The Strategic Nature of Compliance: An Empirical Evaluation of Law Implementation in the Central Monitoring System of the European Union

Thomas König  
University of Mannheim

Lars Mäder  
University of Mannheim

This compliance study models correct and timely implementation of policies in a multilevel system as a strategic game between a central monitoring agency and multiple implementers and evaluates statistically the empirical implications of this model. We test whether compliance is determined by the anticipated enforcement decision of the monitoring agency and whether this agency is responsive to the probability of enforcement success and the potential sanctioning costs produced by noncomplying implementers. Compared to other monitoring systems, the centralized monitoring system of the European Union (EU) is praised for exemplary effectiveness, but our findings reveal that the monitoring agency refrains from enforcing compliance when the probability of success is low, and the sanctioning costs are high. This results in a compliance deficit, even though the selective enforcement activities of the monitoring agency are almost always successful before court.

The enforcement of compliance by a central agency, which monitors lawmaking activities for implementing agreed policy goals and eventually litigates noncompliance by starting an infringement proceeding, is an effective ex post control mechanism for combating violations in multilevel systems, thus “reducing non-compliance to a temporal phenomenon” (Tallberg 2002, 609). In addition to international agreements and treaties, which come into force through ex post ratification by the signatories (Calvert, McCubbins, and Weingast 1989; Chayes and Chayes 1993, 1995; Downs, Rocke, and Barsoom 1996; Putnam 1988; Simmons 1998, 2009), the goals of federal policies are implemented by lower-level local or regional lawmakers (Elazar 1987; Filippov, Ordeshook, and Shvetsova 2004; Liiphart 1999; Scharpf 1988). Multilevel systems are prone to collective action problems as the number of collaborating actors increases (Scharpf 1988, 1997), which raises disputes on the optimal design of their material competencies (e.g., Alesina and Spolaore 1997; Boadway and Shah 2009; Frey and Eichenberger 2004; Weingast 1995). Because “free riding is the dominant strategy for large groups in the absence of a leviathan or of countervailing norms that can induce actors to monitor and punish defection” (Hooghe and Marks 2003, 239), compliance is a central challenge for the functioning of multilevel systems and may also impact the decision on their optimal design for lawmaking.

One way to cope with this challenge is to establish a court which can punish defections and facilitate compliance with the agreed policy goals at the federal and international level (Carrubba 2005, 2009; Staton 2006; Vanberg 2004). This literature considers courts to act as a fire alarm by providing a venue where the actors can bring cases and signal possible violations.1 For the...

1 Even in restrictive standing systems as with the World Trade Organization (WTO), in which only governments can bring cases, a court ruling will allow a less informed public to observe noncompliance and in the event of unsatisfactory government behavior will impose sanctioning costs, which may increase compliance (Carrubba 2009; Dái 2005).
collaborating actors, however, it is often too costly to monitor compliance with the agreed goals and begin infringement proceedings when the violation incentives of the other actors are not common knowledge. And since courts can neither monitor lawmaking activities nor enforce compliance without a litigator, one solution is to pool these costs by creating a central monitoring agency to detect noncompliance. Ideally, the agency does not pursue its own policy goals, but it may reduce information deficits on lawmaking activities and initiate infringement proceedings, leading finally to litigation to enforce compliance.

However, monitoring and bringing cases to court also implies costs, and if pooling provides limited resources for the monitoring agency, the agency’s decision to start infringement proceedings may establish a strategic game in which both sides base their decisions on the probability of enforcement success and the sanctioning costs of noncompliance. The EU is a prime example of the implementation of policy goals by lawmaking in a multilevel system, which has established a central monitoring agency to safeguard compliance (Börzel 2001; Falkner et al. 2005; König and Luetgert 2009; Siedentopf and Ziller 1988; Thomson, Torenvlied, and Arregui 2007). The collaborating member states agree about their policy goals with the adoption of EU directives (often by supermajorities), which they have to implement by transposing them into their domestic law. These lawmaking activities are monitored by the Commission, which can start infringement proceedings to enforce compliance, if necessary through the European Court of Justice (Carrubba 2005; Carrubba, Gabel, and Hankla 2008; Stone Sweet and Brunell 2012).

This study examines and evaluates compliance of lawmaking activities in a centralized monitoring system from a strategic perspective. Similar to the “policy-patrol” supervision of the U.S. Congress over the executive branch and its agencies, the Commission monitors the implementation activities and can start infringement proceedings against noncomplying member states (Tallberg 2002). The effectiveness of the central monitoring system of the EU, which also uses “fire alarms” (Jensen 2007), is shown by the high levels of compliance documented in annual reports of the Commission (1983–2010) and has been praised by many scholars (Bergman 1997; Martin 2000; Rhodes 1986; Steunenberg 2010; Sverdrup 2004; Tallberg 2002; Wallace 1984). The monitoring activities of the Commission mainly concern compliance with the goals of directives, which offer the implementing member states some discretion to achieve the policy goals adopted at the EU level. Directives are proposed by the Commission and adopted by the member states, sometimes in collaboration with the European Parliament, through a supermajority vote. Once directives are adopted, member states must correctly implement their goals by transposition into domestic law in due time (König and Luetgert 2009; Thomson, Torenvlied, and Arregui 2007).

Even though the member states have agreed on these goals and the general rules for compliance, including the Commission’s monitoring and infringement power, they may still have incentives to pursue their own policy interests in the implementation process at the expense of others. This generates a strategic compliance game, in which member states may attempt to anticipate the ability and willingness of the Commission to litigate noncompliance. Until now, however, compliance research has not yet accounted for the implications of the strategic interaction between the Commission as the central monitoring agency and the implementing member states, because the extensive literature has focused either on the lawmaking activities of the member states (Bailey 2002; Berglund, Gange, and van Waarden 2006; Borghetto, Franchino, and Giannetti 2006; Börzel 2000; Börzel et al. 2006; Bursen 2002; Duina 1997; Duina and Blithe 1999; Falkner et al. 2005; Haverland 2000; Haverland and Romeijn 2007; Hille and Knill 2006; Kaeding 2008; Knill and Lenschow 1998; König and Luetgert 2009; Lampinen and Uusikyla 1998; Luetgert and Dannwolf 2009; Mastenbroek 2003; Mastenbroek and van Keulen 2006;
When the payoffs are fixed and known in Neumayer 2007; Sedelmeier 2008). Giuliani 2003; Jensen 2007; Mbaye 2001; Perkins and 2007; Toshkov 2008; Zubek 2005) or on the decisions Thomson 2007, 2009; Thomson, Torenvlied, and Arregui Steunenberg 2006; Steunenberg and Toshkov 2009; 248 the discretion granted to them by each directive. Implementation delays by the preferences of the member states and in terms of either the number of infringement proceedings or im- stage. The study, however, also separately explains noncompliance present a novel dataset of 299 cases of compliance or implementation is more generally relevant for compliance decision of the Commission, our results reveal that the member states’ willingness to risk an enforcement conflict with the Commission if they strongly disagree with the commonly adopted policy goals. In addition, our analysis reveals that the member states’ willingness to risk an enforcement conflict with the Commission also increases with the preference diversity among them, since the preference diversity decreases the chances of the Commission forming an overwhelming enforcement coalition. For the decision of the Commission, our results reveal that the agency does not pursue its own policy goals; however, the agency’s likelihood to litigate noncompliance also decreases with the preference diversity among the member states and the absence of a pluralist interest group pattern in the domestic arena. Under these conditions, we

In this study, we bring together these two approaches by presenting a unified theory and test of the compliance game, constructing a strategic model that includes both the decision of the member states to correctly implement a directive within their own legal system in a timely fashion and the central monitoring agency’s decision to file an infringement proceeding. More precisely, our strategic model postulates that compliance is determined by the probability of enforcement success and the level of sanctioning costs, which we approximate by the utility losses of both types of actor. The standard enforcement approach assumes that the likelihood of noncompliance increases with the distance between an actor’s own policy position and the commonly adopted policy goals. However, when the diversity of the member states’ positions reaches a level at which the Commission fails to find enough support from other member states to successfully enforce compliance, our strategic model predicts that the central monitoring agency will refrain from initiating infringement proceedings. Here we follow the logic of Carrubba, Gabel, and Hankla (2008), which we complement by adding private information and by regarding the decisions of the Commission and the member states as part of a single strategic decision-making process.

In addition to this strategic compliance game, we present a novel dataset of 299 cases of compliance or noncompliance with directives, taking into account the implementation deadline and (in)correct implementation of their policy goals into domestic law. We believe that our data-generation process of (in)correct implementation is more generally relevant for compliance research because it allows a substantial evaluation of reported lawmaking activities in a large number of countries. For the evaluation of the empirical implications of our strategic model, we follow Signorino (1999, 2003) and apply a strategic probit model, derived from the McKelvey and Palfrey (1995) quantal response equilibrium (QRE). To our knowledge, previous applications of QRE models mostly used estimates generated by experimental design. When the payoffs are fixed and known in the laboratory, QRE estimates the variability of perceptual errors of the actors. In contrast to QRE analyses of experimental data, we use average utilities from observational data for estimating the bias of outcome probabilities and associated parameter estimates under misspecification. Wand (2006) shows that the key element of these models is determined by the configuration of actors’ preferences.

The results of our analysis suggest that both the member states and the Commission (as the central monitoring agency) act strategically. We find for each type of actor that the probability of enforcement success and relative sanctioning costs explain the strategies and outcomes of this process, which distinguishes between compliance by correct and timely implementation, enforced compliance of late and/or incorrect implementation, and a compliance deficit where the central monitoring agency failed to enforce compliance. On closer inspection of the decisions of the member states, we confirm a central argument of the enforcement school of thought, i.e., member states are less likely to comply with the goals of directives and hence more willing to risk an enforcement conflict with the Commission if they strongly disagree with the commonly adopted policy goals. In addition, our analysis reveals that the member states’ willingness to risk an enforcement conflict with the Commission also increases with the preference diversity among them, since the preference diversity decreases the chances of the Commission forming an overwhelming enforcement coalition. For the decision of the Commission, our results reveal that the agency does not pursue its own policy goals; however, the agency’s likelihood to litigate noncompliance also decreases with the preference diversity among the member states and the absence of a pluralist interest group pattern in the domestic arena. Under these conditions, we

5Only recently, Thomson, Torenvlied, and Arregui (2007) have drawn attention to the impact of factors related to the prior legislative negotiation stage on the later enforcement and implementation stage. The study, however, also separately explains noncompliance in terms of either the number of infringement proceedings or implementation delays by the preferences of the member states and the discretion granted to them by each directive.

6Because QRE models are based on individuals (or their “agents”) making random perceptual errors each time they face a decision, they were originally applied to experiments where participants knew the fixed payoffs to the game but nonetheless made stochastic choices (McKelvey and Palfrey 1995). Quinn and Westveld (2004) have provided a method for relaxing parametric assumptions about the distribution of errors in QRE models. Signorino (2003) proposes a pair of models. The first is an alternative agent-like theory where individuals do not make perceptual errors, but rather information is revealed to individuals when they face a decision. The second is a regressor error model based on a game where actors know each other’s type, but only limited information about actor types is available to a researcher observing the game.
identify a compliance deficit when the Commission, despite its general role as the “guardian of the treaties,” is either unable or unwilling to pursue an infringement against member states in this centralized monitoring system. Our analysis makes three contributions to the study of compliance. First, we specify the role of the enforcing actor by modeling the decision of a central monitoring agency, which can facilitate compliance by reducing the information deficit about possible violations and can undertake infringement proceedings in court to enforce compliance. Since we find that the central monitoring agency selects these cases, our results suggest a “conditionality” of the compliance-enhancing role of courts. Second, we present and test a strategic compliance game that incorporates implementation and infringement decisions into a single unified model. While previous studies analyzed the reasons for implementation and infringement separately, our model examines the causes of both decisions simultaneously. By this, we test whether and how these decisions are interrelated. This model allows us to identify the outcome of a compliance deficit that qualifies the effectiveness of the central monitoring agency. Third, we develop a novel data-generation strategy for the evaluation of lawmaking activities which distinguish between correct and timely implementation for the study of compliance. We combine tools from document analysis with standardized legal expertise to identify unobserved noncompliance outcomes for reported lawmaking activities—cases for which the central monitoring agency is unable or unwilling to start infringement proceedings. This is particularly relevant for our understanding of compliance in centralized monitoring systems where the actors’ decisions are influenced by their anticipation of the monitoring agency’s decision.

The first section of this article analyzes the procedural provisions of the implementation and enforcement stages. We introduce our game-theoretical compliance model of the strategic interaction between a central monitoring agency and implementing actors. We formulate hypotheses about the factors we expect will influence the decisions of the member states and the central monitoring agency. In the second section, we explain how we generated the data for our dependent variable of (non-)compliance and how we tested our hypotheses empirically with a sample of 299 cases. The last section consists of an extensive discussion of our findings, and we conclude by discussing their relevance for similar monitoring systems and by providing suggestions for further research on compliance.

A Strategic Model of the Two-Stage Implementation Process

Theoretically, the likelihood for strategic interaction increases when collaborating actors pursue different interests, have incomplete information, and take decisions over more than one procedural stage. Under these circumstances, actors attempt to anticipate the reactions and decisions of other actors in subsequent stages in order to optimize their own decision in the present stage. Member states have to implement EU directives within their own legal system in a correct and timely manner, before the Commission scrutinizes their lawmaking activities and decides whether or not to enforce compliance by starting infringement proceedings. Because member states are only required to report their activities without indicating whether they have correctly completed the implementation process, the Commission has established an infringement procedure to reduce the information deficit. This consists of: first, a formal letter requesting clarification, and second, the formulation and delivery of a reasoned opinion which includes the legal justification for enforcement actions and also a deadline after which the central monitoring agency may refer the case to the European Court of Justice. The European Court of Justice may then decide on the case and can impose financial sanctions on a noncompliant member state.

Even though this procedure establishes several interdependent implementation and enforcement decisions, the literature has investigated the implementation of directives separately from their enforcement stage. Studies on enforcement conventionally refer to the number of infringement proceedings, which indicate when and to what extent the Commission has reached enforcement decisions against member states. Using the notified lawmaking activities of the member states, implementation scholars investigate when and how member states report their activities to transpose directives into domestic law.7 A good example is the recent analysis of Thomson, Torenvlied, and Arregui (2007), which points to the relationship between the prior negotiation of directives and the subsequent implementation stage by separately evaluating the factors which influence implementation delays and infringement proceedings. Because the decision to implement a directive can also be influenced by the anticipated decision of the Commission to enforce compliance, we integrate the decisions of these two stages into a single model. The model postulates a strategic interaction...

7For a more detailed discussion, see König and Luetgert (2009).
between the member states and the central monitoring agency, contingent on the potential costs of sanctioning and the probability for enforcement success in the event of a compliance conflict.

Following McLean and Whang (2010), we model the strategic interaction between the member states and the agency as a stylized two-actor game. Figure 1 displays the structure of this game, which starts with a member state’s decision whether to correctly implement a directive into domestic law or to defect by late or incorrect implementation (Comp, ~Comp). If a member state correctly implements the directive within the prescribed time period, there is no potential conflict with the Commission. While the central monitoring agency receives the highest payoff with a value of 1 in this case, the complying member state accepts the directive’s outcome, which has been adopted at the prior negotiation stage. Consequently, the member state does not gain additional benefits but avoids any sanctioning costs from the Commission, meaning that its utility losses are set to 0. If the member state decides not to implement the directive in a timely and correct manner, the threat of the central monitoring agency to enforce compliance has failed, and the Commission needs to decide whether or not to take action.

This first stage of the game is equivalent to the design of existing studies, which investigate the implementation activities of the member states by the directives’ transposition records (König and Luetgert 2009; Thomson, Torenvlied, and Arregui 2007). Our model further includes a second stage in which the Commission decides whether to begin infringement proceedings against a noncomplying member state by means of a reasoned opinion (RO, ~RO). If the Commission as the central monitoring agency takes action by filing a reasoned opinion, it decides to litigate a noncomplying member state in order to enforce the goals of a directive. This leads to a compliance conflict between the Commission and the noncomplying member state, in which the member state succeeds with the probability of p and the central monitoring agency with 1 − p. In this conflict, the Commission’s payoff for enforced compliance is $p \times 0 + (1 - p) \times 1 - C_{\text{Comp}} = (1 - p) - C_{\text{Comp}}$, where 1 − p is the central monitoring agency’s probability of success and $C_{\text{Comp}}$ the costs of sanctioning. When the Commission decides to proceed against a member state, the latter can succeed with probability of p, but it may have to bear the sanctioning costs from the Commission $C_{\text{MS}}$. Hence, the noncomplying member state’s payoff for enforced compliance is $p \times 1 + (1 - p) \times 0 - C_{\text{MS}} = p - C_{\text{MS}}$. However, if the Commission decides not to initiate an infringement proceeding, a compliance deficit exists, and the noncomplying member state remains without sanctions. In this case, the noncomplying member state receives additional benefits with the highest payoff having a value of 1 and the Commission’s utility losses a value of 0 (no sanctioning costs). Table 1 summarizes the actors’ actions and compliance outcomes.

This game, which controls for a selection process by the monitoring central agency, has four actions (Comp, ~Comp, RO, ~RO) and three compliance outcomes: (1) compliance by timely and correct implementation of a directive (Compliance), (2) enforced compliance through litigation by the Commission (Enforced Compliance), and (3) compliance deficit by not proceeding against a noncomplying member state (Compliance Deficit). The actions of each member state and of the Commission depend on the evaluation of their probability of enforcement success (p, 1 − p) and the sanctioning costs for each type of actor ($C_{\text{MS}}, C_{\text{Comp}}$). Hence, we expect only to observe activity by the central monitoring agency when $p < 1 - C_{\text{Comp}}$ and $p \geq C_{\text{MS}}$, i.e., when the potential costs of sanctions are sufficiently low relative to the expected success of the noncomplying member state in the event of a conflict with the Commission. But when the respective costs are high, we expect that either the member state will comply and implement the directive in light of the high probability for the agency’s enforcement success, or the Commission will avoid a conflict with the noncomplying member state.

### Table 1 Theoretical Effects of Potential Costs on Outcomes

<table>
<thead>
<tr>
<th>Commission’s Enforcement Condition</th>
<th>Member State’s Compliance Condition</th>
<th>Equilibrium Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p \geq 1 - C_{\text{Comp}}$</td>
<td>For all $p \in [0, 1]$</td>
<td>Compliance Deficit</td>
</tr>
<tr>
<td>$p &lt; 1 - C_{\text{Comp}}$</td>
<td>$p \geq C_{\text{MS}}$</td>
<td>Enforced Compliance</td>
</tr>
<tr>
<td>$p &lt; 1 - C_{\text{Comp}}$</td>
<td>$p &lt; C_{\text{MS}}$</td>
<td>Compliance</td>
</tr>
</tbody>
</table>

### Operationalization

In order to test the empirical implications of this model, we need to gather information on the member states’ implementation and the Commission’s enforcement activities, together with measures of our two main explanatory variables, the potential costs of sanctioning (c) and the probability of enforcement success (p). Furthermore, our dependent variables need to distinguish between correct and timely implementation of the goals of a directive: the incorrect or late but subsequently enforced compliance...
The Dependent Variables

Although member states are formally obliged to implement the commonly adopted policy goals of directives within their own legal system in a correct and timely manner, it remains an empirical challenge to identify the level of compliance in the centralized monitoring system of the EU. A major reason for this difficulty is that member states only have to notify the Commission of their lawmaking activities without indicating whether they have completed the implementation process in a correct and timely manner. Furthermore, the extensive paperwork that member states usually produce when reporting on multiple activities complicates the evaluation of compliance (Hartlapp and Falkner 2009). The notifications of these activities may span over a long time period, are written in different languages, and include various types of legal acts reflecting national lawmaking systems. Hence, the evaluation of compliance with the goals of the directives is like looking for a needle in a growing haystack to which member states constantly add hay by notifying further activities.

To evaluate compliance with the goals of the directives, we therefore propose to focus only on the implementation of issues that proved controversial during the negotiations of the policy goals during the EU decision-making process. This information is available for 37 contested issues in 21 directives included in the dataset of the research project “Decision Making in the European Union” (DEU) (Thomson et al. 2006). The original aim of the DEU project was to compare the explanatory power of different decision-making theories by a set of contested cases, whereby experts were asked to rate the positions of the member states, Commission, and adopted outcomes on issue-specific scales ranging from 0 to 100. Using the description of these issues and their adopted outcomes, we ran a computer-based content analysis, which tagged each sentence in each notified document that includes a reference to the implementation of the adopted outcome of the respective issues. All tagged sentences were then evaluated through cross-validated legal expertise by asking law school graduate students to fill in a standardized questionnaire.\(^8\) The students assessed whether and to what level the content of these tagged sentences correctly implemented the goal of each outcome of the 37 contested issues included in the 21 directives of our sample. Finally, we aggregated the issue-specific evaluation on the level of each directive. We classified the implementation process as correct if the goals of all issues of a directive were transposed into domestic law or as incorrect if implementation failed for at least one issue.

In addition to the evaluation of compliance, we also collected and added the respective enforcement activities by the Commission for our sample of 21 directives. This generates our final dataset, including 299 compliance cases with all implementation and enforcement activities.

---

\(^{8}\) Approximately 50 graduate law students from these countries were trained with respect to the evaluation of the country-specific implementation process. Within the first weeks, students studied the formal rules and goals of legal implementation, followed by a period in which they worked out the implementation characteristics and processes of their respective countries (depending on their language skills). Finally, they learned about the contents of the 21 EC directives and were asked to interpret the formulation of each standard in comparison to the outcome, using a standardized questionnaire. For some countries—for example, Austria and Germany—more than five students evaluated the record of the issues with an overall intercoder reliability of almost 95%. The evaluation took place in 2007, almost three years after the deadline for the last of the directives expired.
of the actors involved. The member states implemented 195 cases correctly and in a timely manner, while they failed to comply in the remaining 104 cases. Among these, the Commission, as the central monitoring agency, filed infringement proceedings against noncomplying member states in 85 cases. Compared to the overall positive evaluation of the effectiveness of this central monitoring system, this dataset provides a more complex picture of compliance with directives.

In brief, only two directives were implemented correctly and in a timely manner by all member states. We also found that all member states failed to comply at least three times by either incorrect or delayed implementation. With 12 defections Portugal has the poorest record, failing to comply in 12 cases, followed by Belgium, France, Ireland, and Luxembourg, all with compliance failures for nine directives. In contrast to this group, Denmark and Finland have a much better record, with only three non-compliances. On closer inspection, there is also considerable variation for the Commission’s enforcement decisions across the member states.

A typical example concerns the implementation of directive 2001/55/EC, which defines temporary protection for asylum seekers for a mandatory period of two years with a possible extension of one extra year for refugees from armed conflicts. The proposed directive caused considerable controversies during the decision-making process at the EU level. Some member states, such as Austria and the Netherlands, favored a relatively short period of only 18 months, others preferred much longer periods, and Ireland and Denmark even declared that they would not be bound by the directive’s goals nor would they be subject to its application. The spread of interests across the member states already indicated potential compliance difficulties. Austria, which strongly disagreed with the directive’s adopted goal on temporary protection for asylum seekers, notified timely implementation through the 2002 amendment to its 1997 immigration law (BGBl. 126/2002). However, this incorrectly implemented the directive’s goals because it delegated the definition of the period of protection to an Austrian parliamentary committee instead of implementing the mandatory period with a possible extension. In spite of this incorrect implementation, the Commission did not commence infringement proceedings against Austria.

The Independent Variables
For measuring the potential costs of sanctioning (c) and the probability of enforcement success (p), we also use the data from the DEU project on the member states’ and the Commission’s issue-specific preferences, which proved very robust in terms of cross-validity (König and Proksch 2006; Thomson and Torenvlied 2011). Regarding sanctioning costs (c), we argue that even though member states may support the adoption of a directive when they agree on a compromise solution across the goals of multiple issues, they can still pursue issue-wise individual interests in the implementation process. We accordingly assume that a member state is more willing to bear the costs of sanctions when its level of disagreement with the directive is high. In other words, a member state with a high level of disagreement will regard sanctioning costs as sufficiently low relative to the expected gains from a conflict with the central monitoring agency. For the Commission, this perspective suggests that the monitoring central agency is more likely to enforce a directive against a noncomplying member state when its own distance to the directive’s goals is small. The gains from its enforcement decision should increase and the costs of sanctions decrease with the Commission’s level of agreement.

In essence, we suppose a positive relationship between both the level of a member state’s disagreement and the central monitoring agency’s agreement on the probability of the enforcement outcome. We measure the level of the member state’s disagreement by the preference distance between its most preferred policy goals and those finally adopted in a directive. For the Commission’s agreement, we subtract its absolute distance from 100 in order to express the agency’s benefit. Because several directives contain more than one issue-specific policy goal, we calculate the directive-specific level of (dis)agreement as the sum of the issue-specific (dis)agreement levels (Thomson 2007; Thomson, Torenvlied, and Arregui 2007; Zhelyazkova and Torenvlied 2009). We coded the few missing values on the level of disagreement for a member state as zero and the Commission’s agreement as 100, expressing their indifference with the outcomes. We further...
consider the saliencies that both types of actors attach to a directive. We assume that the relative costs of sanctioning decrease with their saliency and therefore expect a negative relationship between member states’ and the Commission’s saliency and the probability of enforcement.11

Regarding the probability of success (p) in the enforcement stage, we follow Carrubba, Gabel, and Hankla (2008) and argue that the European Court of Justice is constrained by the defendant member state’s response to a ruling since member states have the possibility to circumvent undesirable rulings through noncompliance (Garrett, Kelemen, and Schulz 1998). The credibility of such a noncompliance threat, however, depends on pressures against the violating member state from other member states and/or from their own domestic audiences. Compared to voters, interest groups have more resources to generate information on the implementation process and to exert pressure on their own member state by bringing a case to court (Dai 2005). If the European Court of Justice values compliance with its ruling, the likelihood of a ruling against a violating member state and the enforcement success of the central monitoring agency should increase with the likelihood of third-party intervention (i.e., either by other member states and/or by domestic interest groups). When almost all member states pursue the same interests in a directive, we expect that the likelihood for an overwhelming enforcement coalition against a single violating member state will increase. This should also increase the likelihood of a court ruling against a noncomplying member state and thus promote the (expected) success of the central monitoring agency.12 Conversely, we expect that the size of the enforcement coalition and the likelihood of a favorable ruling decrease when the interests of the member states differ or are in opposition to one another.13

In addition to other member states, domestic actors such as interest groups can bring a case to court, which will be determined by the distribution of the groups’ interests too. To measure this distribution, we use Lijphart’s index of interest group pluralism, with high values indicating countries with diverse and pluralistic interest group patterns (Lijphart 1999). Corporatist interest group patterns with low values indicate high intragovernmental collaboration and coordination with interest groups in a member state. This suggests that a member state’s decision to comply with a directive is more likely to already reflect the interest groups’ preferences in a corporatist system. Even if an interest group is against this decision in a corporatist system, it is more likely to refrain from bringing a case to court in view of the group’s risk for exclusion from future intragovernmental collaboration and coordination. In a pluralistic system with diverse and competing interest groups, it is therefore more likely that an interest group protests against the compliance decision of a member state and even brings a case to court. Thus, we expect that a pluralistic interest group pattern positively influences the Commission’s enforcement utilities, while it should reduce the enforcement utilities of the member states.

In addition to these variables, which specify the two main terms of our strategic game, we include three control variables. Regarding the member states’ enforcement utility, we control for the number of controversial issues contained in each directive. This controls for possible compensation, which is more likely for cases with multiple issues. In situations with two or more issues, the utility loss of a member state on one issue may be compensated for by a gain on another issue. We expect that the compensation likelihood increases with the number of issues and therefore positively influences the compliance utility of a member state. We also use a dummy variable in the Commission’s enforcement utility when the directive contains an issue concerned with the deadline for implementation. Since noncompliance cases of late implementation are much easier to detect than cases of incorrect implementation, we assume that an explicit reference to the deadline increases the enforcement utility of the Commission (Börzel 2001). Furthermore, we also expect that the type of directive affects compliance. Directives that are adopted by the Council and the European Parliament have a higher likelihood for compromise than those that are decided by the Council alone (Ciavarini Azzi 2000; Mastenbroek 2003). Thus, this type of directive should increase a member state’s utility for enforcement, while with a broader legitimization of directives, the Commission’s enforcement utility increases.

For the outcome of compliance by correct and timely implementation in our strategic game, we fix the utilities of the member states at 0 and of the Commission at 1, in accordance with our theoretical discussion. Hence, we expect that only directive-specific and state characteristics matter for this outcome (cf. Falkner et al. 2005; Mastenbroek 2003). In our strategic game, however, directive-specific characteristics such as complexity can affect the two types of actors in different ways: the central monitoring agency may have more difficulties to scrutinize

---

11 Similar to the measure of (dis)agreement, we sum the issue-specific saliency measures in the DEU dataset for each actor and for each directive.

12 For this argument, see Torenvlied (2000).

13 Following Zhelyazkova and Torenvlied (2009), we measure this diversity by the sum of the standard deviations of all member states’ most preferred outcome for the issues of each directive.
complex directives, while a member state may have more leeway for interpretation. Furthermore, Franchino (2004) reports that directives typically delegate more power to member states in those areas where implementation uncertainty is high (see also Epstein and O’Halloran 1999). Thus, the level of the delegated authority should be associated with a higher risk of noncompliance (Thomson, Torenvlied, and Arregui 2007; Zhelyazkova and Torenvlied 2009). For scholars from the management school of thought, member states with high administrative capacity face lower implementation constraints (Falkner et al. 2005; Kaufmann, Kraay, and Mastruzzi 2006; Mbaye 2001; Pridham 1994; Thomson, Torenvlied, and Arregui 2007). Because they have more administrative resources for coping with the implementation requirements, the likelihood for noncompliance is lower. Table 2 provides a detailed summary of our main explanatory and control variables with some explorative statistics.

The Empirical Evaluation

The Statistical Model

The game-theoretical model assigns probabilities to the three outcomes (compliance, enforced compliance, and compliance deficit) that result from the two types of actors’ expectations for the costs and benefits of their interdependent strategic decisions in the implementation process. To evaluate the empirical implications of their strategic interaction, we follow Signorino (1999, 2003) and apply a structural estimation model based on the probabilities of the compliance game (for specification of payoffs and utilities, please see the appendix). The equilibrium probabilities represent the strategic interaction with uncertainties between member states and monitoring Commission in the extended form game depicted in Figure 2.

From the structure of the game, which assumes utility-maximizing behavior of the two types of actors and independence between outcome probabilities, we can predict the equilibrium choices of an actor where the probability for each outcome is directly determined by the action probabilities along the game path. Thus, the probabilities are defined as:

\[ P(\text{Compliance}) = p_{\text{comp}}, \]
\[ P(\text{Enforced compliance}) = p_{\text{~comp}} \times p_{RO}, \]
\[ P(\text{Deficit}) = p_{\text{~comp}} \times p_{\sim RO}, \]

where \( p(\text{Compliance}), p(\text{Enforced compliance}) \) and \( p(\text{Deficit}) \) define the probabilities for compliance, enforced compliance, and compliance deficit, respectively. We translated the equilibrium-based strategic model into a statistical (probabilistic) model for our statistical estimation.

Figure 3 illustrates the general specification of the actor’s utilities in terms of regressors, i.e., their utilities with our explanatory variables. We estimate a member state’s utility for compliance as a linear function \( X_{11} \beta_{11} \), where \( \beta_{11} \) is a vector of coefficients to be estimated. The observed utility of the member states for a conflict with the monitoring Commission \( U_{MS}(\text{Enforced compliance}) \) is estimated as a linear function \( X_{13} \beta_{13} \) and the observed

<table>
<thead>
<tr>
<th>Theoretical Concepts</th>
<th>Operationalization</th>
<th>Min</th>
<th>Low</th>
<th>Mean</th>
<th>Moderate</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanctioning costs (Enforced compliance utility)</td>
<td>Member state’s disagreement</td>
<td>−0.96</td>
<td>−0.66</td>
<td>0.00</td>
<td>0.39</td>
<td>4.03</td>
</tr>
<tr>
<td>Member state’s saliency</td>
<td>−1.25</td>
<td>−0.66</td>
<td>0.00</td>
<td>0.64</td>
<td>2.76</td>
<td></td>
</tr>
<tr>
<td>Commission’s agreement</td>
<td>−1.12</td>
<td>−0.78</td>
<td>0.00</td>
<td>0.46</td>
<td>2.70</td>
<td></td>
</tr>
<tr>
<td>Commission’s saliency</td>
<td>−1.15</td>
<td>−0.76</td>
<td>0.00</td>
<td>0.78</td>
<td>2.06</td>
<td></td>
</tr>
<tr>
<td>Probability of enforcement success</td>
<td>Diversity of member states’ interests</td>
<td>−1.35</td>
<td>−0.78</td>
<td>0.00</td>
<td>0.42</td>
<td>2.57</td>
</tr>
<tr>
<td>Control variables (Enforced compliance utility)</td>
<td>Interest group pattern</td>
<td>−1.73</td>
<td>−0.79</td>
<td>0.00</td>
<td>0.87</td>
<td>1.31</td>
</tr>
<tr>
<td>Issue of timeliness</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of directive (1 - CM &amp; EP)</td>
<td>−1.30</td>
<td>−0.62</td>
<td>0.00</td>
<td>0.14</td>
<td>2.19</td>
<td></td>
</tr>
<tr>
<td>Delegation ratio</td>
<td>−1.85</td>
<td>−0.97</td>
<td>−0.03</td>
<td>0.66</td>
<td>1.80</td>
<td></td>
</tr>
<tr>
<td>Bureaucratic efficiency</td>
<td>−3.01</td>
<td>−0.34</td>
<td>−0.01</td>
<td>0.78</td>
<td>1.27</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Summary of the Independent Variables
utility of a member state for the central monitoring agency’s inaction $U_{MS}(\text{Deficit})$ as a constant $\beta_{14}$. The Commission’s utility for nonenforcement is normalized to zero, and its utility for enforced compliance is defined as a linear function $X_{23}\beta_{23}$ of explanatory variables. Following Signorino and Tarar (2006), we use the equilibrium-outcome probabilities of $p(\text{Compliance})$, $p(\text{Enforced compliance})$, and $p(\text{Deficit})$ as a basis for the maximum-likelihood estimation. The log-likelihood to be maximized with explanatory variables ($\beta$) is defined as:

$$
\ln L = \sum_{i=1}^{N} \left[ y_{\text{Compliance},i} \ln p_{\text{Compliance},i} + y_{\text{Enforced compliance},i} \ln p_{\text{Enforced compliance},i} + y_{\text{Deficit},i} \ln p_{\text{Deficit},i} \right], 
$$

where $y_{\text{Compliance},i} = 1$ if the strategic game in situation $i$ leads to compliance, $y_{\text{Enforced compliance},i} = 1$ if it results in enforced compliance by the Commission, and $y_{\text{Deficit},i} = 1$ if the outcome is compliance deficit.

**Empirical Analysis**

We test the empirical implications of our theoretical expectations on the outcomes that we derived from the strategic game between the monitoring Commission and the member states. In order to provide a unified analysis of compliance with directives and to capture strategic interaction, we adopt the strategic estimation approach of Signorino (1999, 2003). In spite of a skewed distribution of outcomes with about 65% compliance, 29% enforced compliance, and 6% compliance-deficit cases, our strategic model has a relatively high fit with over 71% correctly predicted outcomes. On closer inspection, we correctly
### Table 3 Strategic Probit Regression Based on the Model in Figure 3

<table>
<thead>
<tr>
<th></th>
<th>$U_{\text{MS}}$ (Enforced Compliance)</th>
<th>$U_{\text{MS}}$ (Compliance Deficit)</th>
<th>$U_{\text{Com}}$ (Enforced Compliance)</th>
<th>$U_{\text{MS}}$ (Compliance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member state’s disagreement</td>
<td>0.586**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.245)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member state’s saliency</td>
<td>−0.080</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.265)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commission’s agreement</td>
<td></td>
<td>−0.619</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.630)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commission’s saliency</td>
<td></td>
<td>3.837**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.327)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diversity of member states’ interests</td>
<td>0.282</td>
<td>−3.849**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.346)</td>
<td>(1.276)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interest group pattern</td>
<td>0.232</td>
<td>0.670**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
<td>(0.285)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of issues</td>
<td>−0.743**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.311)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issue of timeliness</td>
<td></td>
<td>4.228</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.347)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of directive (1 - CM &amp; EP)</td>
<td>1.063*</td>
<td>0.191</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.565)</td>
<td>(0.744)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of major provisions</td>
<td></td>
<td>−0.554***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.143)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delegation ratio</td>
<td></td>
<td>−0.352**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.173)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bureaucratic efficiency</td>
<td></td>
<td>0.134</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.139)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>−0.294</td>
<td>0.984</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.588)</td>
<td>(0.598)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>299</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>−197.607</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCP Outcomes</td>
<td>71.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCP Compliance</td>
<td>86.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCP Enforced Compliance</td>
<td>48.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCP Compliance Deficit</td>
<td>21.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* Standard errors in parentheses. Significance levels: **p < 0.001; *p < 0.05; *p < 0.10.

Predict more than 86% of the 195 compliance cases, while we overestimate their correct and timely implementation behavior vis-à-vis enforced compliance (48% correctly predicted) and compliance deficit (21% correctly predicted). Even though a direct comparison of the model fit between our strategic and standard probit models (estimates not shown here) is difficult and seems to be in favor of the standard probit models—since these models need to predict only two in contrast to three outcomes—our strategic model possesses an equally high model fit. Regarding the prediction sensitivity of the models, our strategic model for the enforcement outcome even outperforms the standard probit model with 48% to 24% correctly predicted enforcement cases, while it shows equally high prediction sensitivity for the compliance outcome with 86% to 85%.
see Rosendorff 2005). Table 3 reports the estimates of the four utility functions as specified in Figure 3. These coefficients are simultaneously estimated as part of a unified model of strategic interaction between the member states and the Commission.\(^\text{15}\)

According to the estimates of our strategic model, we find that the member states’ implementation decisions are mainly influenced by (dis)agreement and directive-specific features such as complexity, delegation ratio, level of discretion, and type of directive. We also find that the diversity of member states’ interests, the Commission’s saliency, and the influence of domestic interest groups significantly affect the enforcement decision of the central monitoring agency. Our estimates of the member states’ compliance utility confirm existing findings. Member states are more likely to defect when they disagree with a directive. The results further reveal that the member states’ utilities are higher for Council and EP directives than for directives decided solely by the Council. Furthermore, a member state’s utility for conflict with the Commission significantly decreases with the number of controversial issues included. All other variables, such as the diversity of the member states’ preferences and saliency for the directive, have no statistically significant effect on the member states’ compliance utility.

The results of our analysis reveal that the central monitoring agency acts strategically in the compliance game. Confirming our expectation, the Commission’s decision to litigate a noncomplying member state is also significantly influenced by the probability of success and the associated costs of sanctioning. In the event of preference diversity among the member states and a corporatist interest group pattern in the noncomplying member state, the Commission has difficulty finding enough support for successful enforcement. This decreases the central monitoring agency’s enforcement utility and hence decreases the probability of filing a reasoned opinion to a noncomplying member state. Likewise, the Commission is only willing to bear the costs of sanctioning when its saliency for the respective directive is high.

The compliance utility of the member states is significantly affected by two directive-specific factors: (1) the complexity of the directive and (2) the delegation ratio of implementation authority of the directive. This utility decreases with complexity and level-of-delegation ratio. Hence, member states are less likely to comply with very long and complex directives, which involve substantial implementation duties within their own legal system and government. However, the results suggest that the efficiency level of their administration has no significant effect on their decision. At first sight, this appears to contradict the management school of thought, but the coefficient indicates that the compliance utility of the member states still increases with the efficiency of their administration.\(^\text{16}\)

For a more substantial interpretation of our findings, we take a closer look at the marginal effects of our independent variables, i.e., the change in the estimated probabilities of the outcomes with the change in the values of the explanatory variables. According to our findings, the member states’ decision to comply with directives not only depends on their own compliance utility but also on the Commission’s utility to enforce compliance. Member states comply if and only if their true expected utility for noncompliance is greater than their true expected utility for compliance, which means that \(p_{\text{RO}}U_{\text{MS}}^s(\text{Enforced compliance}) - p_{\text{RO}}U_{\text{MS}}^s(\text{Deficit}) > U_{\text{MS}}^s(\text{Compliance})\). It is important to note that this is a function of the explanatory variables in the Commission’s enforcement utility because these variables affect not only the Commission’s decision but also indirectly the compliance decision of the member states. Compared to the variables that only reflect the central monitoring agency’s utility, and therefore only have an indirect effect on the member states’ decisions, the variables that enter in both the member states and Commission’s utilities have a direct and an indirect effect.

For the substantial interpretation of our main theoretical arguments, we plot the probability of enforced compliance and compliance deficit as a function of the diversity of member states’ preferences, their disagreement, and the level of the Commission’s saliency, while holding the other independent variables fixed at their low (dashed lines), mean (solid lines), and moderate (dotted lines) values.

Figure 4a displays the effect of the diversity of member states’ preferences on the probability of enforced compliance and compliance deficit. This strategy allows us to control for the possible heterogeneity effect of a directive by including a random effect for directives in the decision equation of the member state and the Commission. The results of these sequential probit regressions with random effects for directives show essentially identical results as our systematic estimator, confirming the previous significant results for our explanatory variables. However, here we only report the results of our system estimator, which is more efficient than the SBI estimator.

\(^{15}\)Column 1 of Table 3 displays the estimate \((\hat{\beta}_{13})\) for the member states’ enforcement utility, column 2 the member states’ utility for a compliance deficit \((\hat{\beta}_{14})\), column 3 the Commission’s enforcement utility \((\hat{\beta}_{15})\), and column 4 the member states’ implementation utility \((\hat{\beta}_{11})\), with standard errors in parentheses.

\(^{16}\)In order to check the robustness of our findings, we reestimated our model using a different estimation strategy including the sequential estimation of two probit regression models for the decision choices of the Commission and the member states in a manner analogous to backward induction (Bas, Signorino, and Mastruzzi 2008). This strategy allows us to control for the possible heterogeneity effect of a directive by including a random effect for directives in the decision equation of the member state and the Commission. The results of these sequential probit regressions with random effects for directives show essentially identical results as our systematic estimator, confirming the previous significant results for our explanatory variables. However, here we only report the results of our system estimator, which is more efficient than the SBI estimator.
compliance, where an increase in diversity is associated with a decrease in the utility of the Commission, making it less likely to enforce compliance, and with an increase in the member state’s utility for enforced compliance, making it them more likely not to comply and to risk an enforcement decision by the central monitoring agency. Accordingly, the diversity of interests has a nonmonotonic effect on the probability of enforced compliance. Considering the moderate scenario (dotted line), an increase in the diversity of member states’ preferences only for average to high values leads to a decrease in the probability of enforced compliance. In other words, in cases of a homogeneous preference constellation, where a large group of member states supports a directive, the noncompliance of only a few member states with the adopted policy goals increases the likelihood of enforced compliance. In such cases, the Commission still perceives itself as having considerable support for its enforcement decision among the other member states and will not hesitate to take action. However, the more member states do not comply with common policy goals, the less support will the Commission find for its enforcement decision, so that the likelihood of enforced compliance decreases.

In contrast to Figure 4a, Figure 4b shows a strict monotonic relationship between the diversity of member states’ preferences and the probability of a compliance deficit. Here, an increase in the diversity is always associated with an increase in the probability of a deficit. This is to be expected because an increase in diversity always favors the likelihood for noncompliance and always decreases the chance of strict enforcement by the Commission, even in cases where noncomplying behavior is detected. However, the point at which this effect becomes visible mainly depends on the values at which the other variables are held constant. This effect is largest when we hold all other values at their moderate value (dotted lines), while an increase in the probability of a compliance deficit is visible earlier for the low-value line (dashed line), followed by the mean-value line (solid line), and the moderate line (dashed line).

17 These results are in sharp contrast to the findings of Zhelyazkova and Torenvlied (2009) that the diversity of member states’ interests should have a strictly negative, monotonic effect on implementation delay, i.e., it should speed up the domestic implementation processes, and therefore should decrease the likelihood of enforcement and of compliance deficit.
Figures 4c and 4d show the relationship between a member state’s disagreement and the probability of enforced compliance and compliance deficit. Here, an increase in the level of disagreement is always associated with an increase in the probability of enforced compliance or compliance deficit. According to Figure 4d, however, the level of disagreement has only a weak effect on the probability of compliance deficit for low and mean values of the other variables, while it is absent for moderate values. This confirms the results of other studies which have pointed to the negative effect of a member state’s disagreement for compliance (Falkner, Hartlapp, and Treib 2004; Thomson, Torenvlied, and Arregui 2007; Zhelyazkova and Torenvlied 2009).

Finally, we consider the combined effect of our two main explanatory variables, i.e., the probability of conflict success, operationalized by the diversity of member states’ interests, and the costs of sanctioning, operationalized by disagreement of the member states. Figure 5 plots the probability of enforcement as a function of these two variables while holding the other variables at their mean value. It is clear that the likelihood of an enforced compliance increases substantially with a member state’s disagreement and its willingness to bear the costs of an enforcement conflict with the Commission. This effect, however, is only relevant in cases where the probability of successful enforcement by the Commission is high, i.e., when the diversity of member states’ interests is low.

When the central monitoring agency faces member states with diverse interests, the probability of enforced compliance is low regardless of the member states’ incentive for compliance.

**Conclusion**

How does the strategic interaction between a central monitoring agency and implementing actors affect compliance? Do low-level implementing actors consider and anticipate the decisions of the central monitoring agency to act against noncompliance? And is the central monitoring agency responsive to the chances of enforcement success and sanctioning costs by noncomplying actors? This study attempted to answer these questions by introducing a unified theory and test of the compliance game between a central monitoring agency and implementing actors. Using the centralized monitoring system of the EU as an example, we evaluated the empirical implications of our model, which considers the Commission as the central monitoring agency that decides about starting proceedings against noncomplying member states. Our strategic analysis reveals that the compliance outcome significantly depends on the probability of enforcement success and the potential sanctioning costs. We show that the Commission refrains from enforcement against a noncomplying member state in situations where the probability of success is low and/or the potential sanctioning costs are very high.

Although the central monitoring system of the EU offers some discretion for implementing the goals of directives, ranging from using country-specific legal acts to exempting some member states from implementing a directive within its own legal system, we believe that our analysis characterizes a mechanism with the probability of enforcement success and the potential sanctioning costs that also exists in other multilevel systems, such as the strictly mandatory U.S. implementation of federal law or the voluntary implementation of international agreements, such as the Kyoto Protocol. A typical example is the U.S. Environmental Protection Agency (EPA), which is responsible for scrutinizing implementation and enforcing compliance with environmental laws and regulations. Compared to the Commission, the EPA gathers data through on-site visits by qualified inspectors and also reviews information which has to be submitted. However, when the probability of enforcement success is low, e.g., due to large and powerful support of a defendant in U.S. Congress, or if the potential sanctioning costs are very high, e.g., due to long court procedures,
our model will also predict a compliance deficit. Similarly, our model suggests that the compliance committee of the Kyoto Protocol considers enforcement success with respect to coalition size of supporters and the potential sanctioning costs of the defendant, e.g., for exiting from the protocol.

Our identification of a compliance deficit by a strategic analysis of the compliance game between the central monitoring agency and the decentralized implementing actors might also explain a well-known puzzle in compliance research, according to which the central monitoring agency is almost always successful when it brings a noncomplying member state to court—indeed of the power relationship between the two types of actors. Because the central monitoring agency knows the distribution of the power and preferences of the member states, it can easily anticipate the probability of enforcement success and the potential sanctioning costs. We have shown that the central monitoring agency is prevented from taking action from taking action in less favorable situations where the enforcement success is low respectively and sanctioning costs are very high. This means that official compliance statistics may only provide a partial picture of the effectiveness of a central monitoring agency when these statistics ignore the level of compliance deficit resulting from the unwillingness or inability of the central monitoring agency to enforce compliance in less favorable situations.

We found that the central monitoring agency pre-selects noncompliance cases to take to court, which suggests a “conditionality” of the compliance-enhancing role of courts in those systems. According to the literature, courts can facilitate compliance with the adopted policy goals at the federal and international level because they act as a fire alarm by providing a venue where the actors can bring cases and signal possible violations (Carrubba 2005, 2009; Staton 2006; Vanberg 2004). However, courts cannot enforce compliance without a litigator, who needs to generate information on the compliance behavior of the collaborating actors. Since enforcement success may not only depend on the compliance behavior of one but of all collaborating actors, acquiring information to enforce compliance is costly in multilevel systems. Our findings show that a central monitoring agency cannot detect or is unwilling to do so and act in all cases of noncompliance. Whether each collaborating actor in horizontal monitoring systems or the voters, interest groups, or other domestic audiences in the monitoring system from below are able to bear these costs and can thus provide for a higher level of compliance by bringing more cases to court remains a topic for future research.

References


Hille, Peter, and Christoph Knill. 2006. “‘It’s the Bureaucracy, Stupid.’” *European Union Politics* 7(4): 531–52.


---

**Supporting Information**

Additional Supporting Information may be found in the online version of this article at the publisher’s website:

**Appendix**