they are confined to certain policies or policy areas which are party politically salient and they are often the exception rather than the rule since a large part of the directives coming from Brussels are of a technical nature and are dealt with by bureaucracies in routine fashion (Berglund et al. 2006; Berghlund 2009).

The second variable that has an impact on member states’ willingness to transpose EU directives is the **degree of fit between the policy goals enshrined in European legislation and pre-existing domestic policy legacies**. This factor has found support in quite a number of qualitative and quantitative studies. The research synthesis of 37 studies by Angelova et al. (2012) found a robust effect of the ‘goodness of fit’ variable. Although two other summary analyses yielded a slightly different picture (Toshkov 2010; Toshkov et al. 2010), it is true that quite a number of studies have confirmed the ‘goodness of fit’ as a relevant predictor of transposition performance. It seems, however, that a large part of the confirmatory evidence comes from qualitative studies, which have used some rather specific indicators to measure the ‘goodness of fit’ and which have transformed a deterministic into a probabilistic argument.

The original idea behind the concept was to grasp fundamental clashes between the goals enshrined in European pieces of legislation and domestic policy legacies and institutional structures, and the original argument was formulated in a rather deterministic way: fitting policies were expected to be implemented without any problems whereas misfitting policies would lead to protracted implementation processes involving long delays or substantive flaws. Indeed, such patterns could be found in individual cases, although it already transpired in the original studies that there are cases where other, more actor-related variables had to be added to the parsimonious misfit model (Duina 1997, 1999; Duina and Blithe 1999; Knill and Lenschow 1998, 2000a, 2001; Börzel 2000, 2003a). Some other studies also found evidence that conformed to this original specification of the misfit argument (Bailey 2002; Dimitrova and Rhinard 2005; van der Vleuten 2005; Di Lucia and Kronsell 2010; Siegel 2011). However, several studies from the qualitative camp presented evidence that placed doubt on the misfit argument (Falkner et al. 2002, 2005; Treib 2003a, b, 2004; Bahr 2006; Mastenbroek and Kaeding 2006; Mastenbroek and van Keulen 2006; Liefferink et al. 2011; Bourblanc et al. 2013).

In contrast to the contradictory findings of qualitative research, quantitative studies largely seem to support the ‘goodness of fit’ argument. This type of research, however, has given the argument a different twist. If indicators seeking to measure misfit are included in multivariate statistical models, the logic turns from a deterministic bivariate argument (‘serious mismatches cause transposition problems, matching policies are transposed without problems’) into a probabilistic multivariate hypothesis (‘all other things being equal, higher degrees of misfit increase the likelihood of delayed or flawed transposition’). In this transformed version, it was argued that misfit was a significant explanatory variable even in Falkner et al.’s own data (Thomson 2007, 2009), suggesting that if one controls for various other factors, higher degrees of misfit tend to increase the likelihood of long transposition delays. This finding, however, is quite different from the original formulations of the argument.

A second shift that went along with the increasing use of the ‘goodness of fit’ in quantitative research is related to operationalisation. Due to a lack of case-specific data on legal reform require-
ments, their practical significance, institutional or procedural mismatches, and cost implications, scholars started to operate with what they had. Thus, the distinction between Commission directives and Council directives or directives of the Council and the European Parliament was used as a proxy for the significance of different pieces of legislation. Indeed, several studies showed Commission directives caused less trouble in transposition than Council or Council/EP directives, presumably because Commission directives are of a more technical nature and thus imply less wide-ranging reform requirements (Mastenbroek 2003; Borghetto et al. 2006; Borghetto and Franchino 2010; Kaeding 2006; 2008a; Steunenberg and Rhinard 2010; Haverland et al. 2011; Spendzharova and Versluis 2013).

A second indicator frequently used in quantitative studies to measure the ‘goodness of fit’ relates to the distinction between new-amending directives, the idea being that amending directives are less far-reaching than new directives. Here, the picture is a bit more mixed, with many studies finding significant effects (Mastenbroek 2003; Borghetto et al. 2006; Borghetto and Franchino 2010; Kaeding 2006; 2008a; Luetgert and Dammwolf 2009; Steunenberg and Rhinard 2010; König and Mader 2013; Spendzharova and Versluis 2013), while a few others did not (Haverland and Romeijn 2007; Haverland et al. 2011). This might be due to the differing scope of individual studies and to different statistical specifications.

Scholars using transposition data retrieved from Celex/EUR-Lex or domestic databases have also employed a third indicator to measure the ‘goodness of fit’. They distinguish between cases where the notified laws included a measure that had come into force before the adoption of the relevant directive from cases where this was not the case. The idea behind this indicator is that cases where laws existed before a directive came into effect would be associated with less reform requirements than cases where no laws existed before. This effect could be demonstrated in some studies (Mastenbroek 2003; Linos 2007; Thomson, Torenvlied, and Arregui 2007).

In general, however, we have to conclude that indicators such as these are a far cry from the original misfit concept. They only measure rather indirect aspects of legal reform requirements, as it is far from obvious that Commission directives are always less far-reaching than Council directives, that amending directives carry less significant policy substance than new directives, or that cases where pre-existing domestic legislation was notified are associated with lower degrees of misfit than cases where only new laws were notified. Moreover, the indicators do not say anything about the non-legal aspects of misfit – the practical significance of legal reform requirements, institutional mismatches, clashes with established state-society relations, or cost implications. Finally, the probabilistic version of the misfit hypothesis used in quantitative studies is quite different from the original formulations of the argument. These limitations should be kept in mind when interpreting the growing quantitative evidence confirming the ‘goodness of fit’ argument.

Some quantitative research has been devoted to the impact of public opinion on member states’ transposition performance or their behaviour in relation to infringement proceedings. Most often, this was done with regard to public opinion towards European integration. The findings, however, are rather disconfirming. Only one study so far has found a robust effect showing that public support for European integration facilitates compliance (Mbaye 2001), while several others could not uncover significant effects (Lampinen and Usikyla 1998; Kaeding 2006; Börzel et al. 2010, 2012). As a consequence, the process of incorporating EU directives into domestic law does not seem to be a matter of general support of, or opposition to, European integration. However what about more policy-specific aspects of public opinion? Surprisingly little research has been devoted to this issue so far. One of the few exceptions is a study by Spendzharova and Versluis (2013) on transposition in the field of environmental policy, which showed that countries where many citizens attached high priority to environmental protection tended to transpose environmental directives faster.

The impact of interest groups, finally, was identified in a number of studies as a key factor in transposition. However, there is wide disagreement on whether interest group activities are bene-
Implementing and complying with EU governance outputs

Several qualitative studies have shown that groups whose members profit from a particular piece of EU legislation can help overcome the resistance of unwilling governments or administrations by means of lobbying, public shaming, litigation, or lodging complaints with the European Commission (Börzel 2000, 2003a, 2006; van der Vleuten 2005; Pankó 2007; Slepečević 2009). Conversely, other qualitative studies have argued that interest groups whose members are affected negatively by a given EU directive may also hamper domestic adaptation to EU policies (Risse et al. 2001; Héritier 2001; Héritier and Knill 2001; Treib 2004; Falkner et al. 2005; 303–309).

Quantitative studies have struggled with the problem of finding appropriate indicators to measure interest group influence on the transposition process. Since no good indicators of the relative strength of groups that may be positively or negatively affected by an EU policy are available, and since the population of relevant interest groups as well as the direction of their influence is likely to at least depend on the policy area at hand if not on the individual piece of legislation, quantitative scholars have resorted to structural indicators of state-society relations, notably corporatism or pluralism indices. Given that there is not even a clear, theoretically well-founded argument about the expected effects of corporatism or pluralism on transposition, and considering the fact that state-society relations are highly sectoralised, it is no surprise that the results have been inconclusive. Two quantitative studies so far have found a significant effect of the degree of corporatism, showing that corporatist countries are more likely to transpose directives on time (König and Luetgert 2009) and are less likely to fail to comply with individual provisions of directives (Thomson 2010). Other studies have found no significant effect, some of them at least confirming the positive impact of corporatism (Thomson 2007, 2009), while others even found negative effects (Lampinen and Uusikylä 1998; Mbayé 2001). A case-sensitive analysis of social partner involvement in the transposition of EU social policy directives did not find any significant effects either (Falkner et al. 2005; 303–309; Leiber 2005).

In sum, it seems that among the factors influencing the willingness to comply with EU directives, the most widespread and systematic effects relate to the reform requirements associated with individual directives. Party politics only seem to come into play in some cases, whereas public opinion towards European integration appears to be largely irrelevant for transposition and the impact of interest groups is too heterogeneous to show up in a consistent manner in large-N studies.

3.3.3 Domestic factors related to member states’ capacity to comply

The literature has identified two main factors influencing the capacity of member states to transpose EU directives in a timely and correct manner: the number of veto players and administrative capabilities.

The number of veto players was first identified in second-wave scholarship as a relevant factor for implementation and, more broadly speaking, Europeanisation (Haverland 2000; Risse et al. 2001; Héritier 2001; Héritier and Knill 2001). In line with the veto player argument developed by Tsebelis (2002), the idea behind this line of reasoning was that the more actors had to agree to domestic implementation measures, the more likely it became that the implementation process would get stuck or dragged on beyond the deadline set by EU directives.

Although other case study based accounts found no delaying effects of veto players (Falkner et al. 2005; 296–298; Falkner, Hartlapp, and Treib 2007a), a finding which was also corroborated by some quantitative studies (Mbayé 2001; Borghetto et al. 2006; Börzel et al. 2010; Sprungk 2013), it turned out that a considerable number of other quantitative studies found confirmatory evidence. Indeed, the research synthesis by Angelova et al. (2012) confirmed that what they call “institutional decision-making capacity”, a concept that includes the number of veto players and related concepts such as federalism, is a robust predictor of transposition performance across the 37 articles they analysed. Significant effects of the number of veto players were found by Lampinen...
and Uusikylä (1998); Giuliani (2003); Kaeding (2006, 2007, 2008b); Linos (2007); Perkins and Neumayer (2007); Di Lucia and Kronsell (2010); and Börzel et al. (2012). The same is true for federalism, which was shown to contribute to transposition problems by Mbayel (2001); Haverland and Romeijn (2007); Linos (2007); Thomson (2007, 2009, 2010); König and Luetgert (2009); and Borghetto and Franchino (2010).

Some of these findings are open to criticism since they establish statistical effects for factors that may not be causally relevant for the cases analysed. For example, federalism is not relevant for transposition in cases where central government is responsible for adopting transposing laws and federal chambers do not hold veto power, as was the case for most of the transposition processes analysed by Falkner et al. (2005). This places doubt on the causal relevance of the findings by Thomson (2007, 2009, 2010) about the impact of decentralisation on transposition performance (Falkner 2007). Similarly, most veto player indices cover aspects of political systems that may be relevant for cases where directives are transposed by formal legislation adopted by parliaments, but they do not seem relevant if directives can be transposed by ministerial orders (Hartlapp 2009; Steunenberg and Rhinard 2010).

To account for the procedural effects of different transposition instruments, Steunenberg and Rhinard (2010) proposed a “procedural veto-player index”, which is sensitive to the different types and numbers of veto players involved in the making of ministerial orders, government decrees or acts of parliament. This variable turned out to be a powerful predictor of transposition timing (Steunenberg and Rhinard 2010). Other studies simply tested for the effect of different transposition instruments, but came to similar conclusions (Mastenbroek 2003; Kaeding 2006; Haverland et al. 2011). On the other hand, some authors have presented countervailing evidence, suggesting parliamentary involvement does not have a significant impact on transposition or might even facilitate transposition (König and Mader 2014).

The second widely-used variable influencing the capacity of member states to transpose directives on time and in a correct manner is administrative capabilities. Here, the picture is a bit more mixed, probably because some studies analysed sectors or countries where politicised transposition processes are more widespread whereas in other contexts, bureaucratic modes of transposition are more predominant. Altogether, however, many studies have confirmed that administrative capabilities are an important factor influencing transposition performance. Administrative capabilities involve various aspects. The first is administrative capacity or efficiency, which was shown to facilitate transposition by Mbayel (2001); Falkner et al. (2004); Falkner et al. (2005: 302–303); Berglund et al. (2006); Hille and Knill (2006); Kaeding (2006); Haverland and Romeijn (2007); Linos (2007); Perkins and Neumayer (2007); Toshkov (2007c); Berglund (2009); Knill and Tosun (2009); Börzel et al. (2010, 2012); König and Mader (2013); and Spendzharova and Versluis (2013), while no effect could be established by Thomson (2007, 2009, 2010) and König and Mader (2014). The second aspect is administrative experience with transposing EU law, a factor that involves learning effects of sectoral administrations that become more and more acquainted with EU law the longer their sector is affected by EU directives, which was confirmed as an important explanatory factor by Berglund et al. (2006); Kaeding (2006); Berglund (2009); Steunenberg and Rhinard (2010); and Haverland et al. (2011). The third aspect relates to effective administrative organisation and co-ordination. Several studies have shown that the presence or absence of central administrative bodies or procedures seeking to co-ordinate the tasks of incorporating EU legislation into domestic law has an impact on transposition performance (Ciavarini Azzi 1988; Rasmussen 1988; Dimitrakopoulos 2001, 2008; Bursens 2002; Zubek 2005, 2008; Dimitrova and Toshkov 2009).

A related directive-level factor is the degree of complexity a directive involves for domestic administrations. This factor goes back to first-wave studies, which suggested that complex directives are more problematic to transpose (Schwarze et al. 1993; van den Bossche 1996; Demmke 1998; Ciavarini Azzi 2000). The most frequent indicator for complexity employed in quantitative compliance research is the number of recitals. The empirical record for this factor is mixed,
Implementing and complying with EU governance outputs

with a number of studies showing complex directives involve more delayed transposition (Kaeding 2006, 2007, 2008a,b; Steunenberg and Rhinard 2010), while others did not find significant effects (Haverland and Romeijn 2007; Toshkov 2008).

In sum, structural factors determining the capacity of member states to process the legal incorporation of directives into domestic law play an important role in transposition. This primarily relates to administrative capabilities and to the (sector- or even case-specific) number of veto players involved in law-making processes. Since many of the empirical findings on these factors stem from quantitative studies using data of dubious validity, further research efforts should be devoted to confirming these results on the basis of other data.

3.3.4 Sectoral patterns

One of the advantages of quantitative research on the transposition of EU directives is that it allows scholars to search for broader patterns in the data that may not yet be fully explainable on the basis of existing theoretical arguments. One of these patterns relates to sectoral differences. Several studies report significant differences in compliance patterns between different policy areas (König and Luetgert 2009; Steunenberg and Rhinard 2010; Haverland et al. 2011). These sectoral differences have been partly explained by sector-specific configurations of different types of directives and of different types of legal instruments used to transpose these directives. In short, the argument made by Steunenberg and Rhinard (2010), and by Haverland, Steunenberg, and van Waarden (2011) is that different sectors are characterised by instruments that are easier or harder to be transposed (more or less Commission or Council directives and more or less amending or new directives). Moreover, different sectors rely on different typical transposition instruments involving different actor constellations. The combination of policy significance and procedural veto players can account for some of the sectoral variance to be observed. At the same time, Haverland, Steunenberg, and van Waarden (2011) also note there is further cross-sectoral variance which they cannot fully explain.

This suggests that further research into the reasons for sectoral differences in transposition performance is needed. This research should also take into account issues such as the varying degree of legal certainty that seems to apply to compliance with positive and negative integration, the different sectoral profiles of state-society relations, the different distribution of societal costs and benefits associated with distributive, redistributive and different types of regulatory policies, and the varying degrees of party politicization and the different political direction of party political effects in different policy areas.

3.3.5 Country patterns

Many studies have also reported widely divergent transposition patterns between different member states (Falkner, Treib, Hartlapp, and Leiber 2005; Falkner, Hartlapp, and Treib 2007c; König and Luetgert 2009; Börzel, Hofmann, Panke, and Sprunk 2010). Falkner, Treib, and their colleagues (2005, 2008) argued that these country differences not only reflect divergent favourable and obstructive factors for successful transposition in different countries but also that there are cultural differences between country clusters in terms of collectively shared norms of complying with the law. Several studies have tried to find empirical evidence that would be in line with Falkner, Treib, and et al.’s typology of different worlds of compliance, but not much confirmatory evidence could be found.

Re-examining the dependent variable of the ‘Complying with Europe’ data, neither Toshkov (2007a) nor Thomson (2007, 2009, 2010) could establish major differences in the transposition performance of the countries belonging to the worlds. The same is true for analyses using broader datasets. Examining the transposition of 24 directives in fifteen member states on the basis of Celex notification data, Thomson, Torenvlied, and Arregui (2007) again found that the worlds did
not perform in a significantly different way. Analysing annual non-notification rates between 1998 and 2005, [Toshkov 2007a] did identify significant differences between the country clusters, but he could not find the expected patterns in terms of the variability of transposition timeliness.

Also, efforts to identify the alleged differences between the worlds in terms of compliance cultures did not yield confirmatory results. Using opinion poll data measuring people’s attitudes towards law obedience, [Toshkov 2007a] showed that the differences between countries in general are rather marginal and that, to the extent that clusters are at all visible, they pinpoint the worlds of law observance and neglect on one side against the world of domestic politics on the other, which is not in line with the underlying theoretical argument. Thomson’s attempts to re-examine the dependent variable of the ‘Complying with Europe’ data also raised doubts about the argument that certain theoretical mechanisms work differently in the individual worlds. Instead, his analysis suggested that all explanatory variables exhibit the same direction of influence ([Thomson 2009]).

Although some of the disconfirming evidence was based on data of limited validity and used independent variables that Falkner, Treib et al. found to be causally irrelevant in their in-depth case studies ([Falkner 2007] [Falkner et al. 2007a]), the accumulated evidence of studies using different data sources suggests that efforts to test the worlds of compliance argument have not yielded much empirical support. This could mean that the country patterns observed by Falkner, Treib and their colleagues are specific to the field of social policy, that there are different country clusters, or that these clusters are not held together by the cultural factors that informed the worlds of compliance typology.

This conclusion leaves future research with the task of exploring the reasons for the significant country differences in transposition processes and outcomes. Since other studies have confirmed that the Nordic countries stand out as exceptionally good compliers ([Rasmussen 1988] [Sverdrup 2004]) while countries like Greece or Portugal perform significantly worse than most other countries ([König and Luetgert 2009] [Steunenberg and Rhinard 2010]), it will be up to further research to find out whether these differences are merely due to the distribution of favourable or unfavourable domestic institutional conditions in different countries or whether something beyond these differences accounts for cross-country variance.

Research on compliance in Central and Eastern Europe countries may shed further light on these questions. Several studies, both qualitative ([Leiber 2007] [Falkner and Treib 2008] [Falkner et al. 2008]) and quantitative ([Sedelmeier 2008] [Toshkov 2008] [Spendzharova and Versluis 2013]), have demonstrated that the new member states are exceptionally good compliers when it comes to transposition. While EU pressure related to accession conditionality could explain the good transposition record in the run-up to EU membership, the apparent continuation of this trend after accession must be caused by something else. One explanatory factor could be the creation of administrative co-ordination and supervision systems in many of the new member states’ core executives. Although the organisational properties of these systems seem to vary considerably, with stronger co-ordination systems being more favourable to transposition than weaker ones ([Zubek 2005] [2008] [Dimitrova and Toshkov 2009]), it is well possible that the institutional investments new member states made in order to come to grips with having to transpose the acquis communautaire continue to facilitate transposition also after accession ([Sedelmeier 2008]) – at least in routine cases that are not hampered by politicised debates ([Dimitrova and Toshkov 2009]).

Of course it is also possible that other factors account for the relatively good transposition record of Central and Eastern European countries. For example, it could be that the serious shortcomings of law enforcement detected in many CEE countries ([Sissenich 2005] [2007] [Falkner and Treib 2008] [Falkner et al. 2008] [Trauner 2009] [Toshkov 2012]) makes it easier for political opponents to agree to a certain legal reform as their actual impact can be softened by lax application. It will be up to future research to scrutinise whether the new member states from Central and Eastern Europe
continue to perform better in terms of transposition than the rest of the EU and if so, what accounts for this better transposition record.

3.4 Theoretical insights II: Application and enforcement

Given that comparatively few studies have been devoted to the later stages of the implementation process when transposed EU legislation or directly applicable EU law such as regulations are being applied in practice and violations are to be remedied through effective enforcement mechanisms, we have as yet comparatively little evidence on the extent to which there is non-compliance beyond transposition and on the factors that are conducive to effective application and enforcement. At the same time, there are some findings that seem relatively undisputed. They mostly relate to enforcement since the rule-compliant behaviour of EU rules by citizens, companies or administrative agencies is even harder to analyse empirically than more general features of court systems or supervisory bureaucracies.

Given the different actors involved in transposition as opposed to application and enforcement, it seems clear that different theories need to be applied to both phases of the implementation process. At the same time, scholars studying application and enforcement need not start from scratch since most studies agree that the application and enforcement of EU law is not fundamentally different from putting into practice policies that have a purely domestic origin (Ciavarini Azzi 1988; Falkner et al. 2005, 2008; Versluis 2003, 2004, 2007). Often administrative or judicial enforcement actors and societal target groups do not even know that a particular rule to be applied and enforced has European origins. This means analysing enforcement and application can build on theoretical and empirical insights from traditional domestic implementation research.

The first important parameter to account for is the type of enforcement system required for a given policy. Regulatory policy, which makes up the largest part of EU law, can be divided into two basic categories of norms. The first type of norms consists of provisions that demand or prohibit certain behaviour by private actors. Violations of such norms are usually sanctioned under criminal law by imposing fines or even prison sentences. This type of norms requires active enforcement by way of public inspections. Research on Congressional oversight of executive agencies has dubbed this monitoring mode ‘police-patrol oversight’ (McCubbins and Schwartz 1984).

The second type of norms grants individual rights to citizens, workers, consumers or companies. Such norms are usually sanctioned under civil law, for instance by granting disadvantaged persons the right to claim damages in court. In the American literature on Congressional oversight, this passive form of ensuring rule adherence was dubbed ‘fire-alarm oversight’ (McCubbins and Schwartz 1984). As much of the EU’s market integration agenda has been pursued through negative integration, which effectively declares certain discriminatory practices illegal but leaves it to private actors to assert their rights in court, and as some of the re-regulation coming from Brussels also relied on the stipulation of individual rights to be enforced through litigation, European integration has contributed to the spread of this second type of enforcement in Europe (Kelemen 2006, 2011; van Waarden 2009; van Waarden and Hildebrand 2009). Yet, there are still important areas of EU policy-making where active enforcement is required, including environmental protection, agriculture, health and safety at work, or food safety. However, enforcement through litigation also applies to areas where active enforcement is the dominant form. It allows citizens to sue norm violators for damages under civil law in addition to the state’s liability to enforce non-compliance under criminal law.

The effectiveness of these two basic types of enforcement depends on very different parameters. Active enforcement through monitoring and criminal sanctions is influenced by many of the factors highlighted by the top-down school of implementation research (Pressman and Wildavsky 1973; Bardach 1977; van Meter and van Horn 1975; Sabatier and Mazmanian 1981; Mazmanian and Sabatier 1983). It requires sufficient resources on the part of enforcement agencies, well-
trained inspectors, effective sanctions, and a functioning system of co-ordination between different enforcement agencies. Major shortcomings in these parameters have been shown to make active enforcement systems ineffective (Falkner et al. 2005, 2008; Jensen 2007; Hartlapp 2014).

Passive enforcement through individual litigation is influenced by many of the variables uncovered by the predominantly American literature on legal mobilisation (Zemans 1983; Epp 1998; for an overview, see McCann 2006; Vanhala 2013). Conducive factors include norm addressees that are well-informed about their rights, civil society organisations that actively support individual litigants, court systems or extrajudicial conciliation and arbitration bodies that provide easy access to individual litigants and civil society organisations, and supporting agencies or other actors that raise public awareness for the policy problem at hand (Alter and Vargas 2000; Caporaso and Jupille 2001; Falkner et al. 2005, 2008; Slepečević 2009).

In contrast to these findings, which apply to the general enforcement capabilities of individual member states in particular sectors, there are also factors that produce variance between individual cases. The first of these factors was highlighted by Versluis (2003, 2004, 2007) in the context of a study on the domestic enforcement of chemical safety rules. She argues that the salience of a particular piece of legislation is crucial for the extent to which it is actively enforced. Since enforcement actors cannot monitor everything, they need to prioritise their activities and are thus more likely to pay attention to rules that are highly salient, probably because there were major accidents that caught the attention of the general public or because the adoption of the law to be enforced was surrounded by fierce controversies. In this view, low-profile rules are less likely to be enforced than high-profile policies.

This finding is in slight contrast to the second factor that may have a bearing on the ease or complication with which different legal rules are put into practice: the degree of policy and institutional fit between a piece of EU legislation and existing domestic practices. The misfit argument, which was advanced by several second-wave scholars, not only applies to transposition but also to application and enforcement. It is in line with earlier top-down implementation research, which argued that laws requiring a vast change of behaviour on the part of target actors are harder to implement in practice than laws that only require gradual changes of behaviour (see, e.g., Sabatier and Mazmanian 1981). Although second-wave scholars often did not distinguish between implementation problems related to transposition, application or enforcement, a close reading of some of the case studies reveals that many of the problems with misfitting EU policies are related to policy application, especially in the context of ensuring that local environmental agencies comply with procedural and substantive provisions of EU environmental policy measures (Knill and Lenschow 1998, 2000a; Knill 2001; Börzel 2003a; Bailey 2002). It is here where the misfit argument is most plausible, since it is much easier to find a coalition of political actors to support the enactment of a major piece of new legislation than to make a large state machinery fundamentally change its bureaucratic routines or to ensure that economic actors comply with unfamiliar and costly new regulations.

A third factor that may account for divergent enforcement and application records in individual cases was highlighted by Dimitrova and Steunenberg (2013) [see also Dimitrova 2010]. Based on qualitative evidence from implementing EU rules on movable cultural heritage in Bulgaria, they present an argument that couples the different stages of the implementation process and highlights the impact of preference heterogeneity among domestic policy-makers on policy execution. Dimitrova and Steunenberg (2013) argue that the practical implementation of transposed EU legislation crucially depends on the preferences of both domestic policy-makers responsible for transposition and implementing actors on the ground. Domestic policy-makers holding highly divergent preferences on a given EU policy may be forced to transpose the policy under the threat of Commission enforcement, but their divergent preferences will not allow them to force implementing actors to stick to the transposed policy. If an implementing actor deviates from the transposed policy and sticks to the old policy regime, domestic policy-makers would have to take further action to
force this implementing actor into compliance, but given their heterogeneous preferences they are unlikely to succeed in doing so.

In sum, previous research has shed some light onto the logic of how EU law is implemented in practice. Many of the problems observed at these stages seem to be general problems of implementation in countries with administrations that lack resources or expertise, administrative structures that involve cumbersome co-ordination processes, court systems that are overburdened or hard to access by litigants, or a lack of civil society mobilisation to help individuals invoke their rights in court. On the other hand, there are also case-specific problems to do with policies that require major changes to the traditional way of doing things, that are too contentious to guarantee implementation on the ground, or that are perceived to be too unimportant to warrant major enforcement efforts.

It will be up to future research to produce more systematic, comparable data on the effectiveness of enforcement and practical application of EU legislation. Moreover, further work on the systematic differences in application and enforcement of different types of norms could contribute to a fuller understanding of compliance with EU law, and it will be interesting to see the results of ongoing research addressing the role of EU-level agencies in improving administrative policy delivery.

4 Conclusion

The process of implementing policies enacted at the EU level is a particularly interesting object of study. The EU is marked by a highly decentralised implementation structure that leaves responsibility for policy execution to the member states. Given the heterogeneity of interests among the actors involved in EU decision-making and the high consensus requirements, EU policies often contain fuzzy concepts and leave certain issues to the discretion of member states in order to facilitate agreement. What applies to implementation in general is thus particularly true for the domestic execution of EU policies: crucial decisions that may determine the success or failure of a policy are regularly taken at the implementation stage.

It was not until the mid-1980s that EU scholars discovered this interesting issue. Since then, the field has developed into one of the growth industries within EU research. In light of the considerable proliferation of EU compliance studies, this essay has provided a systematic overview of the four waves of research that have unfolded in the past decades, and it has identified the most important theoretical, empirical and methodological lessons to be drawn so far.

Despite some notable exceptions, research heretofore has focused on the transposition of EU directives to a remarkable degree, while still comparatively little is known about issues of enforcement and application. The focus on directives has also meant we lack information on how the implementation of regulations compares to directives. Moreover, there has been a heavy bias on positive integration while research on member states’ reactions to negative integration, which often comes in the form of ECJ rulings and seems much more prone to legal uncertainty than measures of positive integration, has only just begun. Also, existing studies have overrepresented individual countries and policy sectors and in general have only begun to systematically explore cross-sectoral differences. There is thus a huge empirical field to be explored by future research.

With regard to transposition, scholars meanwhile seem to agree we need to address factors that influence both the capacity of member states to comply and the willingness of domestic actors to fulfil the requirements stemming from EU legislation. The most important capacity variables appear to be administrative capabilities and veto players. With regard to willingness, the most significant factor appears to be policy misfit, especially the amount of legal changes required by directives, whereas party political or ideological variables only seem to be relevant in certain sectors and in certain high-profile cases. The main task to be accomplished by future research is
to establish under what conditions which configurations of factors prevail, especially with regard to systematic sectoral differences.

Concerning application and enforcement, it seems clear that research on the logics underlying the phase of practical implementation requires different theoretical approaches than those applied to transposition. At the same time, putting laws emanating from Europe into practice is done by the same administrative agencies and actors that also execute domestic laws. Therefore, approaches from domestic implementation research are the natural theoretical starting point for research on this phase of the EU policy cycle. Although parts of this literature are rather inconclusive and tend to operate on the basis of overtly complex theoretical models, it would seem worthwhile to invest some more time and energy in order to trace relevant parameters that could inform research on the practical implementation of EU policies. Further theoretical insights could be gained from management and enforcement approaches. These approaches, which were developed in the field of International Relations, have been applied to transposition research but have not yet found much resonance in research on practical implementation. Empirically, future research will have to focus more on the specific requirements of executing different types of norms, and there needs to be much more comparative research on the parameters influencing enforcement and application.

In methodological terms, especially quantitative transposition research will have to improve the data used to measure the dependent variable. Available data shed light on transposition timing but not correctness, or represent the actions of the Commission against non-compliant member states rather than the actual transposition performance of member state governments. Therefore, scholars are well advised to explore better data sources and, most importantly, invest more energy in producing their own data measuring transposition timing and correctness. Research on application and enforcement, on the other hand, needs to go beyond case studies and should instead search for or produce data with which the practical phase of implementation can be analysed on a broader, more comparative scale.
References


Ciavarrini Azzi, Guiseppe (Ed.), 1985, *L’application du droit communautaire par les états membres*, European Institute of Public Administration, Maastricht. (Cited on page 5)


*Living Reviews in European Governance*


Curtin, Deirdre and Egeberg, Morten (Eds.), 2009, *Towards a New Executive Order in Europe?*, Routledge, London. (Cited on page 15)


Implementing and complying with EU governance outputs


Knill, Christoph and Lenschow, Andrea (Eds.), 2000b, *Implementing EU Environmental Policy. New Directions and Old Problems*, Manchester University Press, Manchester. (Cited on pages [8] and [17].)


Living Reviews in European Governance


Mayntz, Renate (Ed.), 1983, Implementation politischer Programme II. Ansätze zur Theoriebildung, Westdeutscher Verlag, Opladen. (Cited on page 7)


---


*Living Reviews in European Governance*

http://www.livingreviews.org/lreg-2014-1


