doi:10.1017/S0003055412000020

Understanding the Role of the European Court of Justice in European Integration

CLIFFORD J. CARRUBA *Emory University*
MATTHEW GABEL *Washington University in St. Louis*
CHARLES HANKLA *Georgia State University*

In 2008 we published an article finding evidence for political constraints on European Court of Justice (ECJ) decision making. Stone Sweet and Brunell (this issue) argue that our theoretical foundations are fundamentally flawed and that our empirical evidence supports neofunctionalism over intergovernmentalism "in a landslide." We respectfully disagree with Stone Sweet and Brunell regarding both their conclusions about our theoretical arguments and what the empirical evidence demonstrates. We use this response to clarify our argument and to draw a clearer contrast between our and their perspective on the role the ECJ plays in European integration. Finally, we reevaluate their neofunctionalist hypotheses. Ultimately, we do not find support in the data for Stone Sweet and Brunell's empirical claims.

In 2008 we presented evidence for political constraints on the European Court of Justice's (ECJ) decision making (Carrubba, Gabel, and Hankla [CGH] 2008). Using a novel data set[1] of ECJ decisions from 1987–97, we proposed tests to identify the presence and substantive impact of threats of legislative override and noncompliance by member state governments (MSGs) in the European Union (EU). In this issue Stone Sweet and Brunell (2012) argue that the theoretical foundations of our article are fundamentally flawed and that the empirical evidence supports neofunctionalism over intergovernmentalism "in a landslide."

Stone Sweet and Brunell raise several important questions, but two points are particularly clear to us now. First, our article would have benefited from a variety of examples and a clearer exposition of the argument. We provide those here. Second, we could have done a better job of describing the "big picture" and answering this question: From our findings, what can we conclude about how the ECJ influences integration? In some important ways, our views contrast with those of Stone Sweet and Brunell (2012). In this article, we highlight these differences and discuss how our empirical evidence can help adjudicate among these and other accounts of ECJ decision making.

SUMMARY OF OUR ARTICLE

Our 2008 article focused on the ECJ, but it also addressed the more general question of whether governments influence judicial decision making. Scholars have debated this question for years, focusing their research primarily on domestic courts. So far, the evidence for legislative override as a constraint has been mixed (e.g., most recently Bailey and Maltzman 2011; Harvey and Friedman 2011; Segal, Westerland, and Lundquist 2010), and the debate remains unresolved. Scholars have found evidence that threats of noncompliance influence domestic courts (e.g., for Germany, Vanberg 2005; for Mexico, Staton 2010; for the United States, Carrubba and Zorn 2010). Unsurprisingly, scholars have argued that these same threats may influence ECJ rulings (e.g., Carrubba 2005; Garrett 1995; Garrett, Kelemen, and Schulz 1998; Garrett and Weingast 1993; Tsebelis and Garrett 1996; Wasserfallen 2010). In fact, a number of scholars argue that political constraints are important (CGH 2008, 436). Our article was an effort to provide quantitative evidence of the presence and magnitude of these constraints. In contrast to Stone Sweet and Brunell's (2012) characterization, it was not an effort to adjudicate between intergovernmentalism and neofunctionalism.[2]

Our Argument: The Threat of Legislative Override

Our argument for the threat of legislative override is based on the separation of powers literature (e.g., Ferejohn and Weingast 1992). An override occurs when a court's ruling is modified in subsequent legislation or treaty revisions. According to Wasserfallen (2010, 1131), the potential for such a legislative reaction can cause the ECJ to balance legal considerations against

Clifford J. Carruba is Associate Professor, Department of Political Science, Emory University, 327 Tarbutton Hall, 1555 Pierce Drive, Atlanta, GA 30322 (ccarrub@emory.edu).

Matthew Gabel is Professor, Department of Political Science, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130 (mgabel@artsci.wustl.edu).

Charles Hankla is Associate Professor, Department of Political Science, Georgia State University, P.O. Box 4069, Atlanta, GA 30302 (chankla@gsu.edu).

We gratefully acknowledge the financial support of the National Science Foundation (SES-079084). We thank Jeffrey Staton and Nathan Jensen for valuable comments on an earlier version of this response.

[1] We are unaware of any publicly available data set that covers the same time period, includes the same breadth of issues areas, and codes the same information.

[2] The only time we link our argument with this debate appears as a subpoint in a concluding paragraph (CGH 2008, 449). We meant to connect to claims over the likely magnitude of government influence, not to refute neofunctionalism.

the concerns of key legislative actors and particularly those of the MSGs. For example, Wasserfallen (2010, 1139) describes how the ECJ ruling in *Grzelczyk* (Case C-184/99) anticipated the formation of subsequent legislative coalitions by moderating its jurisprudence on student mobility.

In our 2008 article we argue that the court's responsiveness to such threats depends on the likelihood that a legislative coalition will form to counter a ruling. In the European Union, a legislative override requires that a decisive set in the council—all member states under unanimity, a supermajority under qualified majority—vote for a new policy. Revising a treaty requires unanimity. We assume that MSGs can construct such a coalition through logrolls. If a subset of governments considers an issue sufficiently important, it can trade concessions in other areas in exchange for legislative support from disinterested or otherwise opposed governments. In fact, a sizable portion of EU legislation involves logrolls, especially under unanimity (Aksoy n.d.).[3] Further, individual states, even small ones, have succeeded in building logrolls for treaty reform.[4] Thus, we argue, override is *possible* when at least one MSG wants to revise a court ruling.

Of course, the likelihood of success of a potential logroll depends on the initial distribution of MSG positions on the ruling. Consequently, we argue that the *credibility* of a threat of override depends on these initial conditions: The more MSGs expressing initial opposition to the ruling and the fewer expressing support, the more credible the threat. This conception of credibility is continuous.

This argument contrasts sharply with that of Stone Sweet and Brunell (2012). They dismiss the potential for logrolls and instead posit that override can only be achieved if enough MSGs explicitly support an override. They note (correctly) that, in the data, the distribution of MSG observations does not meet that threshold, and thus they dismiss override as a constraint on the ECJ.

Stone Sweet and Brunell (2012, 205) also claim that there is no evidence of an override of a significant case in the history of the ECJ. We disagree and provide an example later. It is important to note, however, that because our argument is about courts curtailing their behavior in anticipation of a possible override, providing more such examples or counting the number of instances of successful override is not informative. This is a standard issue with studying deterrence. If the court is responsive to, and accurate in its assessment of, threats of override, we should not see many instances of override.

Our example comes from the area of social security benefits.[5] In 1971, the European Community (EC) passed a regulation on the applicability of social security schemes to workers moving within the EC; it potentially extended member states' social security provisions to non-national residents. In a series of rulings in the late 1980s and early 1990s, the ECJ ruled that national welfare policies enjoy "exportability," and therefore a variety of national residency requirements for eligibility were invalid.[6] In the last of these rulings (*Stanton Newton*) in 1991, the ECJ issued a preliminary ruling against a national restriction. The restriction in question was supported by the two MSG observations (from the United Kingdom and Belgium). In 1992, the council amended regulation 1408/71 explicitly to insert language countering the effect of recent rulings by the ECJ. The court then capitulated. Conant (2002, 193) summarizes these events as follows: "The Member States overruled the ECJ with secondary legislation, which the ECJ could have overturned on the basis of Article 42 (51 EEC) of the treaty. The EJC declined this opportunity, however, when it chose to uphold the compatibility of the amended regulation with the treaty, leaving Member States with the last word in this particular dispute."

Thus, it is possible for MSGs to override an ECJ ruling through legislation. In addition, in this example the override was successful even though MSG observations against the ruling did not surpass the legislative threshold.

Our Argument: The Threat of Noncompliance

We define MSG "compliance" with EU law as the promulgation of community norms within their national territory. This process typically requires implementation, application, and enforcement by national institutions and may require the state to refrain from action (Krislov, Ehlermann, and Weiler 1986). Our use of the term "compliance" is consistent with several studies of compliance with ECJ decisions (e.g., Beach 2005; Krislov, Ehlermann, and Weiler 1986; Obermaier 2009; Steiner 1993). To clarify the relationship between our and Stone Sweet and Brunell's argument, we summarize our theory.

Our argument about compliance comes from Carrubba (2005), which argues that many international regulatory regimes are created to realize mutually beneficial cooperation by overcoming collective action problems. A common example is a treaty to liberalize trade. Such an agreement almost inevitably presents governments with a form of a prisoner's dilemma. All of the participant states perceive a benefit from mutual adherence with the proposed agreement, but they have no incentive to follow the rules unilaterally. A regulatory regime can help overcome this challenge

[3] Achen (2006, 293) also shows that EU legislative process is consistent with an "exchange model" that reflects cross-issue compromises typical of logrolling.

[4] The Maastricht Treaty included protocols on abortion in Ireland and property acquisition in Denmark.

[5] Other examples of override involve environmental policy and patient mobility. The MSGs countered the court's restriction of national government discretion in determining protective habitats for wild birds in its 1991 *Leybucht Dyke* ruling, Case C-57/89 (Cichowski 2007, 156–57; Nollkaemper 1997, 276–78). They also effectively nullified the *Pierik* judgment (Case C-17/177), which limited national discretion in authorizing payments for medical services (Martinsen 2006, 1038).

[6] *Stanton Newton* (C-356/89), *Commission v. France* (C-236/88; C-307/89), *Giletti and others* (C-381/85 and C-93/86).

by allowing states to tie the threat of retaliation or punishment to violations of the regime's rules (see also Simmons 1998).

The success of such a regime hinges on whether governments can effectively monitor and punish each other. Carrubba (2005) argues that courts can facilitate monitoring and punishment by acting as a *fire alarm* (a centralized venue in which to challenge possible acts of noncompliance) and an *information clearinghouse* (a venue in which participants make arguments in an attempt to establish whether a contested action is punishable).

A court knows its decisions do not automatically gain compliance. Rather a litigating government's decision to comply depends on whether it anticipates retaliatory punishment. Thus, the threat of noncompliance increases as the likelihood of retaliatory punishment decreases. Assuming that a court wants to avoid noncompliance, it is more likely to rule for the government as the threat of noncompliance by that government increases. Importantly, Carrubba (2005) argues that the decision to punish depends on the domestic context of the litigant government. The governments created the regulatory regime for a reason; they want to realize mutually beneficial gains from overcoming a collective action problem. But sometimes, for political, social, and economic reasons, the costs of compliance will be unusually high, and compliance no longer will be mutually beneficial. Thus, when the cost of complying with a ruling exceeds the benefits, the litigant government may be let "off the hook."

The empirical prediction (Hypothesis 2) in CGH (2008) follows. If a government can credibly threaten noncompliance (i.e., it is a physical option), the more third-party government support and the less opposition a litigating government receives, the less retaliation for noncompliance should be anticipated, and the more likely the court is to defer to the litigating government.[7]

This model of a regulatory regime has several important empirical implications for our tests and for several of Stone Sweet and Brunell's (2012) observations. First, the theory predicts that governments will be taken to court, be ruled against, and comply with court decisions on a regular basis. Second, we do not expect to see third-party governments routinely signaling their position on a case. The frequency of filings will depend on how often governments consider the cost of compliance to be especially high. Thus, the observed frequency of observations highlighted by Stone Sweet and Brunell (2012) is not dispositive. Third, the court in the model is regularly expected to adjudicate cases consistent with the substantive provisions of treaty or legal statutes defining the agreement. The governments created the regime because they wanted the rules enforced when doing so is mutually beneficial, and governments will have adopted rules so that the legally sanctioned application is generally mutually beneficial. Thus, we should observe many decisions made on the "legal merits." Fourth, the capacity of the court to promote its own agenda is limited. Governmental willingness to punish noncompliance determines the boundaries of the application of the agreement, and when those boundaries are reached, the court will tend to be deferential. Finally, this model does not mean that the court cannot innovate. If the court identifies new ways of applying the law that go beyond the written mandate and that governments see as mutually beneficial, governments may support the ECJ ruling or even enshrine it in future treaties.

Compare this argument to Stone Sweet and Brunell's claims. They (1998) argue that pro-liberalizing ECJ rulings increase opportunities for cross-border trade. By taking advantage of these new economic opportunities, pro-liberalizing interests grow and bring more cases challenging trade barriers to the ECJ, which in turn lead to more liberalization, and so on. Although MSGs do not act as a gatekeeper in their argument, Stone Sweet and Brunell's argument could be reconciled with our theory quite easily. Suppose, as pro-liberalizing domestic interests grow, that they induce reelection-seeking MSGs to support more liberalization via the standard electoral constraint. By shifting these induced preferences, new opportunities for mutual benefits may arise. If properly identified by the ECJ, these opportunities may allow it to promote an innovative liberalizing agenda, although the ECJ still cannot move beyond the newly identified mutually beneficial opportunities.

Stone Sweet and Brunell (2012, 205) also argue that noncompliance begets more litigation, which serves as a vehicle for expanding the ECJ's ability to construct EU law. They see this mechanism as contrary to our theoretical argument. But our theory allows for noncompliance followed by innovative jurisprudence. In particular, consider rulings that expand EU law and suffer noncompliance, but that are supported by third-party governments. In that setting, the punishment (or the threat thereof) should eventually end the noncompliance, ensuring implementation of the court's novel interpretation.

Where we differ from Stone Sweet and Brunell (2012) is in not expecting the court to develop EU law that enjoys compliance in areas where noncompliance with its rulings is supported by third-party governments. In particular, we predict that the ECJ is less likely to issue a ruling that requires national government compliance as the number of MSGs supporting the noncompliance increases. Our theory further implies that, when such rulings occur in the face of credible threats of noncompliance, the court will typically not be successful in developing EU law beyond the wishes of the MSGs. In sum, although we see tensions between our arguments and those of Stone Sweet and Brunell, we do not see the need for hard lines to be drawn. These arguments frequently can be complementary.

Stone Sweet and Brunell (2012, 205) assert that the development of the court's jurisprudence has never been stunted by threats of noncompliance and emphasize that the absence of such an example is critical

[7] The theory is necessarily about MSG observations against a litigant *constraining* the ECJ from making a ruling and for a litigant *permitting* the Court to make particular rulings (Stone Sweet and Brunell 2012, 207).

evidence against our argument. Once again, we are happy to provide an example to illustrate our point.[8] In two 1998 rulings, *Kohll* and *Dekker*, the ECJ determined that national laws requiring patients to obtain prior authorization for medical services delivered outside their home country violated EU law.[9] These rulings and the broader principles on which they were based had the potential to disrupt dramatically national health care systems. Numerous MSGs submitted observations in opposition to the court's position. We interpret these actions to mean that threats of noncompliance were likely credible. Indeed these rulings were met with widespread noncompliance (Obermaier 2009, 84) and the continued threat of noncompliance for similar subsequent rulings. In response, the ECJ significantly tempered its position in later rulings to satisfy the concerns of member state governments and thereby increase compliance (Obermaier 2009, 170). Obermaier (2009, 173) concludes, "Since the ECJ narrowed down the doctrinal content of the original rulings and limited the scope *de facto* to out-patient care, the actual effect diminished accordingly. Ultimately, the tremendous impact which was postulated by legal and political science scholarship, that is the massive change of the institutional configuration of domestic health care systems, did not come true."

Thus, the potential for noncompliance with its rulings seems to have caused the ECJ to alter its original decisions to placate the MSGs' interests. We also see that the noncompliance generated subsequent rulings, which is consistent with Stone Sweet and Brunell (2012, 205). Yet, the additional litigation did not serve to expand EU jurisprudence or to develop law beyond the desires of the MSGs. The result was quite the opposite. This illustrative example is useful for clarifying our argument and demonstrating its plausibility. However, as with the reasoning on threats of override, we do not consider the frequency of such events material to our argument. If our argument holds, the court should usually react to credible threats of noncompliance by tempering its rulings to reflect MSG interests and thereby ensure compliance. Therefore, we will not typically observe noncompliance when the court is influenced by such threats.

OUR EMPIRICAL ANALYSIS

Scope

We test our hypotheses with data from all rulings (1987–97), which consist primarily of three types: infringements, annulments, and preliminary rulings. Stone Sweet and Brunell (2012) claim that noncompliance is never an issue in annulment rulings, and they question its saliency to preliminary rulings. We suspect this disagreement reflects, at least in part, different uses of the term "compliance." We use "compliance" to mean the promulgation of a community norm within a national territory. Government compliance is therefore relevant when the promulgation of the norm in the ruling requires reform of existing national law or depends on some form of executive behavior. This kind of compliance is front and center in an infringement ruling, because the MSG is the defendant and the failure to comply with its legal obligations is the cause of action. On this we agree. Both annulments and preliminary rulings can also contest existing national laws or, more generally, require implementation by national institutions. In preliminary rulings, the ECJ ruling often articulates a general right, principle, or rule with application to the specific litigants in the national case (Steiner 1993, 4). Stone Sweet and Brunell (2012, 205) comment that national courts typically "comply" in the sense that the national court decision comports with the ECJ ruling; however, the national court can implement the ruling only in the instant case. The broader effect of the ruling may require changes to national law or administrative practices—that is, government compliance (Steiner 1993, 4). National courts cannot ensure the compliance of their governments. Indeed, according to Conant (2002, 69), when government compliance is in question in preliminary references, national officials often evade or ignore the ruling. It is this potential noncompliance that is at issue in our article. We do not, however, dismiss the potential effect that national courts could have on the credibility of threats of noncompliance; our Hypothesis 3 specifies and tests that expectation.

Similarly, noncompliance can be at issue in annulments where the action in question (e.g., a commission decision) may depend on a national government for implementation, especially when that government brings an annulment against the commission or the council. The bulk of these actions have addressed agriculture or competition policy (Bauer and Hartlapp 2010), both of which areas often rely on national government implementation. For example, a commission determination that state aid is invalid (i.e., in breach of competition policy) requires the national government to terminate the current subsidy and desist from creating a substitute policy. In particular, member states are responsible for the recovery of state aid that the court deems unlawful (Jestaedt, Derenne, and Ottervanger 2006). The inherent conflict of interest for member state governments that both grant and are charged with recovering state aids raises an obvious concern with the implementation of commission decisions requiring recovery. According to a comprehensive survey of the enforcement of these decisions, member state compliance is particularly problematic after an ECJ ruling confirming the commission determination—that is, when the MSG loses an annulment action (Jestaedt, Derenne, and Ottervanger 2006, 538).

Member states also are intricately involved in the implementation of the common agricultural policy. For

[8] Another example of the court responding to a threat of noncompliance involves the *Sheep Meat* case (C-232/78). France blatantly ignored the adverse ruling in that case. In a subsequent related enforcement action requesting an interim order, the court backtracked. Hartley (2007, 316) explains the turnaround as follows: "[O]ne suspects that the Court, knowing that any order it gave would be ignored, decided that it would be better to save what was left of its tattered authority by refusing the order."

[9] Cases C-120/95 and C-158/96.

example, in the annulment action *France v. Commission* (C-366/88), the French government challenged a measure adopted by the commission that obligated national officials to cooperate with commission officials in specific aspects of supervising and monitoring the European Agricultural Guidance and Guarantee Fund. Clearly, the effectiveness of an ECJ ruling for the commission would have depended on the compliance of the French government.[10]

Our use of "compliance" in annulment cases is not unique. Consider, for example, the conventional account of the annulment action brought by the United Kingdom in 1994 against the council on the Working Time Directive (Case C-84/94). The directive defined minimum rest periods and maximum weekly working hours. The United Kingdom then enjoyed an exception from EU social policies. Nevertheless, the council had adopted the measure under a treaty article that applied to the United Kingdom and that allowed the council to decide without the support of the United Kingdom. The UK government brought an action to annul this directive on the grounds that it had been adopted with the wrong legal basis.

The Advocate General's (AG) opinion in the case advised the ECJ to find the directive valid. In response, the UK government announced that it would not implement the directive if the court followed this recommendation: It threatened noncompliance. Pollack (2003, 333) described the context this way: "[T]he defiant mood of John Major's government, and the significant financial and political costs likely to be imposed by an adverse ECJ ruling, posed a non-trivial risk of overt *noncompliance* by the UK. This possibility of *noncompliance* was highlighted, moreover, by the British government's reaction to the preliminary opinion" [emphasis added].[11]

Of course, not all rulings involve threats of MSG noncompliance. Annulment actions can involve one EU institution challenging an act of another institution or a private litigant challenging the act of an EU institution. In these settings, government compliance is less often relevant to implementation (e.g., see *Partie Ecologiste 'Les Verts' v. European Parliament*; Case 294/83). Similarly, preliminary rulings involving two private parties (a frequent event) also are less likely to raise compliance challenges. In short, compliance figures more prominently in cases with national government litigants than otherwise.

Stone Sweet and Brunell (2012) gloss over this distinction, although it is central to distinguishing between our first and second hypotheses. Although override is always a possibility, the threat of MSG noncompliance is not. Because we posit that the court should be more sensitive to threats of noncompliance than threats of override—they are easier to execute—and because threats of noncompliance are primarily present when governments are litigants, we predict that the court's decisions will be more sensitive to MSG observations when a government is a litigant than when one is not.

Measurement and Statistical Tests

We operationalized our measure of third-party government preferences by using a weighted function of government observations on cases (CGH 2008, 440). For the purposes of this discussion, the tests of our argument are the following. First, we tested for threats of override when governments are not litigants and for threats of noncompliance when they are. We found statistically significant relationships between government observations and ECJ decisions when governments are not litigants, and a much stronger relationship when they are (Table 1, first column). Second, we included the Advocate General as a control for the legal merits of the case and found the exact same relationships (CHG 2008, 446). In both analyses we controlled for the position of the commission.

We make two points. First, in CGH (2008), our operationalization of government observations followed the logic of our theories. We measured the impact of both MSG support and opposition to litigant governments in the "net weighted observations for the plaintiff" and in the unweighted net measure used in the supplemental analyses presented in the Appendix of CGH (2008).[12] Thus, we captured net support for a particular litigant's position. The more net support for a position, the more credible the threat to a ruling against that position, and therefore the more likely the court is to defer to that position.

Second, we treated government briefs as a proxy for the underlying concept we wanted to measure: government preferences on the outcome of the case. If a direct measure of an MSG's preferences were available, we would certainly use it. Nonetheless, these briefs can be credible signals of MSG preferences. Might briefs correlate with court decisions for reasons other than political pressure? Yes. In particular, briefs may reflect, support, and/or substantiate the side with the stronger legal argument. Accounting for these alternatives is critical, and we addressed this issue in two ways. First, if the briefs are simply a function of the legal merits, the influence of briefs on court decisions should not depend on whether a government is a litigant. However, we did observe such a conditional effect. We also included the Advocate General's position as a control for the merits of the case and found support for political constraints.[13]

Second, we included the commission's position as a control, because we have theoretical reasons to believe its position might be correlated with both the court's decision and with the positions that governments take.

[10] The court ruled for the French government.

[11] The two MSG observations supported the council position. According to our theoretical expectations, this should have increased the likelihood the ECJ would rule against the United Kingdom and the United Kingdom would comply with the ruling. This is exactly what happened. See also Beach (2005) for treatment of this case.

[12] The unweighted measure obviates Stone Sweet and Burnell's (2012, 206) claim that we stack the deck by assuming qualified majority voting in our analysis.

[13] The Advocate General is a nonpolitical officer of court who recommends a decision on each case before the ECJ rules. See CHG (2008, 447–49) for further discussion.

TABLE 1. Probit Models of ECJ Ruling for the Plaintiff

	Original Model	With AG control
Net weighted observations for plaintiff	1.18*	0.94*
	(0.33)	(0.36)
Net weighted observations for plaintiff when government litigant	2.19*	1.80*
	(0.90)	(0.96)
Government litigant	0.20*	0.19*
	(0.08)	(0.08)
Commission is plaintiff	0.04	−0.54*
	(0.10)	(0.12)
Commission is defendant	−0.74*	−0.73*
	(0.10)	(0.12)
Commission observation for plaintiff	0.99*	0.79*
	(0.08)	(0.09)
Commission observation for defendant	−0.60*	−0.26*
	(0.08)	(0.09)
Advocate General for plaintiff		1.70*
		(0.08)
Chamber size	−0.21*	−0.22
	(0.12)	(0.14)
Constant	−0.08	−1.01*
	(0.08)	(0.10)
N	3,176	3,176

Notes: Robust standard errors reported in parentheses. *p < .05, one-tailed test.

Its position may be an indicator of the legal merits of the case; the commission's observation may serve as a "political bellwether" indicating to the court what would be acceptable to the MSGs (Burley and Mattli 1993, 71), or the court might agree with or be responsive to a commission pro-integrationist agenda. Thus, the commission's position may be correlated with the ECJ's decision, and for several of these reasons (i.e., the merits argument or responsiveness to MSG interests), the commission's position could also be correlated with the position of the government briefs.

Thus, Stone Sweet and Brunell's (2012) claims about our interpretation of MSG observations are incorrect. First, we agree that these briefs do not "directly evaluate the proposed *explanation* of this influence" and that one should not "presume that the outcome can only be explained by 'political' threats" (Stone Sweet and Brunell 2012, 206). We minimize the odds that our analysis suffers from these problems by accounting for plausible alternative explanations for a correlation between government briefs and court rulings. Second, we do not, as Stone Sweet and Brunell (2012, 206) claim, count all instances of congruence between the position of MSG observations and the ECJ ruling as evidence of a political constraint. Only congruence independent of these other various factors counts as evidence supporting our hypotheses.

MEMBER STATE GOVERNMENTS, THE COMMISSION, AND THE ECJ

Stone Sweet and Brunell (2012) use our data to examine whether a neofunctional argument provides a better account of ECJ decision making than either an intergovernmental argument or our specific claims. This is an important question, but their empirical analysis does not provide appropriate evidence to answer it. We use the same data to reanalyze their claims.[14]

Does Neofunctionalism Account for Our Evidence Regarding MSG Observations?

Stone Sweet and Brunell (2012, 211) assert that neofunctionalism predicts that MSGs have more influence on the ECJ's rulings when MSGs encourage the ECJ to punish a member state and when they are joined by the commission. Consequently, they contend that, in preliminary rulings, MSG observations for the plaintiff should be more influential than ones for the defendant, and the impact of MSG support for the plaintiff should be strongest when the commission also supports the plaintiff. They present what amount to cross-tabs (without test statistics) showing that, in preliminary rulings, the ECJ is more likely to side with the position supported by the MSGs if they support the plaintiff. The plaintiff's success rate is even greater if both MSGs and the commission support the plaintiff.

[14] Stone Sweet and Brunell (2012, 210) incorrectly assert that we "declare that, although 'the ECJ may, on the margin, favor the commission,' the MSGs are the actors who count, since they 'systematically' constrain the Court," and they criticize us for never pitting the commission against the MSGs in a head-to-head battle. Although we include the commission's position as a control, nothing in our argument asserts an expectation over its importance relative to that of the MSG observations. That commission support also matters does not change our conclusion; it only suggests that other variables matter as well.

TABLE 2. Probit Models of ECJ Ruling for the Plaintiff

	Model 1	Model 2	Model 3
Commission observation for plaintiff	0.58*	0.62*	0.65*
	(0.11)	(0.11)	(0.11)
Commission observation for defendant	−0.36*	−0.38*	−0.41*
	(0.11)	(0.11)	(0.11)
Pro-plaintiff government observations	1.97*	3.34*	1.25
	(0.71)	(0.96)	(0.82)
Pro-defendant government observations	−0.89*	−0.88*	−0.36
	(0.52)	(0.52)	(0.69)
Pro-plaintiff government observations with commission observation for plaintiff		−2.23*	
		(1.24)	
Pro-defendant government observations with commission observation for plaintiff			−1.22
			(0.96)
Pro-plaintiff government observations with commission observations for defendant			2.21*
			(1.32)
Advocate General for plaintiff	1.99*	1.98*	1.98*
	(0.09)	(0.09)	(0.09)
Chamber size	0.016	−0.002	0.001
	(0.19)	(0.19)	(0.19)
Constant	−1.09*	−1.09*	−1.09*
	(0.12)	(0.12)	(0.13)
N	2,048	2,048	2,048

Notes: Robust standard errors in parentheses. *$p < .05$, one-tailed test. The hypotheses all specify a directional expectation, which calls for a one-tailed test (Kelstedt and Whitten 2009, 176).

This does not constitute an appropriate test. First, their measure of the MSGs' position does not provide a test of our hypotheses. Our expectation is that the ECJ is more responsive to MSG observations for the plaintiff or defendant as the number of such observations grows. We therefore constructed a continuous measure of net MSG support. Stone Sweet and Brunell (2012, 211) use a dichotomous measure, indicating only whether MSGs were net on the side of the plaintiff or defendant. This measure ignores important variation in the net number of MSGs on the litigants respective sides and fails to distinguish the economic or political clout of the observers. This variation is also relevant to their argument; MSG observations indicate potential punishment of a defendant state (Stone Sweet and Brunell 2012, 211).

Second, they do not control for the legal merits of the ruling. The commission and the MSG position may convey information about a ruling's legal merits. If so, the greater frequency of success of plaintiffs supported by these actors (Stone Sweet and Brunell 2012, 211) could simply indicate that the legal merits were more often on the plaintiff's side. We therefore control for the legal merits.

In Models 1 and 2 in Table 2 we report the results for our tests of Stone Sweet and Brunell's hypotheses with corrections for the two problems.[15] The statistical model includes separate variables for each litigant, which allows us to assess whether the effect of MSG observations is stronger when they support the plaintiff. These two variables are derived from the original variable, *net weighted observations for the plaintiff*, defined in CGH (2008, 440). Each variable is coded as zero in the absence of net observations supporting the specific litigant and as the absolute value of net weighted observations for that litigant otherwise. Thus, both measures are continuous. In Model 2, we add an interaction term to capture the conditional effect of MSG observations for the plaintiff when the commission supports the plaintiff. If Stone Sweet and Brunell's hypotheses are correct, we should find that the coefficient on MSG observations for the plaintiff is larger than the absolute value of the coefficient for MSG observations for the defendant and that the interaction term is positive.

The results do not support Stone Sweet and Brunell's expectations. First, as shown in Model 1, the size of the positive impact of an increase in pro-plaintiff MSG observations is not statistically different from the size of the negative impact of an increase in pro-defendant MSG observations ($p > .26$). Second, as shown in Model 2, we find that the impact of MSG observations for the plaintiff is not greater when the commission also supports the plaintiff. To see this, one must consider the coefficients from both the main effect for pro-plaintiff MSG observations and the interaction term (Brambor, Clark, and Golder 2006). The sum of these two coefficients ($3.34 + -2.23 = 1.11$) is smaller than the coefficient for MSG observations for the plaintiff when the commission does not favor the

[15] Following the analysis in CGH (2008), we also include a control for the chamber size.

plaintiff (3.34). These findings do not support Stone Sweet and Brunell's (2012, 211) claims in favor of neofunctionalism.

Stone Sweet and Brunell (2012) also assessed whether the ECJ is more likely to rule for the commission position or the MSG position when they are in conflict. They claimed (211) that the critical question is how the ECJ responds when the commission's brief is in opposition to the MSG observations and that neofunctionalism predicts that the commission position trumps the MSG observations. They do not provide appropriate tests for this claim: for reasons cited earlier.[16]

Model 3 of Table 2 presents an appropriate test. It adds two interaction terms to the variables in Model 1,[17] each of which captures the condition when the commission and MSG observations are in opposition. The results are counter to those of Stone Sweet and Brunell (2012). First, with the commission supporting the defendant, an increase in MSG observations for the plaintiff is associated with a greater increase in the likelihood of the court ruling for the plaintiff than when the commission does not support the defendant. The sum of the coefficients indicating the impact of MSG observations for the plaintiff when the commission supports the defendant ($2.21 + 1.25 = 3.46$; $p < .01$) is greater than the coefficient for government observations for the plaintiff (1.25) when the commission does not support the defendant. Second, when the commission supports the plaintiff, an increase in MSG observations for the defendant is associated with a decrease in the likelihood of a ruling for the plaintiff, and the magnitude of this effect is at least as large as when the commission does not support the plaintiff. The sum of the coefficients capturing the impact of MSG observations for the defendant when the commission supports the plaintiff ($-0.36 + -1.22 = -1.58$; $p < .05$) indicates a larger negative impact than the coefficient for government observations for the defendant (-0.36) when the commission does not support the plaintiff. That is, the presence of an opposing commission observation fails to temper the impact of MSG observations on the court. More generally, the three models presented in Table 2 do not support Stone Sweet and Brunell's claims that our data favor their neofunctionalist hypotheses.[18]

Substantive Significance of the Commission, MSG, and AG positions

Because CGH (2008) focused primarily on estimating the effect of MSG observations, we did not assess or compare the impact of other factors on ECJ rulings. Yet clearly, our understanding of ECJ decision making would be enhanced by such an assessment. Stone Sweet and Brunell pursued one such assessment: a comparison of the substantive significance of commission and MSG observations. Our data can shed light on the broader question of what accounts for ECJ rulings.

Unfortunately, Stone Sweet and Brunell's statistical evidence does not make appropriate comparisons and omits consideration of a key factor in ECJ decision making. First, their Figure 1 (210) is a visual comparison of the change in the predicted probability of a ruling for the plaintiff associated with changes in net MSG observations and with the presence of an observation from the commission. This figure is based on the first model presented in Table 2 in CGH (2008). That model lacks both the interaction term with the government litigant used to capture threats of noncompliance and a control for the legal merits. This is not the appropriate statistical model.

Second, this sort of visual inspection is potentially misleading (Gelman and Hill 2007). The basic problem is that the predicted probabilities calculated at different values on the variable of interest (e.g., commission observations for plaintiff) must assume particular values for the other variables in the statistical model. The selection of these values can dramatically alter the predicted probabilities. For this reason, Gelman and Hill (2007, 100–3) recommend assessing the substantive significance at each observed value for the other variables and then averaging across all of these values to generate a summary of the substantive effect. This is called the "average predictive comparison." In the context of our data, this measure requires estimating the change in the predicted probability of a pro-plaintiff ECJ ruling due to a shift (e.g., from its lowest to its highest observed value) in the independent variable of interest, averaged across all observed values of the other variables.

We can also evaluate the impact of the position of the AG, which serves as a proxy for the legal merits. Because students of the court routinely assert that the ECJ rules in response to the merits, our accounting of the factors shaping ECJ decision making should include the impact of the AG's position.

We use the probit model presented in the last column of Table 1 to generate average predicted comparisons for these three factors. This model is based on Model 2 in Table 2 in CGH (2008), which is reproduced in the first column of the table, but also includes a control for the position of the AG.[19] Consequently, this model allows us to assess the threat of noncompliance and

[16] A general problem with their statistical analysis is that the parsing of the data results in very few observations in the categories with a neutral commission position (30 with a pro-plaintiff MSG position; 85 with a pro-defendant MSG position). The obvious consequence is imprecise estimates (i.e., large standard errors). Nevertheless, Stone Sweet and Brunell (2012) focus on the relatively large standard errors on coefficients for these categories as evidence that MSG observations have no impact.

[17] We also have estimated a fully saturated model with all possible interactions between MSG and commission variables. The results are consistent with the conclusions drawn here. Such a model is extremely complicated to interpret and parses the data unnecessarily.

[18] These results reported in Table 2 obtain with or without the weighting of government observations.

[19] CGH (2008) include a model with a control for the AG's position. That model (Model 2, Table 4) also includes a set of very complicated interaction terms designed to test Hypothesis 3. Because Hypothesis 3 is not in question here, we focus on this simpler model.

the impact of the legal merits. We calculated average predictive comparisons on the variables of interest. The average increase in the likelihood of a pro-plaintiff ruling associated with a change from a commission observation for the defendant to one for the plaintiff is 0.28. The average change associated with a low to high shift for MSG observations is 0.31 without a government litigant and 0.52 with a government litigant.[20] The average change associated with a change in AG position between supporting the defendant and the plaintiff is 0.53. These results are not specific to the time period studied in CGH (2008). In a larger data set (1960–1999) the average predictive comparisons support the same conclusions.[21]

In sum, we find no evidence of a landslide in favor of any one factor: All three have substantively important independent effects on the direction of ECJ rulings. However, we are intrigued by the impact of commission observations. Because the statistical model controlled for the position of the AG, we are fairly confident that the observed impact of the commission does not simply reflect the legal merits on the specific legal question at issue. This rules out one plausible interpretation of the relationship between the ECJ rulings and the commission's position. However, as we discussed earlier, several plausible interpretations remain. Given the sizable impact of commission observations on rulings, identifying the appropriate interpretation is important to our understanding of ECJ decision making.

CONCLUSION

The value of our research depends on scholars understanding our arguments and empirical strategy. We hope the clarification of our argument and the illustrative examples of ECJ cases have helped on this front. We also hope our statistical analysis of the hypotheses from Stone Sweet and Brunell (2012) distinguishes among alternative accounts of patterns in our data.

How does our understanding of the ECJ compare to that of Stone Sweet and Brunell? They argue that the ECJ is a strongly independent court that can and does promote European integration; although the ECJ is influenced by government preferences, that influence is only on the margin.[22] We agree that the ECJ promotes European integration. However, we argue that ECJ rulings help facilitate integration *because* of government preferences, not in spite of them. Only if the ECJ tries to go beyond the role of facilitator will it struggle with noncompliance and override. Stone Sweet and Brunell also argue that the ECJ's rulings have ultimately empowered domestic forces that use the international system to further their integrative agendas. We see no obvious conflict with our argument on this issue. Our only difference is that we see this effect working because those domestic forces, through the electoral constraint, induce their governments to internalize the benefits for further integration. Thus, ultimately, we do not disagree on much of the big picture, but on some of the "whys" and "hows."

[20] The range is −.87 to .53 without a government litigant and −0.5 to 0.33 with one.

[21] The larger database, including the disaggregated information for variables of interest, has been available online since March 2010 at http://www.polisci.emory.edu/faculty%20pages/carrubba.htm. We posted an audit of a random subsample of cases at the same web address. The average error rate was less than 1%. The maximum error rate for any variable was less than 2.5%.

[22] It is not clear from their argument exactly how and why this should be true, but it is clearly stated.

REFERENCES

Achen, Christopher. 2006. "Evaluating Political Decision-making Models." In *The European Union Decides*, eds. Robert Thomson, Franz Stokman, Christopher Achen, and Thomas Koenig. New York: Cambridge University Press, 264–99.

Aksoy, Deniz. N.d. "Institutional Arrangements and Logrolling: Evidence from the European Union." *American Journal of Political Science*. Forthcoming.

Bailey, Michael, and Forrest Maltzman. 2011. *The Constrained Court*. Princeton, NJ: Princeton University Press.

Bauer, Michael, and Miriam Hartlapp. 2010. "Much Ado about Money and How to Spend It! Analysis of 40 Years of Annulment Cases against the European Union Commission." *European Journal of Political Research* 49: 202–22.

Beach, Derek. 2005. "Why Governments Comply: An Integrative Compliance Model that Bridges the Gap between Instrumental and Normative Models." *Journal of European Public Policy* 12 (1): 113–42.

Brambor, Thomas, William Clark, and Matthew Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14 (1): 63–82.

Burley, Anne-Marie, and Walter Mattli. 1993. "Europe Before the Court." *International Organization* 47 (1): 41–76.

Carrubba, Clifford. 2005. "Courts and Compliance in International Regulatory Regimes." *Journal of Politics* 67 (3): 669–89.

Carrubba, Clifford, Matthew Gabel, and Charles Hankla. 2008. "Judicial Behavior under Political Constraints." *American Political Science Review* 102 (4): 435–52.

Carrubba, Clifford, and Christopher Zorn. 2010. "Executive Discretion, Judicial Decision-making, and Separation of Powers in the United States." *Journal of Politics* 72 (3): 812–24.

Cichowski, Rachel. 2007. *The European Court and Civil Society*. New York: Cambridge University Press.

Conant, Lisa. 2002. *Justice Contained*. Ithaca, NY: Cornell University Press.

Ferejohn, John, and Barry Weingast. 1992. "A Positive Theory of Statutory Interpretation." *International Review of Law and Economics* 12: 269–79.

Garrett, Geoffrey. 1995. "The Politics of Legal Integration." *International Organization* 49 (1): 171–81.

Garrett, Geoffrey, R. Daniel Kelemen, and Heiner Schulz. 1998. "The European Court of Justice, National Governments, and Legal Integration in the European Union." *International Organization* 52 (1): 149–76.

Garrett, Geoffrey, and Barry R. Weingast. 1993. "Ideas, Interests, and Institutions: Constructing the EC's Internal Market." In *Ideas and Foreign Policy*, eds. Judith Goldstein and Robert Keohane. Ithaca, NY: Cornell University Press, 173–206.

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Hartley, T. C. 2007. *The Foundations of European Community Law*. 6th ed. New York: Oxford University Press.

Harvey, Anna, and Barry Friedman. 2011. "Pulling Punches: Congressional Constraints on the Supreme Court's Constitutional Rulings, 1987–2000." *Legislative Studies Quarterly* 31 (4): 533–62.

Jestaedt, Thomas, Jacques Derenne, and Tom Ottervanger (coordinators). 2006. *Study on the Enforcement of State Aid Law at National Level*. DG Competition, European Commission. Luxembourg: Office for Official Publications of the European Communities.

Kellstedt, Paul, and Guy Whitten. 2009. *The Fundamentals of Political Science Research*. New York: Cambridge University Press.
Krislov, Samuel, Claus-Dieter Ehlermann, and Joseph Weiler. 1986. "The Political Organs and the Decision-making Process in the United States and the European Community." In *Integration through Law*, *Vol. 1*, eds. M. Cappelletti, M. Seccombe, and J. Weiler. Berlin: Walter de Gruyter, 3–110.
Martinsen, Dorte. 2005. "Towards an Internal Health Market with the European Court." *West European Politics* 28 (5): 1035–56.
Nollkaemper, Andre. 1997. "Habitat Protection in European Community Law: Evolving Conceptions of a Balance of Interests." *Journal of Environmental Law* 9: 271–86.
Obermaier, Andreas. 2009. *The End of Territoriality?* Burlington, VT: Ashgate.
Pollack, Mark. 2003. *The Engines of European Integration*. Oxford: Oxford University Press.
Segal, Jeffrey, Chad A. Westerland, and Stephanie A. Lundquist. 2010. "Congress, the Supreme Court and Judicial Review: Testing a Constitutional Separation of Powers Model." *American Journal of Political Science* 55 (1): 89–104.
Simmons, Beth. 1998. "Compliance with International Agreements." *Annual Review of Political Science* 1: 75–93.
Staton, Jeffrey. 2010. *Judicial Power and Strategic Communication in Mexico*. New York: Cambridge University Press.
Steiner, Josephine. 1993. "From Direct Effect to Francovich." *European Law Review* 18: 3–22.
Stone Sweet, Alec, and Thomas Brunnell. 1998. "Constructing a Supranational Constitution: Dispute Resolution and Governance in the European Community." *American Political Science Review* 92 (1): 63–81.
Stone Sweet, Alec, and Thomas Brunell. 2012. "The European Court of Justice, State Non-compliance, and the Politics of Override." *American Political Science Review* 106 (1): 204–14.
Tsebelis, George, and Geoffry Garrett. 1996. "An Institutional Critique of Intergovernmentalism." *International Organization* 50: 269–99.
Vanberg, Georg. 2005. *The Politics of Constitutional Review in Germany*. New York: Cambridge University Press.
Wasserfallen, Fabio. 2010. "The Judiciary as Legislator? How the European Court of Justice Shapes Policy-making in the European Union." *Journal of European Public Policy* 17 (8): 1128–46.