How can international organizations shape national welfare states? The answer depends on why national governments comply with international organization mandates. Enforcement theories predict that states’ policy preferences determine implementation, whereas managerial theories attribute noncompliance to states’ capability limitations and to institutions. This article derives specific compliance mechanisms from these theories and examines the implementation of European Union social policy directives through a new quantitative data set and qualitative case studies of Greece and Spain. Countries whose preferences diverge from social policy directives, specifically countries lacking related early national legislation, and countries with low labor costs, delay implementation. However, delays from capability limitations are much greater—poor bureaucracies, federal states, coalition governments, and parliaments that do not prepare for directives cause big delays. These findings suggest that international organizations can shape national social policies by reorienting the axes of contestation from left-right to supranational-subnational.

Keywords: compliance; transposition; social policy; employment policy; European Union

Globalization challenges national welfare systems and stretches the creative capacities of citizens, political leaders, and academics to invent suitable policy responses. Some labor advocates, fearing the erosion of worker rights, fight to limit free trade, whereas others seek to strengthen domestic labor laws. Other groups propose compensating displaced workers or toughening labor standards in multinationals’ corporate codes of conduct. Parallel to these national efforts, a more coordinated alternative is on
the rise—international legislation regulating domestic labor and social practices. Although the International Labour Organization (ILO) promotes a new wave of core labor rights, and the EU, NAFTA, MERCOSUR, and ASEAN develop regional responses, academics delay in including these increasingly important solutions in their studies. As supranational organizations promote new and potentially powerful labor norms, a key question emerges: Will national governments comply with their international obligations, and how will this shape national social systems?

In Europe, key employment and social policy decisions are already made supranationally. EU directives are legal instruments that bind states to reach specific policy objectives within a given timetable but permit some flexibility in the specific means employed to reach these common goals. From maternity leave minima to maximum working-week hours, from the concentration of chemicals in the workplace air to the structure of works councils, social policy designed at the EU level can have a large impact on member state activity in this domain. Almost all directives eventually get implemented. Critically, however, only a minority get implemented on time, and there is substantial variation in implementation delays. Indeed, an implementation gap threatening the European project has been much debated and decried (Börzel, 2001; Snyder, 1993; Tallberg, 2002). Spurred by a wealth of fascinating but frequently contradictory case studies concerning the implementation of particular directives, this study examines factors that shape the implementation of more than 50 directives in 12 member states during a 15-year period and then focuses in on 2 countries where compliance difficulties often arise—Greece and Spain. Understanding the obstacles to directive implementation in the EU can in turn offer insights into likely obstacles to compliance with other emerging international and regional initiatives.

Compliance depends on states’ rational calculation of the costs of changing their policy to meet international commitments, weighed against the benefits this change might bring, according to Downs, Rocke, and Barsoom (1996), key exponents of the enforcement model. As compliance can often be costly,
careful monitoring and strict enforcement by international organizations best tips the balance in its favor. In contrast, the managerial thesis, associated most clearly with Chayes and Chayes (1993, p. 188), attributes compliance failures not to rational calculation but instead to treaty ambiguity, capacity limitations of states, and uncontrollable social and economic changes. In Young’s (1992) formulation, “The effectiveness of international institutions varies directly with the capacity of the governments of members to implement their provisions” (p. 183). Whether a coercive strategy of monitoring and sanctions is preferable to a cooperative managerial approach, promoting capacity building and rule clarification, is a live debate whose resolution has been hindered by severe theoretical and methodological obstacles. Drawing testable hypotheses from these overarching theories moves the theoretical debate forward by demonstrating the relative importance of specific factors in particular contexts. For clarity, hereafter variables associated with the enforcement model are labeled “preferences measures,” and variables associated with the managerial model are labeled “capability variables.”

Several authors examine the factors that influence the adoption of EU directives in a quantitative fashion (Bergman, 2000; Lampinen & Uusikyla, 1998; Martin, 2000; Mbaye, 2001). These important works shed light on some of the facets of directive implementation and point to the issue’s significance. Nevertheless, none of these works explicitly juxtaposes major theories of compliance, and none of them directly examines the dependent variable of interest—the delay in directive implementation. These two methodological limitations call for a reexamination of the data.

This study overcomes obstacles to the testing of dominant theoretical approaches through three innovations: a research design that permits the explicit juxtaposition of dominant theories, a data set that measures the dependent variable directly, and novel measures for relevant independent variables. Instead of using the entire set of directives and trying to surmise government preferences from the limited information of the voting record, this study focuses on a subset of directives on which government preferences can be specified well. Because the literature on social policy is well developed, the combination of information on party orientation, national labor costs and unemployment levels, government voting records, and past legislative efforts at the national level should provide substantial information on national preferences. At the same time, information on federal structures, government coalitions, parliamentary procedures, and overall bureaucratic transparency permit us to better understand institutions that can influence government capacity for prompt implementation.
Hypotheses and Operationalization

This section transforms broad international relations theories into testable propositions.

Hypotheses Concerning Directive Congruence With National Preferences

It seems very plausible that countries whose national objectives coincide with particular directives are more likely to implement these. Nonetheless, advocates of the enforcement model clash with supporters of the managerial model on the centrality of this theme as an explanatory variable. A first measure of the congruence between a directive and the national interest of a member state is an affirmative vote in the Council of Ministers. We expect negative votes and abstentions to correlate positively with implementation delays but not to explain much of the variance in these delays. This is because negative votes and abstentions are infrequent, as both enthusiastic supporters and more hesitant members often vote affirmatively unless the issue is of great national importance.

Hypothesis 1: Countries that voted against a particular directive are more likely to implement it slowly.

A second proxy for national preferences on a particular subject is national-level legislation. The hypothesis states that countries that have pioneered legislation in the field of a directive will be more likely to implement the directive promptly. This hypothesis rests on the extensive literature on European welfare states, which describes the process of policy development as one of accretion, with strong institutional dynamics impeding cutbacks and past choices constraining future options (Pierson, 1996). An analogous argument is made in several case studies on EU directive implementation, which argue that the fit of the directive with national institutions determines the likelihood of its timely implementation (Duina & Blithe, 1999; Knill, 2001). Testing for this hypothesis leaves open the possibility that we will discover a negative relationship, as reform may in some cases be costlier than starting from scratch. This, however, seems unlikely to be the dominant trend as directives set minimum standards, permitting countries that already have higher national standards to keep these intact.
The ILO Natlex database offers a thorough compilation and categorization of national legislation on employment and social security. To create a measure that required as few assumptions about the content of directives and about the appropriate national law as possible, we used a broad classification. We classified each EU member as an early or late developer in three fields: occupational health and safety, gender equality, and employment.

_Hypothesis 2:_ Countries that introduced national legislation in the subject area of the directive early on are more likely to implement it quickly.

By limiting the directives examined to the subset concerning work and social policy, one can substantially improve a model of preferences. The directives in question add to labor protection and thus to the cost of production. Thus, one expects governments with low levels of protection to hesitate to adopt them, given that they will face a greater increase in labor costs. Similarly, one expects governments experiencing high levels of unemployment to be hesitant to introduce additional protection. Unemployment and labor cost data came from Eurostat (2001). Labor costs are mean hourly costs for workers in industry in ecus, whereas unemployment rates are harmonized annual averages.

_Hypothesis 3:_ Countries with low labor costs should implement social policy directives more slowly.

_Hypothesis 4:_ Countries with high unemployment rates should implement social policy directives more slowly.

Finally, the ideological position of the government should matter. Directives on social policy either call for standards higher than those already in place in many member states or set minima according to which states with more developed protections need (and ought) not lower their standards. Left-wing governments are expected to support additional worker protection and thus to be more favorable toward directive implementation in employment and social policy matters (Bradley, Huber, Moller, Nielsen, & Stephens, 2003; Esping-Andersen, 1990).

_Hypothesis 5:_ Left-wing governments should implement directives more quickly.

Similarly, governments with a more favorable attitude toward the European Union should make greater efforts to implement directives. Attitudes toward the EU can serve as an indirect proxy for reputation effects. As Downs and Jones (2002) explain the reputation thesis,
A major—if not the major—reason why states keep commitments, even those that produce a lower level of returns than expected, is because they fear that any evidence of unreliability will damage their current cooperative relationships and lead other states to reduce their willingness to enter into future agreements. (p. S69)

Governments that are supportive of the EU should be more concerned about the success of current, and the prospects for future, cooperative arrangements than governments hesitant about further EU integration. Thus, pro-EU governments should be more willing to implement directives punctually.

Dummy variables for left-wing and pro-Europe governments were used. This article follows the Bohrer and Tan (2000) classification of left-wing parties and the Ray (1999) classification of pro-Europe parties. Governments including at least one left-wing party were coded as left-wing (1), whereas governments with at least one party scoring at least a 6 on Ray’s 7-point scale were coded as pro-Europe (1).

Hypothesis 6: Pro-Europe governments should implement directives more quickly.

Hypotheses Concerning National Capacities to Implement EU Directives

In contrast to the enforcement model, which predicts that countries’ preferences on international agreements will determine compliance, the managerial model sees countries’ capacities to comply as the primary factor explaining compliance. How can we model national capacity to comply? First, one might expect countries whose decision-making processes require the consent of many actors to have greater difficulty in implementing legislation (Haverland, 2000; Tsebelis, 2002b). Both the number and ideology of coalition partners in particular governments and institutions, such as federalism, should raise the threshold of consensus needed for directive implementation and thus delay this process. Qualitative work supports both these hypotheses (Siedentopf & Ziller, 1988).

Data on multiparty governance come from Tsebelis (2002a). For practical purposes, these veto-point data correspond to the number of parties in governing coalitions, although they are adjusted slightly for the ideologies of the partners and for the existence of other veto players. Data on the representation of regional players come from Hooghe and Marks (2001) and range from 0 (no representation of regional interests at the central level) to 4.
**Hypothesis 7:** Countries whose governments are composed of many and ideologically diverse parties will implement directives slowly.  
**Hypothesis 8:** Countries with regional governments represented at the central level are expected to implement directives more slowly.

Furthermore, one might expect the ability of national bureaucracies to process directives to matter. The Chayes and Chayes (1995) theoretical analysis centers on capacity, and they present multiple empirical illustrations from a variety of treaty frameworks where scientific, technical, bureaucratic, and financial limitations of developing countries impeded compliance. Case studies on EU directive implementation also support this (Spanou, 1998). As Polidano’s (2000) survey of possible measures of public sector capacity indicates, no such measure exists. Thus, three related proxies were used here. First, following Mbaye (2001), ratings of national bureaucracy effectiveness, derived from Auer, Demmke, and Polet’s (1996) survey of civil services, were used.12 The second proxy for government effectiveness is a transparency score produced by the Internet Center for Corruption Research (2001)—the more transparent the government, the higher the score. The third measure is a score on government responsiveness. The International Institute for Management Development published these scores for the 1990 to 2000 period on the basis of national experts’ answers to the following question: “Does the government adapt its policies to new economic realities effectively?” (World Competitiveness Yearbook, n.d.). Responsive governments were given higher scores. Despite each of these measures’ limitations, their high correlation indicates that they capture a similar underlying dimension and are thus reasonably good proxies for a measure of government capacity.

**Hypothesis 9:** Countries with more efficient and more transparent bureaucracies will implement directives more quickly.

Finally, Martin (2000) and Bergman (2000) argue that countries whose parliaments scrutinize EU directives in the early stages of their development have better implementation records. Early parliamentary examination permits national governments to prepare themselves for the directive and also signals a credible commitment to implementation. Although both find empirical support for this proposition, they do so without controlling for many alternative explanations and invite further research on this matter. Parliamentary involvement in EU affairs was scored from 0 to 3; a 0 indicates that the country had no EU committee in a particular year, a 1 indicates
a committee with minimal involvement, and a 3 indicates substantial involvement. The scoring of parliamentary involvement follows Bergman (1997); dates of committee foundation come from Raunio and Wiberg (1999). Hypothesis 10: Countries that scrutinize directives prior to agreeing to them will implement them more quickly.

Data Analysis and Results

Measuring the Dependent Variables

Two variables are of theoretical interest: whether a country incorporated a directive into its national legal system on time and, if it did not, how long it delayed. Annual information on whether particular directives were incorporated in particular years, available in the European Commission’s *Annual Reports on the Monitoring of the Application of Community Law*, was used to construct these two variables. This project analyzes 53 directives from the field of employment and social policy whose implementation deadlines fell between 1985 and 2000.

Incorporating a directive into national law (transposition) is the first step toward complete compliance, in the sense of changing individual practices in every geography and sector the directive covers. Measuring full compliance in this latter sense is often of great interest, and yet detailed case studies of more than 50 directives in 12 countries would be prohibitively costly. Fortunately, however, Falkner, Treib, Hartlapp, and Leiber (2005) have conducted studies for 5 directives in the social policy field that move us closer to this objective. Although they do not measure on-the-ground implementation, they report the date when a country adopted the directive correctly rather than when it first notified the commission of its initial transposition, as initial transpositions are sometimes incorrect. Although these 5 directives were not chosen randomly, this information nonetheless permits order-of-magnitude estimates about how much more time correct implementation would take. Regressing the Falkner team’s measure of correct transposition on the delay in transposition measure used here yields the following estimates. These are expressed in months and are, as expected, highly significant (at the .99 level):

\[
\text{Correct Implementation} = 30 + 0.79 \times \text{Transposition Delay}
\]

Thus, a 1-year delay in transposition would translate in a little more than 3 years to essentially correct implementation, and so forth. In addition,
Falker et al. (2005) report what follow-up measures the commission initiated in the case of transposition delays and errors. For these five directives, the correlation between the delay in transposition measure used in the analysis below and the first stage of commission action—the letter of formal notice—was .54; the correlation between delay in transposition and the next stage in commission action—the reasoned opinion—was .52; and the correlation between delay in transposition and referral to the European Court of Justice was .46. As expected, the correlation is relatively high and positive, confirming the validity of the delay measure, and highlighting some systematic element in commission behavior. It is far from 1, however, reflecting limitations in the commission’s ability to process violations, and strategic choices by the commission to pursue certain violations over others.

Estimation Techniques and Model Specification

On-time implementation is a binary variable—either a country transposes a directive sometime before the deadline (1), or it fails to do so (0). To determine the effect of independent variables on this probability, the commonly used logit model was estimated using maximum likelihood techniques. This model involves transforming a binary dependent variable into the logarithm of the odds that 0 or 1 will be selected. Delay in implementation is a duration variable—the amount of time (in years) a country takes to implement a directive after the deadline has passed. From these data, a hazard rate can be calculated, indicating the probability that a directive will be implemented in a particular interval provided that it has not already been implemented.

To estimate the effects of independent variables on this hazard rate, a Cox (1972) proportional hazards model was chosen, as unlike parametric models, it does not require an assumption about the baseline hazard rate.\textsuperscript{17} In the Cox model, the hazard rate consists of an unspecified baseline hazard $\lambda_0(t|X)$ and a simple transformation of the independent variables that prevents them from making the hazard rate negative:

$$\lambda(t|X) = \lambda_0(t|X)\exp(x^\prime \beta)$$

Three highly correlated independent variables influenced dependent variables—bureaucratic efficiency, responsiveness, and transparency. Because they were all intended as proxies for the same concept and bureaucratic capacity, and were highly correlated, they could not all be placed in a model at the same time. We report models with the transparency measure as it varies over time and space, is available for the entire period under study, and is easiest to interpret.\textsuperscript{18}
Quantitative Results

Across countries, only 32% of the directives were implemented within the calendar year of the implementation deadline. Denmark and Britain were clear leaders—in these two countries only, directives are more likely than not to be implemented on time. Italy stands out as a laggard—there a directive had less than a 10% chance of on-time implementation.19

These country patterns are striking in light of some national preference theories. For example, the two leaders in directive implementation, Britain and Denmark, have national systems of voluntary agreements at odds with the regulatory model promoted through EU directives (Rhodes, 1996). Moreover, Britain has been the vocal opponent of EU social policy. Country patterns suggest that distance from existing political positions may not be the key predictor of timely implementation. This is explored further in the regression models below.

Table 1 presents a logit model predicting the on-time implementation of directives. Model 1 places all the variables in the regression, whereas Model 2 excludes the early legislation variable. Both the novel nature of this measure and its theoretical connection to labor costs warrant examining whether it substantially changes the size and significance of other coefficients. Because logit coefficients are difficult to interpret, and because the independent variables are coded on different scales, Table 2 indicates what happens when we move from the 20th percentile of an independent variable to the 80th percentile of that independent variable and hold all other independent variables constant at their means.20

Early legislation and labor costs were the two proxies for national preferences that were significant predictors of on-time implementation. As Table 2 shows, moving from the 20th to the 80th percentile on labor costs, a shift from 13 to 23 ecu per hour, increases a country’s chances of on-time implementation by 14%. Shifting from the 20th to the 80th percentile on the early legislation binary variable, that is, shifting from a country without early national legislation in the field of the directive to a country with such legislation, increases the probability of on-time directive implementation by 15%. Because these two variables are closely linked theoretically, and because, empirically, the coefficient on labor costs decreases in size and loses its significance when the national legislation variable enters the model, it makes sense to interpret these variables as different facets of closely related underlying processes. Several predictors of national preferences, including a government’s position on the left–right scale, its position on the pro-Europe scale, and its vote on the directive, never mattered. Instead, labor costs and early legislation have moderate to sizable effects. This pattern...
### Table 1
Predicting On-Time Implementation

<table>
<thead>
<tr>
<th>Preference measures</th>
<th>Exp. Sign</th>
<th>Coefficient (SE)</th>
<th>1&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Coefficient (SE)</th>
<th>2&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vote</td>
<td>+</td>
<td>0.82 (0.95)</td>
<td>0.53 (0.91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early legislation</td>
<td>+</td>
<td>0.65** (0.31)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor costs</td>
<td>+</td>
<td>0.04 (0.04)</td>
<td>0.06* (0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment</td>
<td>–</td>
<td>0.02 (0.03)</td>
<td>0.02 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left government</td>
<td>+</td>
<td>–0.14 (0.29)</td>
<td>–0.28 (0.27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro-Europe government</td>
<td>+</td>
<td>0.30 (0.33)</td>
<td>0.10 (0.32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capability variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiparty government</td>
<td>–</td>
<td>–0.30** (0.14)</td>
<td>–0.21* (0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regional representation</td>
<td>–</td>
<td>–0.30** (0.15)</td>
<td>–0.36** (0.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bureaucracy’s transparency</td>
<td>+</td>
<td>0.24** (0.10)</td>
<td>0.26** (0.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parliamentary involvement</td>
<td>+</td>
<td>0.17 (0.18)</td>
<td>0.12 (0.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>–</td>
<td>–3.75*** (1.27)</td>
<td>–3.51*** (1.23)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. $n = 347$.
b. $n = 360$.


### Table 2
Predicting On-Time Implementation: Moving From the 20th to the 80th Percentile of Each Independent Variable

<table>
<thead>
<tr>
<th>Exp. Sign</th>
<th>Change in On-Time Implementation When Independent Variables Move From the 20th to the 80th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference measures</td>
<td></td>
</tr>
<tr>
<td>Early legislation</td>
<td>+</td>
</tr>
<tr>
<td>Labor costs</td>
<td>+</td>
</tr>
<tr>
<td>Capability variables</td>
<td></td>
</tr>
<tr>
<td>Multiparty government</td>
<td>–</td>
</tr>
<tr>
<td>Regional representation</td>
<td>–</td>
</tr>
<tr>
<td>Bureaucracy’s transparency</td>
<td>+</td>
</tr>
<tr>
<td>Parliamentary involvement</td>
<td>+</td>
</tr>
</tbody>
</table>

* $p = .90$. ** $p = .95$.  


suggests that to the extent that preferences matter, it is long-standing features of the country and past government choices that predict compliance.

Variables measuring national capability to comply have more substantial effects. A shift from the 20th to the 80th percentile in government transparency (i.e., a shift from 5.0 to 8.5 on a 10-point scale) increases the probability of on-time implementation by 20%. Conversely, a shift from the 20th to the 80th percentile in coalition size, from 1 to 3 parties in government, decreases the probability of timely implementation by 13%. The shift from 0 to 2 points on the regional government score, another 20-80 shift, reduced the probability of on-time implementation by 15%. Parliamentary preparedness did not have a significant effect on on-time implementation.

In summary, measures of a country’s capabilities, including the capacity of its bureaucracy and the unitary or fragmented nature of its decision making, had a large impact on whether a country would meet directive deadlines.

If countries fail to meet a deadline, what predicts how long implementation will take? Table 3 presents a Cox model predicting the duration of the delay in directive implementation. The coefficients can be interpreted directly. A hazard ratio of 1 indicates no effect; a value between 0 and 1 indicates a decrease; and a value greater than 1 indicates an increase. Three measures of

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Predicting Delays in Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Preference measures</td>
<td></td>
</tr>
<tr>
<td>Vote</td>
<td>+</td>
</tr>
<tr>
<td>Early legislation</td>
<td>+</td>
</tr>
<tr>
<td>Labor costs</td>
<td>+</td>
</tr>
<tr>
<td>Unemployment</td>
<td>–</td>
</tr>
<tr>
<td>Left government</td>
<td>+</td>
</tr>
<tr>
<td>Pro-Europe government</td>
<td>+</td>
</tr>
<tr>
<td>Capability variables</td>
<td></td>
</tr>
<tr>
<td>Multiparty government</td>
<td>–</td>
</tr>
<tr>
<td>Regional representation</td>
<td>–</td>
</tr>
<tr>
<td>Bureaucracy’s transparency</td>
<td>+</td>
</tr>
<tr>
<td>Parliamentary involvement</td>
<td>+</td>
</tr>
</tbody>
</table>

a. n = 784.
b. n = 856.
preferences, early legislation, labor costs, and unemployment, had correctly signed but insignificant effects. Each of the capability measures had correctly signed and significant effects. When bureaucratic transparency increases by 1 unit (on a 1-10 scale), the rate of directive implementation increases by 12%. Conversely, when an additional regional player is added (on a 0-4 scale), the rate of directive implementation decreases by 12% (100%-88%) and by 9% for an additional coalition partner (on a 1-6 scale). When a country’s parliament gets involved early, the implementation rate improves by 17% for every 1-point increase (on a 0-3 scale).

To show the size of the combined effects of these variables, Figure 1 compares the implementation delays of countries with favorable and unfavorable preferences and the implementation rates of countries with high and low capabilities. In the figure on the left, the top line represents a country at the (bottom) 80th percentile of the preference variables, and the bottom line represents a country at the (top) 20th percentile of the preference variables. In the figure on the right, the top line represents a country
at the (bottom) 80th percentile of the capability variables, and the bottom line represents a country at the (top) 20th percentile of the capability variables. These graphs show that variation in capabilities has much bigger effects on implementation delays than does variation in preferences. For example, in the figure on the left, when the implementation deadline year is over (Year 1), countries with favorable preferences have yet to implement slightly fewer than 65% directives, whereas countries with unfavorable preferences have yet to implement almost 70% of directives. A year later (Year 2), countries with favorable preferences have around 35% of directives left to implement, whereas countries with unfavorable preferences have around 45% left, and so forth. The differences in the figure on the right are far more striking. When the implementation deadline year is over (Year 1), countries with low capabilities have yet to implement almost 80% of directives, whereas countries with high capabilities have fewer than 50% of directives left. A year later (Year 2), countries with low capabilities have yet to implement more than 50% of directives, whereas countries with high capabilities have fewer than 20% of directives left. The pattern continues for subsequent years, though duration model estimates are not as reliable for very long durations because these estimates are based on fewer data points.

**Qualitative Results**

Case studies from the secondary literature, a broad commission survey on directive implementation, and interviews with policy makers make up this study’s qualitative evidence. This qualitative research aims to validate the quantitative measures used and to explore additional compliance mechanisms rather than to test hypotheses. It thus focuses on Greece and Spain, two member states whose poor implementation record has provoked considerable concern in compliance debates. As these states are expected to have both substantial incentives to delay directive implementation as a matter of national social policy and substantial capability limitations, turning to them permits us to examine both accounts more closely.

The commission survey offered great detail on measures governments had put into place to improve their capabilities to comply but contained no questions on politically charged issues—for example, on whether interest groups that opposed the directives had tried to delay implementation. I therefore spoke to actors who had the knowledge and incentives to shed light on how both preferences and capabilities mattered and pointed the questions to direct their attention to puzzling findings. Commission officials and national bureaucrats are generally very informed about the directive implementation process. Unlike their counterparts in more political
roles, they have fewer incentives to hide information on lobby groups and other pressures that might frustrate their work. Labor union representatives should be even more willing to report on possible employer pressure to delay social policy directive implementation. I conducted seven interviews with officials in the Commission Employment and Social Affairs and Health and Consumer Safety Directorates General, seven interviews with officials in the Spanish Health and Employment Ministries, eight interviews with current and former officials in the Greek Health and Employment Ministries, and six interviews with social partner representatives (labor union leaders, labor union research institute directors, and national social and economic council members) in Greece and Spain.

This interview research supported the broad position that national capabilities are at least as important as preferences in explaining directive implementation, and probably more so. The consensus among commission officials was that no country showed bad faith in delaying directive implementation. In follow-up questions, some conceded to the possible existence of national lobby groups impeding directive implementation, but in general, they attributed the delay to national structural factors. Representatives from the national bureaucracies attributed delays almost exclusively to their own incapacity to process directives. Even in pointed follow-up questions, I very rarely elicited a response that hinted at the existence of pressures from above, or from lobby groups, to delay directive implementation. Similarly, interviewees in national bureaucracies cited no differences in pressures put on them by administrations that were not labor or EU friendly. Indicatively, part of a typical interview is included as an appendix.

Besides offering broad support for the importance of capabilities theories, qualitative data also indicate how preferences and capabilities matter. In interviews, commission officials cited various administrative difficulties as the main obstacles to implementation, whereas the commission report highlights good coordination among ministries, strong executives, and bureaucracies where the same civil servant is responsible for a directive from proposal through to implementation as the primary reasons for a good implementation record (European Commission, 2001, pp. 45-46). Both the commission report and an interviewee highlighted Denmark’s policy to negotiate directive texts in parliament early on as facilitating implementation. The report’s case-by-case analysis also highlights opposition from regional veto players as a persistent obstacle to implementation (European Commission, 2001, pp. 51, 69). This evidence suggests that a locus-of-authority contestation becomes increasingly relevant once a social policy issue moves from an established pattern of domestic decision making to other fora. Even more interesting than the activation of new players once a
new dimension of a question becomes salient is the apparent loss of interest from established actors. Most impressive among the interview data was testimony from labor union representatives. Neither in Spain nor Greece did they report much employer opposition to directive implementation. The assumption they shared was that any lobbying would have had to occur prior to the directive implementation phase; they saw the government claim that it was implementing an EU agreement as a trump to all national debate. This evidence supports the assertion that international organizations had succeeded in reorienting the axes of national debates from left–right redistribution into supranational–subnational actor authority.

To the extent that preferences matter, it is through constraints of past choices rather than because of new policy directions and future commitments of current governments. The importance of past choices is well established in purely domestic social policy accounts (Hacker, 1998; Pierson, 1996) and also figures prominently in qualitative work on directive implementation (Duina & Blithe, 1999; Knill, 2001). Spain and Greece are challenging cases for this hypothesis because of their history of dictatorships. Labor protection structures introduced in the Francoist era and a social protection system reinforced under the colonels’ dictatorship should be among the easiest to dismantle in the modernizing efforts of the “new” democracies. This is not what we observe. In both Greece and Spain, directives on working conditions and social dialogue, compatible with the pre-existing legal framework, were relatively easy to incorporate. Instead, the health and safety directives, requiring both substantial legal changes and even greater physical infrastructure upgrades, caused greater implementation delays (Threlfall, 1997; interviews in Greece and Spain).

Conclusions and Implications for Further Research

Why do countries sometimes fail to comply with international organization mandates? Some answers emerge from a study of delays in the implementation of EU directives. Delays in directive implementation are not the product of coherent national policy intended to achieve particular social outcomes. Although national preferences matter, state structures unrelated to national interest on social policy matter more. Multiparty governments, federalism, and limited bureaucratic capacity are the biggest obstacles to timely implementation. This conclusion discusses the implications of these findings for various research questions, focusing on how these results illuminate the impact of international organizations on welfare states.
International organizations reorient the axes of contestation. Employment and social issues fought along left–right lines when initiated at the national level instead end up contested as issues of authority of the national, supranational, and subnational levels.

Only some measures of government preferences predict implementation delays. The clear expression of a government’s short-term preferences on a particular issue, its vote on a directive, does not matter. Nor does the government’s position on the left–right scale or its attitude toward the EU seem to be of import. However, the legacies of past policy choices—proxied by early activism in the social policy field and high current labor costs—constrain government preferences and shape compliance patterns to some extent.

This article finds support for the managerial model’s emphasis on technical and bureaucratic capabilities. More competent bureaucracies are better able to process directives and reduce implementation delays substantially. Multiparty coalitions governing at the national level and powerful subnational actors are equally important predictors of delay. Parliamentary scrutiny prior to the passage of the directive also contributes to shorter implementation delays. The sizable capability effects observed are consistent with both institutionalist and constructivist theories. For example, poor bureaucracies may provide few opportunities for frequent contact between their officials and their counterparts in other EU member states, thus impeding learning. Similarly, EU officials may have greater ease in communicating with countries whose leadership is unified than with countries in which multiple parties must be brought to the table.

Although many compliance theories suggest that international organizations shape policy by permitting future constraints to enter current decision making through reputation and commitments, these data suggest that this only occurs when specific mechanisms to carry out these commitments are in place. Although pro-EU countries, which have most to lose in terms of reputation, are not better at complying, countries whose parliaments prepare for directive implementation even before the directive is agreed on are somewhat better able to reduce delays. Theories on reputation assume an identity of interests and incentives between parties making an agreement and parties implementing it. However, only the former are directly motivated by reputational concerns, and as this study shows, veto players have substantial capacity to induce delays. The greater the disjuncture between treaty signer and treaty implementer, the less reputation matters. Only when a country can overcome this disjuncture—by involving, for example, both the executive and the legislature in the process early on—are its commitments credible.
So how can international organizations bring about change? An extensive literature on social policy initiated at the national level indicates that battles fought principally between labor unions and left-wing parties, on one hand, and employers and more conservative parties, on the other, can lead to change. The above findings suggest, however, that the axes of contestation shift where international organizations take on policy making. Here, national debates between the Left and Right matter little, and instead it is center–region, executive–legislative, and executive–bureaucracy tensions that gain prominence. The large agenda-setting and policy-making literatures explain that major shifts in policy occur when issues are redefined and previously uninterested actors join the debate (Baumgartner & Jones, 1993). International organizations can trigger reorientations in debate participants and thus shape national policy systems. It is difficult to predict how exactly a structural shift translates into policy changes, but one might hazard that this reorientation could result in social policy that is less responsive to organized constituent concerns, notably those of labor unions.

A study of areas where international organizations face resistance from national governments shows international organizations to be nonetheless quite influential in shaping national decisions. First, the simplest model of international organization irrelevance, according to which governments comply or fail to do so on the basis of predefined, national self-interest, is proven false. Moreover, international organizations are shown to work as powerful motors for change as they reorient debate axes at the national level. Although the European Union is a unique community, the obstacles it faces are likely sharper in other international systems. And international and regional labor law coordination is growing in importance, as both trade liberalization efforts and human rights activism begin to focus on labor concerns (Macklem, 2002). Perhaps international organizations interested in improving compliance patterns would be better off de-emphasizing the conclusion of expansive agreements with national executives and focusing more on including the concerns of domestic players who might have the power to delay or stop implementation. In this effort, building relationships with traditional partisan groups may be less fruitful than creating links to different layers of the national governance machinery.

Appendix

Interview Extract

Extract from interview with the general director for health and safety of the Greek Employment Ministry from the late 1980s to the late 1990s (own translation).
Interviewer: Why is it that Greece delayed in implementing social policy directives?
General director: Following the Single European Act, we were faced with a flurry of directives. Although we could transpose directives with presidential decrees, even with ministerial orders, it was just very difficult to do this in a timely fashion. We didn’t have the experience, the manpower.

Interviewer: Were you pressured by industry, labor, or other parties to modify a law in their favor?
General director: Yes, all our policies go through social consultation. But this was consultation; we didn’t have to change our policies because of it.

Interviewer: Did you ever find yourself changing the content of a decree in response to such pressure?
General director: Yes, once the Council of State forced us to change the content of a decree. We thought we had stricken a compromise with the labor view on the question of time limits for the work hours of employment experts. In the end the Council of State required us to limit the hours these could provide services to the 8-hour day.

Interviewer: What about directive 89/391? Why did Greece delay so much in implementing it?
General director: You know that the 1990 to 1993 period was a difficult one for Greek politics. It was just very difficult to get any work done then.

Interviewer: A right-wing government was in power for much of this period. Did they seem any less willing to promote social policy legislation?
General director: No, that didn’t matter. It’s just the government was changing all the time; the political leaders were changing. It’s very difficult to do policy when you’re shut in your office, doing it on your own.

Notes

1. The 1993 North American Agreement on Labor Cooperation, NAFTA’s labor side agreement, sets up a commission to enforce each country member’s national laws, with sanctions possible for violations of health and safety, minimum wage, and child labor laws. In 1998, MERCOSUR governments issued the social labor declaration, which they have been following up with other capacity-promoting efforts in the employment area. The 1976 Declaration of the ASEAN concord calls for cooperation among member states on labor affairs; efforts here, however, have been limited to information sharing, coordination of active labor market policy, and technical assistance (Gitterman, 2003).

2. As a reviewer highlighted, the managerial model also discusses how actor preferences can be transformed through international interactions (Zürn & Joerges, 2005).

3. Not including controls for alternative theories can lead to biases. For example, Mbaye (2001) and Lampinen and Uusikyla (1998) both report the paradoxical finding that pro-Europe public opinion is positively correlated with a weaker implementation record. This could result from omitted variable bias: Less developed EU member states may be both more supportive of the EU because they draw greater advantages from it and have bigger difficulties in and disincentives for implementing EU policy promptly.

4. More specifically, Mbaye (2001) uses European Court of Justice (ECJ) cases initiated and won by the commission, whereas Martin (2000), Bergman (2000), and Lampinen and...
Uusikila (1998) use percentages of directives implemented on time each year. However, directives not implemented by the deadline do not always end up in the ECJ, and the commission chooses this sample strategically. Using percentages of directives implemented on time makes key types of analysis impossible, including analyses incorporating directive-level information and analyses of delays in implementation. Additional excellent studies that report aggregate implementation statistics and draw important insights, without, however, connecting dependent and independent variables through multivariate regression, include Börzel (2001), Tallberg (2002), and Falkner, Treib, Hartlapp, & Leiber (2005).

5. The European Commission typically initiates directive proposals, which the council then approves. Qualified majority voting was introduced for health and safety issues in 1986 and was expanded in 1991 to cover directives on working conditions, information and consultation of workers, equal opportunities, and the integration of people excluded from the labor market.

6. National governments must inform the International Labour Organization (ILO) about legislation on labor and social security matters. Labor law experts conclude that at least for the 1980s and 1990s, the Natlex database offers complete and cross-nationally comparable classifications (Scholz & Trantas, 1995).

7. More detailed coding would have sacrificed reliability. Within the NATLEX database, early introduction of measures in a particular subfield was correlated with ample law production in that field. Eichener (1997) independently classifies the same countries identified here as early and late developers in health and safety.

8. If countries had legislation on occupational safety hazards related to toxic substances and agents prior to 1980, they were considered early developers in matters of occupational health and safety; accordingly, all countries except for Greece, Spain, Portugal, and Luxembourg were early developers. If countries had legislation on equal opportunity prior to 1990s or maternity protection legislation prior to the 1980s, they were considered early developers in matters of women and employment. Accordingly, all countries except for Greece, Spain and Ireland were early developers. Employment measures fell into four ILO categories: conditions of employment, conditions of work, employment, and industrial relations. Countries that were late developers in two or more of these categories were coded as late developers for employment measures; these were Greece, Ireland, Italy, Spain, Luxembourg, Denmark, and Portugal.


10. Using data provided by George Tsebelis (2002a), we also tested for three commonly used alternative codings of left-wing party orientations, but the results did not change.

11. Birchfield and Crepaz (1998) and Lijphart (1999) argue that although competitive veto points, such as regional players, impede the policy-making process, collective veto players, such as multiparty governments, facilitate cooperative policy making and thus speed up the process. Although this theory implies that Hypothesis 7 should be signed in the opposite direction, it underscores the need for two separate hypotheses on veto players.

12. However, this indicator is a direct measure of civil service hiring methods, which captures their effectiveness only indirectly and does not vary over time.

13. Although committee strength may vary over time, Bergman’s (2000) more recent classification shows that at least between 1996 and 1999, these shifts were relatively minor.

14. They also present some data for a sixth directive, but as its transposition deadline was in 2000, neither their data nor these data fully track post-transposition measures.

15. The Falkner et al. (2005) report “essentially correct” transposition, a measure that reflects
essentially successful fulfilment of most requirements (this contains a qualitative consideration of how good adaptation was and whether it captured the essential parts of the Directive) and, simultaneously, of the most central requirements of any Directive (this contains two quantitative considerations; how many requirements are fulfilled and how many of these are crucial in terms of the Directive’s aim). (p. 269)

16. In selecting cases, Falkner et al. (2005) focused on “where the most controversial debates took place during the 1980s and 1990s” and picked directives that “actually alter pre-existing national rules,” rather than directives that extend the geographic scope or amend previous directives or create supranational obligations without altering preexisting domestic laws (pp. 7-9). Unsurprisingly, these five directives show somewhat longer implementation delays than do the other directives in the present data set.

17. Maximum likelihood methods were also used to estimate these models. The exact marginal method was used to resolve ties. Data sample size was not an important concern here, as more than 50 directives were implemented in 12 countries. Linear interpolation (imputing missing values from previous and subsequent ones) was used where possible; otherwise, observations with missing values were not included.

18. Results remain very similar when transparency is replaced by responsiveness or efficiency or by the factor of these three measures. We tried three alternative specifications of the Cox (1972) model to test whether the results were driven by cross-national differences not captured by the independent variables presented. We ran a fixed effects model, a model with a dummy for Italy (following Mbaye, 2001), and a model including GDP per capita as a proxy for other cross-national variation, and results did not change substantially. In addition, the results did not change substantially when the insignificant variables were excluded.

19. These three countries had delay patterns significantly different from the others at the .05 level, as shown by their Kaplan-Meier survivor functions, functions that indicate the probability that a directive would not be implemented 1 year, 2 years, 3 years, and so on following the deadline. Comparisons by year indicate slight variation from year to year, but no overall time trend.

20. Only variables that appear significant in at least one specification are included in this table. All magnitude estimates come from Model 2, except for early legislation, which was excluded from Model 2 and whose coefficient estimate thus comes from Model 1.

21. Estimates are based on Model 1 and include all variables with the correct sign.

22. Peter Lange (1993) shows that south European countries are expected to be consistent losers in the social harmonization process. He proposes they accepted these regulations so as to receive other side payments. Thus, they should have the greatest incentives to use any flexibility in the transposition process instrumentally to delay implementation. For a discussion of the capacity limitations thought to impede directive implementation in the south, see Börzel (2000).


24. For most of the period under study, there was no coordinated and effective union and employer lobbying at the EU level. See Hyman (1999) for a review.

25. Findings that anti-EU and nonlabor governments are just as content to implement social policy directives might also surprise constructivist theorists who expect actors who have unfavorable prior beliefs or who do not see norm sponsors as legitimate to comply less. See Axelrod (1997, pp. 55-59), Johnston (2001, p. 497), Cortell and Davis (2000, p. 74), and Goodman and Jinks (2004).

26. This literature does not focus on how interested parties stop participating, perhaps because it happens only rarely, but even this seems to have occurred here.
27. This seems especially likely if EU activity supplants rather than supplements national policy-making activity. Interviewees also indicated that in areas of high EU activity, such as health and safety at work, their strained bureaucracies were no longer investing many resources, except as necessary to translate and transpose EU directives.

28. 89/391 is a framework directive “on the introduction of measures to encourage improvements in the safety and health of workers at work.”

29. From 1982 to 2004, Greece was ruled by a socialist government almost continuously, but a right-wing government held office between 1990 and 1993.

References


liaments and the legitimacy of EU governance*. ECPR Joint Sessions of Workshops, Mannheim, Germany.

Ray, L. (1999). Measuring party orientations towards European Integration: Results from an

P. Pierson & S. Leibfried (Eds.), *European social policy: Between fragmentation and inte-

security and labour market. In H. Doering (Ed.), *Parliaments and majority rule in Western
Europe* (pp. 628-653). New York: St. Martin’s.


Snyder, F. (1993). The effectiveness of European community law. Institutions, processes, tools

the Greek case. *Journal of European Public Policy, 5*, 467-484.


Threlfall, M. (1997). Spain in Europe: A laggard or compliant member state? *South European
Society and Politics, 2*(2), 1-33.

tsebelis/

University Press.

*World Competitiveness Yearbook*. (n.d.). Lausanne, Switzerland: International Institute for
Management Development.

Young, O. R. (1992). The effectiveness of international institutions: Hard cases and critical
variables. In J. N. Rosenau & E.-O. Czempiel (Eds.), *Governance without government*

beyond the nation-state*. Cambridge, UK: Cambridge University Press.

**Katerina Linos** is a JD and PhD candidate in government and law at Harvard. Her disserta-
tion examines the diffusion of social policies across advanced industrialized countries. Her
recent publications include “Are Socialists a Woman’s Best Friend?” in the *Journal of
European Public Policy* and “Self-interest, Social Beliefs and Attitudes to Redistribution” in
the *European Sociological Review* (with M. West).