Improving the Efficiency of Pretrial Bargaining in Disputes over Noncompliance with International Law: Encouraging Evidence from the European Union

Sivaram Cheruvu[1] and Joshua C. Fjelstul[2]

[1]The University of Texas at Dallas

[2]The University of Geneva

Abstract

International institutions can create pretrial bargaining procedures to resolve disputes over noncompliance with international law. In this article, we consider how international institutions can improve the efficiency of pretrial bargaining, thereby reducing the compliance deficit. We use a quasi-experimental research design and difference-in-differences (DD) estimators to analyze the effectiveness of a recent policy initiative in the European Union (EU) — called EU Pilot — that was designed to improve the efficiency of pretrial bargaining in the EU. By evaluating the effectiveness of EU Pilot, we provide new evidence of a proposed management school solution to noncompliance — proper rule interpretation. By preventing accidental noncompliance, EU Pilot allows the Commission and member states to more quickly process cases involving intentional noncompliance, improving the efficiency of pretrial bargaining in the EU.

Key words: European Commission, European Union law, infringement procedure, management school, noncompliance, policy evaluation

International institutions can create pretrial bargaining procedures to resolve disputes over noncompliance with international law (e.g., Gilligan, Johns and Rosendorff 2010; König and Mäder 2014). Pretrial bargaining allows international institutions and states to settle cases involving accidental noncompliance before going to court (Fjelstul and Carrubba 2018), which frees up docket space for cases involving intentional noncompliance, where international courts play a critical role in resolving disputes and enforcing compliance (e.g., Carrubba, Gabel and Hankla 2008; Carrubba and Gabel 2015). However, pretrial bargaining can be time-consuming (e.g., Kim 2008; König and Mäder 2014), as preparing effective counter-offers is a resource-intensive task that requires the parties to have sufficient administrative capacity. Delays in pretrial bargaining can lead to delays in the resolution of disputes over noncompliance, harming citizens and business who benefit from the proper application of international law.

In this article, we consider how international institutions can improve the efficiency of pretrial bargaining, thereby improving compliance with international law. We use a quasi-experimental research design to analyze the effectiveness of a recent policy initiative in the European Union (EU) — called EU Pilot — that was designed to improve the efficiency of pretrial bargaining in the EU. The EU has a centralized monitoring system with a formal, multi-step pretrial bargaining procedure — called the infringement procedure — in which member states bargain with the European Commission (the EU's executive body) over settlements in disputes over noncompliance with EU law before going to court. In practice, the infringement procedure can be a long, inefficient process, requiring back-and-forth exchanges between the Commission and government agencies in EU member states. Between 2004 and 2010 (the period we consider in our empirical analysis), the average case spent 305 days in pretrial bargaining. The Commission regularly points out that the inefficiency of this procedure undermines the application of EU law.[1]

Implemented in 2008, EU Pilot created an online platform to facilitate communication between the Commission and EU member states during the policy implementation process (as most disputes over noncompliance in the EU relate to the implementation of EU law

by EU member states) about what constitutes a violation of EU law. The objective of this program was to increase the efficiency of pretrial bargaining (i.e., reduce the duration of the pre-litigation phase of the infringement procedure) by preventing accidental noncompliance by member states. According to the Commission, the causal mechanism is that preventing accidental noncompliance would reduce the number of new infringement cases, freeing up resources and allowing the Commission and member states to allocate those resources across the remaining cases (cases primarily involving intentional noncompliance). This would allow the Commission and member states to resolve the remaining disputes over noncompliance more quickly, improving the efficiency of pretrial bargaining.

By evaluating the effectiveness of EU Pilot at improving the efficiency of the infringement procedure, we provide new empirical evidence about a proposed management school solution to noncompliance — proper rule interpretation. The management school argues that noncompliance with international law is often accidental — a product of governments misinterpreting rules or lacking administrative capacity (e.g., Mbaye 2001; Börzel 2003; Hille and Knill 2006; Sedelmeier 2008). EU Pilot promotes proper rule interpretation by improving communication between the Commission and member states about what constitutes a violation of EU law.

Leveraging the fact that only some member states participated in EU Pilot at first, we use a quasi-experimental design and observational data on infringement cases to estimate the effect of the program on the efficiency of pretrial bargaining during the infringement procedure. We use a difference-in-differences (DD) estimator to account for the fact that participation in EU Pilot was voluntary (i.e., the treatment was not randomly assigned). We find that EU Pilot substantially improves the efficiency of the pretrial bargaining, speeding up the initial phase of bargaining by 74 days, on average, among participants relative to non-participants. We then use a difference-in-difference-in-differences (DDD) estimator to check for possible violations of the parallel trends assumption in our DD analysis. We also find empirical evidence consistent with the observable implications of the Commission's proposed causal mechanism.

Our findings have a clear policy implication: domestic and international institutions that use pretrial bargaining to resolve disputes over policy implementation may want to consider investing in policy initiatives aimed at facilitating better communication with implementing actors early in the policy implementation and dispute-settlement processes.

Substantive Context: The EU Pilot Program

EU Pilot complements the EU's formal infringement procedure. In this section, we discuss the infringement procedure, the motivation behind EU Pilot, the mechanism by which the Commission expects EU Pilot to improve the efficiency of pretrial bargaining during infringement procedure, and how EU Pilot was implemented, which is critical for our research design and empirical analysis.

The Commission uses the infringement procedure to prosecute noncompliance with EU law by member states. If the Commission believes a member state is not complying with EU law, it can launch an infringement case by sending a letter of formal notice (LFN), asking the member state to respond to the alleged noncompliance (Article 258 TFEU). If the Commission is not satisfied with the member state's response, it can send a reasoned opinion (RO), explaining in greater detail why the member state is not in compliance. If the Commission is still not satisfied with the member state's response, it can refer the case to the Court of Justice of the European Union (CJEU) — thus ending the pre-litigation phase and starting the litigation phase. The Court will then issue a ruling.

The Commission can launch two types of infringement cases: non-communication cases and non-conformity cases. When the EU enacts a directive (the main type of law in the EU), member states have to implement it by enacting national laws or regulations that achieve the policy objectives set out in the directive (a process called transposition). Directives specify a deadline by which all member states must transpose them. Member states are required by law to notify the Commission when they have done so. If a member state misses a transposition deadline, the Commission will open a non-communication case. These cases

are opened automatically as soon as a member state misses the deadline. If a member state does transpose a directive by the deadline, but does so incorrectly, the Commission can open a non-conformity case.

One problem with the infringement procedure, from the Commission's perspective, is that it can be extremely inefficient, taking up valuable administrative resources. Between 2004 and 2010, pretrial bargaining (the LFN and RO stages) took 305 days, on average. Only 5 percent of cases go to court, so pretrial bargaining accounts for nearly all activity related to the infringement procedure. In 2007, the year before EU Pilot was implemented, the Commission had a backlog of over 2,000 cases because of this inefficiency (see the supplementary material for details).

The Commission's proposed solution to this backlog is EU Pilot, which the Commission started to implement in 2008. The stated objective of EU Pilot was to improve the efficiency of the infringement procedure. As the Commission explains in an internal memo, "the need for recourse to infringement proceedings should be reduced by the problem solving and preventive measures" introduced by EU Pilot, and this "will lead to the more efficient management and resolution of infringement cases."[2]

The centerpiece of EU Pilot is an online platform that facilitates communication between the staff of the Directorates-General of the Commission and the staff of government agencies in EU member states during the policy implementation process — before the infringement procedure ever comes into play. This platform makes it easier for member states to ask questions about how to correctly implement EU law and for the Commission to make sure that member states are staying on track when it suspects an infringement.

The Commission clearly identified the causal mechanism by which it expected the program to improve the efficiency of the infringement procedure. The Commission explained that it would use the online platform to help member states better understand what constituted compliance, which would help member states avoid accidental noncompliance. This would allow member state and Commission lawyers to allocate their time across fewer nonconformity cases (see the supplementary material for further discussion).

One important aspect of EU Pilot for our analysis is that the Commission implemented the program in stages. This provides an opportunity to use a difference-in-differences (DD) design to estimate the effect of EU Pilot on the efficiency of the infringement procedure. EU Pilot launched on April 15, 2008 with 15 member states volunteering.[3] In March 2010, the Commission invited the remaining member states to join and the program expanded to include all member states by 2013 (see the supplementary materials for details). The fact that member states self-selected into the program is the primary inferential challenge we address in our analysis — a challenge that our DD design directly addresses.

Another important aspect of EU Pilot is that it should only decrease the duration of non-conformity cases. We leverage this fact in our empirical analysis by using a difference-in-difference-in-differences (DDD) design to address possible violations of the parallel trends assumption in our DD analysis. EU Pilot is not used for instances of noncompliance that relate to the non-transposition of directives (which go on to become non-communication cases) and the reduction of the Commission's workload due to the reduction in the number of non-conformity cases involving accidental noncompliance should not affect how it takes to resolve non-communication cases.

As the Commission explains, "non-communication cases are registered directly and only in the infringement database" because they are "clear-cut cases for which the use of EU Pilot would not provide added value."[4] In non-communication cases, the duration of the LFN and RO stages is primarily driven by how quickly legislative actors transpose directives. In non-conformity cases, the Commission has to explain why the member state has violated EU law, and the member state has to respond to the Commission's legal arguments, which takes longer and is more resource-intensive, in terms of administrative capacity.[5]

Management School Explanations for Noncompliance

There is an extensive debate in the literature on the EU about why EU member states violate EU law, with different schools of thought pointing to different ways of reforming

the infringement procedure (e.g., Chayes and Chayes 1993; Downs, Rocke and Barsoom 1996; Carrubba 2005; Tallberg 2002). Management school theorists argue that noncompliance is primarily accidental (e.g., Haverland 2000; Mbaye 2001; Börzel 2003; Hille and Knill 2006; Jensen 2007; Sedelmeier 2008, while enforcement school theorists argue that noncompliance is primarily intentional (e.g., Börzel 2000; Falkner et al. 2005; Luetgert and Dannwolf 2009; Thomson 2007; Zhelyazkova 2012; König and Mäder 2013). Constructivists, meanwhile, argue that noncompliance is a function of normative beliefs over rule obedience (e.g., Finnemore and Sikkink 1998; Hurd 1999; Checkel 2001). While these schools are not mutually exclusive (e.g., Tallberg 2002; Börzel et al. 2010; Carrubba and Gabel 2017), we evaluate a management school solution.

Management school explanations pinpoint three factors causing noncompliance with EU law: domestic policy coordination complications and veto points impeding transposition of EU directives (e.g., Chayes and Chayes 1993; Dimitrova and Steunenberg 2000; Haverland 2000; Mbaye 2001; Steunenberg 2006), misfits between EU directives and existing policies causing implementation trouble for domestic administrative institutions (e.g., Börzel 2000; Knill and Lenschow 1998; Knill 1998), and limited administrative capacity generally (e.g., Bursens 2002). In addition to diagnosing these factors causing noncompliance, management scholars proscribe capacity building and proper rule interpretation as solutions to noncompliance (e.g. Chayes and Chayes 1993; Chayes and Chayes. 1995; Keohane and Levy 1996; Tallberg 2002).

Consistent with the management school's proscriptions, the Commission has designed the EU Pilot program to help itself and member states "communicate and clarify the factual and legal background of problems arising in relation to the conformity of national law with EU law or the correct application of EU law."[6] In other words, EU Pilot attempts to solve implementation problems outside the formal infringement procedure by improving rule interpretation. The Commission explains that the program is "a first step to try to resolve problems, so that, if possible, formal infringement proceedings are avoided."[7] By facilitating proper rule interpretation, we expect EU Pilot to prevent accidental noncompliance that

would otherwise lead to non-conformity cases, allowing the Commission and member states to process the remaining cases more quickly.

Many studies have suspected that the Commission is strategic about which cases it prosecutes without rigorously theorizing this selection process or offering quantitative evidence (e.g., Mbaye 2001; Börzel 2003; Thomson 2007; Hartlapp and Falkner 2009; Steunenberg and Rhinard 2010). One mechanism guiding case selection is that the Commission chooses to pursue cases that it expects to win in court (König and Mäder 2014; Fjelstul and Carrubba 2018), but what we propose is a complementary mechanism. We argue the Commission is strategically selecting out of cases involving accidental noncompliance by resolving them through EU Pilot instead of using the infringement procedure.

We focus our theoretical and empirical analysis on the duration of the LFN stage — as opposed to the duration of the RO phase, the duration of the entire pre-litigation phase (the LFN and RO stages together), or the duration of the entire case — for several reasons. First, the LFN stage is the only stage that all cases go through, which makes it the only point of direct comparison across cases. Second, the duration of the LFN stage is substantively more important than the duration of the RO stage. Such a small percentage of cases (only 20 percent in our sample) move on to the RO stage that the efficiency of the RO stage, in terms of its impact on the Commission's overall backlog, is nowhere near as important as the efficiency of the LFN stage.

Third, the Commission decides whether to advance a case from the LFN stage to the RO stage based on whether the member state's response to its LFN is satisfactory. This strategic selection process is very complex (see Fjelstul and Carrubba 2018) and looking at the duration of the entire pre-litigation phase would confound our ability to make inferences. Fourth, once the case is referred to the Court, the duration of the litigation phase no longer depends on actions by the Commission or the member state (the Court controls the schedule). EU Pilot should increase the efficiency of the infringement procedure is by reducing the workload of the Commission and the member state, so the program should have no impact on the duration of litigation phase.

An important aspect of the infringement procedure for our analysis is that the duration of the LFN and RO stages depends on the actions of both the Commission and the member state. Once the Commission sends an LFN or RO, the member state has to respond, and then the Commission has to decide whether to move the process forward to the next stage or close the case. By freeing up Commission resources, EU Pilot should actually allow the Commission to processes non-conformity cases more quickly not just for participants, but also for non-participants.

In this way, EU Pilot could increase the efficiency of the infringement procedure for non-conformity cases across the board — for all member states. (Remember that the duration of non-communication cases is primarily determined by legislative actors.) However, it is important to note that EU Pilot will also free up resources for participants, not just the Commission. Thus, there should be a differential effect between participants and non-participants that operates through the member state, not the Commission. This is the effect we are interested in estimating in our empirical analysis — the effect of EU Pilot on the responsiveness of participating member states during the LFN stage. We will discuss this issue further in the context of our empirical analysis.

Hypothesis 1. The average duration of the LFN stage will be shorter for member states participating in EU Pilot than for member states not participating in EU Pilot.

Next, we derive testable, observable implications of the causal mechanism. If the Commission's proposed causal mechanism is at work, and EU Pilot largely screens out easy non-conformity cases (those involving accidental noncompliance), the implication is that the remaining instances of noncompliance that do make it through EU Pilot and become non-conformity cases are mostly difficult cases (those involving intentional noncompliance). Recent work on strategic selection during the infringement procedure points to what we should observe if the composition of incoming cases changes in this way.

Fjelstul and Carrubba (2018) find that member states that want to avoid going to court will settle cases involving intentional noncompliance at the RO stage. If a member state has

accidentally committed a violation of EU law, it has an incentive to come into compliance as soon as it receives an LFN from the Commission pointing out the problem. If the member state has intentionally committed a violation, on the other hand, but does not want to go to court (a costly and risky proposition), it has an incentive to wait to come into compliance until the RO stage, as there is positive probability that the Commission will drop the case during the LFN stage, either because it does not think the case is worth pursuing, given administrative constraints, or because it does not expect that it will be able to win in court. The testable, observable implication of this causal mechanism, then, is that the cases that do make it to the infringement procedure under EU Pilot are less likely to resolve at the LFN stage and more likely to resolve at the RO stage.

Hypothesis 2. The probability that an infringement case resolves at the LFN stage will be lower for member states participating in EU Pilot than for member states not participating in EU Pilot. The probability that a case resolves at the RO stage will be higher.

We do not expect EU Pilot to have any effect on the likelihood that disputes over noncompliance are resolved during the pre-litigation phase (i.e., before the Commission refers a case to the Court) versus the litigation phase. Building on the insight that member states that commit intentional violations do not have an incentive to come into compliance until the RO stage, we do not expect EU Pilot to change member states' strategic behavior at the RO stage. By preventing accidental noncompliance, which is usually resolved during the LFN stage, EU Pilot just helps cases reach the RO stage more quickly.

Empirical Analysis

We assess the impact of EU Pilot on the efficiency of the Commission infringement procedure using a quasi-experimental design.[8] First, we use a difference-in-differences (DD) estimator to estimate the treatment effect of EU Pilot on the duration of the LFN stage. Next, we use a difference-in-difference-in-differences (DDD) estimator to check for potential violations of the DD estimator's key assumption (the parallel trends assumption). Finally, we present

empirical evidence that is consistent with the observable implications of the Commission's proposed causal mechanism, as described above.

Our unit of analysis is an infringement case, and the treatment in our quasi-experiment is participation in EU Pilot. Since member states self-select into the program, the treatment is not randomly assigned as in a randomly controlled trial. This is the primary challenge to inference in our analysis, and as we discuss in detail below, our DD design directly addresses this non-random assignment.

The treatment group consists of cases involving participants (PILOT = 1) and the control group consists of cases involving non-participants (PILOT = 0). The pre-treatment period runs from May 1, 2004 (the date of the enlargement into central and eastern Europe) to April 15, 2008 (the date EU Pilot started). We code observations in which the Commission sent a LFN during this period as pre-treatment observations (PERIOD = 0). We start the sample after the 2004 enlargement so the composition of the sample, in terms of member states, remains stable. The post-treatment period runs from April 16, 2008 to March 1, 2010 (the date that the second wave of member states joined EU Pilot). Similarly, we code all observations in which the Commission sent an LFN during this period as post-treatment observations (PERIOD = 1). There are 7,054 observations in the sample — 1,777 non-conformity cases and 5,275 non-communication cases. We only include complete cases — cases where we observe the opening of the case, the closing of the case, and all stages in between.[9] There are 3,199 control observations (PILOT = 0) and 3,853 treated observations (PILOT = 1). And there are 6,069 pre-treatment observations (PERIOD = 0) and 983 post-treatment observations (PERIOD = 1).

The outcome variable (DURATION) is the number of days from the date the Commission sends an LFN to the date the Commission concludes the LFN stage of the infringement procedure by sending an RO (if the Commission chooses to pursue the case further) or by closing the case (if it chooses not to pursue it further).

Difference-in-Differences Analysis

We use a DD design to estimate the effect of EU Pilot on the duration of the LFN stage. In this analysis, we subset our sample and only look at non-conformity cases. In our subsequent DDD analysis, we will expand our sample to include all cases. This approach estimates the average treatment effect on the treated (ATT) — the average effect of participating in EU Pilot on the duration of the LFN stage in non-conformity cases involving member states who participated in the program.

Earlier we mention that EU Pilot could actually increase the efficiency of the LFN stage for non-participants by reducing the Commission's overall workload, but that the effect should be greater for participants because it also reduces the member state's workload, and the duration of the LFN stage depends on both actors' workloads because of the back-and-forth communication during the LFN stage. Our analysis will estimate the differential effect for participants, which captures the extent to which the program improves member states' response times, on average, to Commission inquires during the LFN stage.[10]

Our DD design addresses the primary threat to inferences in our analysis — the fact that member states self-selected into the program — by controlling for any difference between the pre-treatment and post-treatment periods and for any differences between participants and non-participants (Angrist and Pischke 2008; Imbens and Wooldridge 2009; Greene 2012). To precisely define the threat to inference, the fact that member states volunteered to participate in EU Pilot means that treatment assignment is non-random. Consequently, there could be systematic differences between cases involving participants, in the treatment group (PILOT = 1), and those involving non-participants, in the control group (PILOT = 0), in terms of the average duration of the LFN stage (DURATION).

If the LFN stage tends to be shorter for participants, and we (incorrectly) treat the treatment as randomly assigned — by comparing the average duration of the LFN stage for participants (PILOT = 1) to that of non-participants (PILOT = 0) during the post-treatment period (PERIOD = 1) — the program could appear to be more effective that it really is.

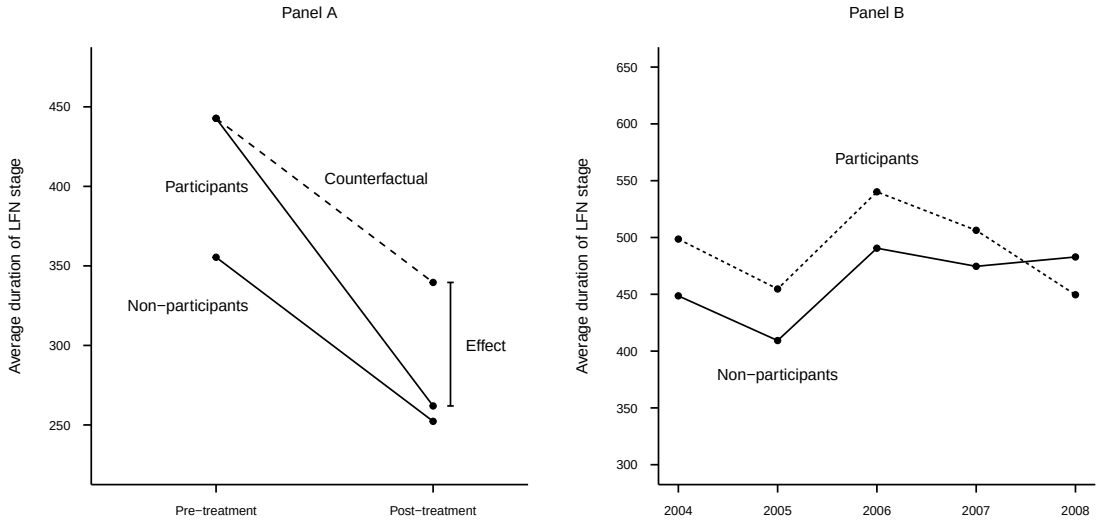

Figure 1. Panel A visualizes the parallel trends assumption. Panel B shows parallel trends in the outcome variable for participants and non-participants between 2004 and 2008.

Conversely, if the LFN stage tends to be longer for participants, the program could appear to be less effective than it really is.

Empirically, the average duration of the LFN stage in cases involving participants (PILOT = 1) is 87 days longer than in cases involving non-participants (PILOT = 0) in the pre-treatment period (PERIOD = 0), which means that the member states that volunteered for the program were not the ones that were already good performers. Thus, if we do not account for the self-selection of member states into the program, we could underestimate the treatment effect.

Our DD design directly addresses this threat to inference. Instead of only comparing pre-treatment observations in the treatment group to post-treatment observations in the treatment group (a before-after design) or comparing post-treatment observation in the treatment group to post-treatment observations in the control group (as in a randomized controlled trial), we make both comparisons to estimate the ATT. This accounts for systematic differences between the treatment and control groups due to non-random self-selection into the treatment (i.e., the fact that we are using observational data).

The identifying assumption of the DD estimator is the parallel trends assumption. Any trend in the average value of the outcome between the pre-treatment and post-treatment periods needs to be the same for the treatment group and the control group. If this assumption is not satisfied, the DD estimator will not yield an unbiased estimate of the ATT. Figure 1 (Panel A) shows the average duration of the LFN stage for the treatment and control groups in the pre-treatment and post-treatment period. The dashed line shows the counterfactual for the treatment group — what we would have expected the outcome to be without the treatment. Under the parallel trends assumption, we assume that the trend that we observe for the control group — a decrease in the average duration of the LFN stage of 103 days — also applies to the treatment group. The estimated treatment effect is the difference between the observed outcome for the treatment group in the post-treatment period and the counterfactual outcome.

A violation of the parallel trends assumption could lead us to overestimate or underestimate the effectiveness of the program. For example, if we are wrong about the counterfactual, and the average duration of the LFN stage for participants would have decreased relative to that of non-participants between the pre-treatment and post-treatment periods (i.e., would have converged to that of non-participants) even without the program, then we would overestimate the effectiveness of the program. Given that the average duration of the LFN stage for participants is longer than for non-participants, it could be that participants had a political preference for improving their cooperation with the Commission in disputes over noncompliance (that non-participants did not share) around the same time as they volunteered for EU Pilot. This could lead us to incorrectly attribute the reduction in average duration of the LFN stage to EU Pilot instead of a change in preferences. In the next section, we use a DDD design to address this possibility.

While the parallel trends assumption is fundamentally untestable, we can increase our confidence that it holds by comparing the trends for participants and non-participants during the pre-treatment period in the lead up to the implementation of EU Pilot in 2008. Specifically, we calculate the average duration of the LFN stage for all non-conformity cases

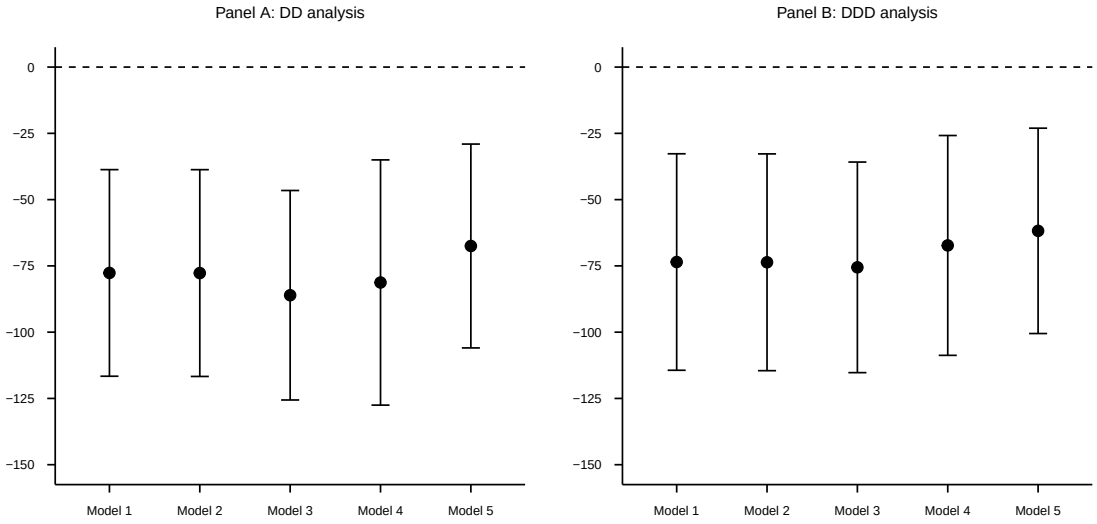

Figure 2. Panel A shows the estimated ATT for each specification in our DD analysis. Panel B show the estimated ATT for each specification in our DDD analysis.

opened in each year from 2004 through 2008 for both groups. Figure 1 (Panel B) shows visually that the trends are essentially parallel until the program goes into effect in 2008, when the duration of the LFN stage starts to improve for participants relative to non-participants. This suggests that we should expect the parallel trends assumption to hold. We address this issue further in our DDD analysis in the next section.

We calculate the treatment effect by estimating parametric DD models (see the supplementary materials for a non-parametric estimate). This allows us to include control variables for robustness. It also provides a way to estimate standard errors for our estimate of the ATT. We model DURATION as a function of the interaction of PILOT and PERIOD, including constituent terms. We estimate heteroskedasticity-robust standard errors. In a parametric DD model, the coefficient on the interaction term is the estimated ATT. Figure 2 (Panel A) shows the estimated ATT for each of our parametric DD models (see the supplementary material for full regression tables). In our base specification (Model 1), we find, consistent with Hypothesis 1, that EU Pilot reduces the duration of the LFN stage by 78 days, among participants. The effect is statistically significant at the 0.01 level. Since our DD estimator has eliminated any difference between the pre-treatment and post-treatment

period that applies to participants and non-participants, it is likely that this effect is due to an improvement in participating member states' response times.

We estimate several other specifications with control variables. Adding covariates to our DD model can control for factors that could contribute to different trends between the treatment and control groups (a violation of the parallel trends assumption). Adding case-level covariates can also reduce the variance of the estimates. Our results are robust across these specifications. The magnitude of the estimated effect varies between a 68 and 86 day reduction in the duration of the LFN stage.

First, we add a control for the workload of member states with respect to infringement cases (Model 2). For each case, we calculate the number of open cases involving that member state in the same policy area when that case was launched. Second, we include member state and Directorate-General fixed effects (Model 3). This controls for any variation in policy preferences or administrative capacity across member states and across Directorates-General that could contribute to different time trends between the treatment and control groups (a violation of the parallel trends assumption).

Third, we include fixed effects for each combination of member state and Directorate-General (Model 4). This controls for variation in member states' preferences with respect to compliance across policy areas (captured by the Directorate-General). It is likely that member states, for domestic political reasons, care more about avoiding compliance in some policy areas than others. It is also likely that the administrative capacity of member states varies across policy areas due to domestic political priorities, which affect administrative budgets. In addition, the willingness of Directorates-General to prosecute noncompliance against any particular member state may depend on the political preferences the relevant Commissioner. Finally, we include year fixed effects (Model 5), which control for time trends within the pre-treatment and post-treatment periods.

Difference-in-Difference-in-Differences Analysis

In this section, we take advantage of the fact that EU Pilot should affect the duration of the LFN stage in some cases (non-conformity cases) and not others (non-communication cases) to assess potential violations of the parallel trends assumption in our DD analysis. As mentioned above, this assumption could be violated if participants had a political preference for improving their cooperation with the Commission in disputes over noncompliance (that non-participants did not share) around the same time that EU Pilot was being implemented. To the extent that such a preference would accelerate the resolution of both non-conformity and non-communication cases, we can account for it using a difference-in-difference-in-differences (DDD) design.

The DDD estimator adds a third difference to the DD estimator, which is the difference is between non-conformity cases and non-communication cases. This removes any effect of a preference for improvement among participants that applies to both types of cases from the estimated treatment effect. Our DDD estimate of the ATT is robust to violations of the parallel trends assumption in which the trend in the treatment group is different from the trend in the control group, but within the treatment group, the trend applies to both types of cases. This narrows the assumption we have to make in order to interpret our estimate as the ATT. Specifically, we only have to assume that, among participants, the trend for non-communication cases is the same as the trend for non-conformity cases.

To estimate our parametric DDD models (see the supplementary material for a non-parametric estimate), we interact a dummy variable indicating whether the case is a non-conformity case or a non-communication case (NON-CONFORMITY) with PILOT and PERIOD. We include all constituent terms. The effect of EU Pilot on the duration of the LFN stage in non-conformity cases is the coefficient on the triple-interaction term. In all specifications, we estimate heteroskedasticity-robust standard errors.

Figure 2 (Panel B) shows the estimated ATT for each of our parametric DDD models (see the supplementary materials for regression tables). In our base specification (Model 1), we find that EU Pilot reduced the duration of the LFN stage by 74 days, among participants.

The effect is statistically significant at the 0.01 level. Our base DDD estimate is very similar in magnitude to our base DD estimate of 78 days.

We conduct the same robustness checks as in our DD analysis. The results are robust across all of these alternative specifications. The magnitude of the estimated effect of EU Pilot in this set of models varies between a 62 and 76 day reduction in the duration of the LFN stage. Since the average duration of the LFN stage for all cases in the sample is 240 days (one standard deviation is 178 days), this effect is substantively large. In sum, our findings indicate EU Pilot has been effective at reducing the duration of the LFN stage of the infringement procedure.

Causal Mechanism

Our DD and DDD results indicate that EU Pilot has increased the efficiency of the Commission infringement procedure, but these tests do not speak to the causal mechanism. If the Commission's proposed causal mechanism is correct, and EU Pilot is reducing the number of incoming non-conformity cases by preventing accidental noncompliance during the policy implementation process, the observable implication is that the cases that do enter the procedure should tend to be the ones that are harder to resolve — those involving intentional noncompliance. These cases should be less likely to resolve at the LFN stage and more likely to resolve at the RO stage (Hypothesis 2). In addition, cases should be no more likely to resolve during the pre-litigation phase.

To test the observable implications of this causal mechanism, we estimate an additional set of DDD models using as outcome variables whether the case was settled during the LFN stage, whether the case was settled during the RO stage, and whether the case was settled pre-litigation. We estimate a linear probability model by OLS so we can include member state and Directorate-General fixed effects (see the supplementary materials for regression tables). If EU Pilot works as expected, the effect on settlement at the LFN stage should be negative and the effect on settlement at the RO stage should be positive. Consistent with Hypotheses 2, we find that EU Pilot cases are approximately 12 percent less likely to

end at the LFN stage and 12 percent more likely to end at the RO stage. With respect to the pre-litigation phase, we find no effect. These findings suggest that EU Pilot works according to the mechanism posited by the Commission.

Conclusion

In this paper, we assess whether EU Pilot was effective at improving the efficiency of pretrial bargaining in the EU by preventing accidental noncompliance during the policy implementation process. This would decrease the number of incoming non-conformity cases and allow member states and the Commission to more efficiently process the remaining cases. Using DD and DDD designs, we find that EU Pilot significantly improves the efficiency of pretrial bargaining. The average duration of the LFN stage in non-conformity cases was approximately 74 days shorter for participants than non-participants.

We also provide evidence consistent with the causal mechanism. By reducing the workload of member states and the Commission, EU Pilot lowers the opportunity costs of bargaining, incentivizing parties to make initial counter-offers more quickly, shortening the LFN stage. This mechanism results in cases reaching the RO stage faster, which is the stage at which member states are likely to settle cases that involve strategic noncompliance (the RO stage is the last opportunity to avoid a costly court case). This mechanism also helps the Commission correct accidental noncompliance more quickly, reducing the period of time during which there is a compliance deficit.

Our research design uncovers a crucial insight for addressing noncompliance: improving communication with implementing actors during the policy implementation process is a viable policy solution to improve the efficiency of bargaining in pretrial dispute-settlement procedures without resorting to sanctioning mechanisms.

Word count: 7977 (7502 for the text plus 475 for two half-page figures)

Supplemental data for this article can be accessed at LINK.

Supporting data and materials for this article can be accessed on the Taylor & Francis website, doi: LINK.

Sivaram Cheruvu is an Assistant Professor of Political Science in the School of Economics, Politics & Policy Sciences, The University of Texas at Dallas, United States.

Joshua C. Fjelstul is a Post-Doctoral Research Fellow in the Department of Political Science and International Relations, University of Geneva, Switzerland.

ORCID iDs:
Sivaram Cheruvu: 0000-0002-4747-2537
Joshua C. Fjelstul: 0000-0002-9494-260X

Corresponding author:
Joshua C. Fjelstul (joshua.fjelstul@unige.ch)

Acknowledgments:
Earlier versions of this paper were presented at the 2019 annual meeting of the Midwestern Political Science Association (MPSA) in Chicago and at the 2019 annual meeting of the Southern Political Science Association (SPSA) in Austin. We thank the participants for their feedback. We also thank Cliff Carrubba, Matt Gabel, the editor, and the anonymous reviewers for their helpful comments and suggestions.

Notes

[1]In 2007, for example, the Commission published an internal memo, COM(2007)502, in which it argued that "delay or error in the application of Community law weakens the system itself [and] reduces the possibilities for its objectives to be achieved, depriving citizens and businesses of the benefits."

[2]See COM(2007)502.

[3]The original participants were Austria, Czech Republic, Denmark, Germany, Finland, Hungary, Ireland, Italy, Lithuania, Netherlands, Portugal, Slovenia, Sweden, Spain, and the United Kingdom.

[4]The 2010 Commission report on EU Pilot.

[5]During the period we analyze (2004 to 2010), the average duration of the LFN stage was 190 days for non-communication days and 387 days for non-conformity cases.

[6]The 2010 Commission report on EU Pilot.

[7]The 2010 Commission report on EU Pilot.

[8]What distinguishes our quasi-experimental design from a standard observational design is that we use a statistical technique, a DD estimator, that allows us to estimate a treatment effect based on observational data where the treatment is not randomly assigned. There is an opportunity to use a DD estimator because only some member states participated in EU pilot, creating treatment and control groups.

[9]The reason for this is the Commission's database of decisions in infringement cases is missing a small number of letters of formal notice, as suspected by Steunenberg and Rhinard (2010) and discussed in Fjelstul and Carrubba (2018).

[10]Our DD design does not capture any effect that applies to participants and non-participants. Such an effect would be absorbed by the differencing in the DD estimator, as this estimator controls for any across-the-board trends between the pre-treatment and post-treatment periods.

References

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Börzel, Tanja A. 2000. "Why there is no 'southern problem' . On environmental leaders and laggards in the European Union." *Journal of European Public Policy* 7(1):141–62.

Börzel, Tanja A. 2003. Guarding the Treaty: The Compliance Strategies of the European Commission. In *The State of the European Union*, ed. Tanja A. Börzal and Rachel A. Cichowski. Oxford: Oxford University Press pp. 197–220.

Börzel, Tanja A, Tobias Hofmann, Diana Panke and Carina Sprungk. 2010. "Obstinate and Inefficient: Why Member States Do Not Comply With European Law." *Comparative Political Studies* 43(11):1363–1390.

Bursens, Peter. 2002. "Why Denmark and Belgium Have Different Implementation Records: On Transposition Laggards and Leaders in the EU." *Scandinavian Political Studies* 25(2):173–195.

Carrubba, Clifford J. 2005. "Courts and Compliance in International Regulatory Regimes." *Journal of Politics* 67(3):669–689.

Carrubba, Clifford J and Matthew Gabel. 2017. "International Courts: A Theoretical Assessment." *Annual Reviews of Political Science* 20:55–73.

Carrubba, Clifford J, Matthew Gabel and Charles Hankla. 2008. "Judicial Behavior under Political Constraints: Evidence from the European Court of Justice." *American Political Science Review* 102(4):435–452.

Carrubba, Clifford and Matthew Gabel. 2015. *Courts and Compliance in International Law: The Role of the European Court of Justice in European Integration*.

Chayes, Abram and Antonia Handler Chayes. 1993. "On compliance." *International Organization* 47(2):175–205.

Chayes, Abram and Antonia Handler Chayes. 1995. *The New Sovereignty: Compliance with International Regulatory Agreements*. Cambridge, MA: Harvard University Press.

Checkel, Jeffrey T. 2001. "Why Comply? Social Learning and European Identity Change." *International Organization* 55(3):553–588.

Dimitrova, Antoaneta and Bernard Steunenberg. 2000. "The Search for Convergence of National Policies in the European Union." *European Union Politics* 1(2):201–226.

Downs, George W, David M Rocke and Peter N Barsoom. 1996. "Is the Good News about Compliance Good News about Cooperation." *International Organization* 50(3):379–406.

Falkner, Gerda, Oliver Treib, Miriam Hartlapp and Simon Lieber. 2005. *Complying with Europe*. Cambridge University Press.

Finnemore, Martha and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 52(4):887–917.

Fjelstul, Joshua C. and Clifford J. Carrubba. 2018. "The Politics of International Oversight: Strategic Monitoring and Legal Compliance in the European Union." *American Political Science Review* 112(1):429–445.

Gilligan, Michael, Leslie Johns and B Peter Rosendorff. 2010. "Strengthening International Courts and the Early Settlement of Disputes." *Journal of Conflict Resolution* 54(1):5–38.

Greene, William. 2012. "Econometric analysis 7th edition.".

Hartlapp, Miriam and Gerda Falkner. 2009. "Problems of Operationalization and Data in EU Compliance Research." *European Union Politics* 10(2):281–304.

Haverland, Markus. 2000. "National Adaptation to European Integration: The Importance of Institutional Veto Points." *Journal of Public Policy* 20(1):83–103.

Hille, Peter and Christoph Knill. 2006. "'It's the Bureaucracy, Stupid' Europeanization implementation policy change transformation countries." *European Union Politics* 7(4):531–552.

Hurd, Ian. 1999. "Legitimacy and Authority in International Politics." *International Organization* 53(2):379–408.

Imbens, Guido and Jeffrey Wooldridge. 2009. "Recent developments in the econometrics of program evaluation." *Journal of economic literature* 47(1):5–86.

Jensen, Christian B. 2007. "Implementing Europe A Question of Oversight." *European Union Politics* 8(4):451–477.

Keohane, Robert O. and Marc A. Levy. 1996. *Institutions for Environmental Aid: Pitfalls and Promise*. Cambridge, MA: MIT Press.

Kim, Moonhawk. 2008. "Costly Procedures: Divergent Effects of Legalization in the GATT/WTO Dispute Settlement Procedures." *International Studies Quarterly* 52(3):657–686.

Knill, Christoph. 1998. "European Policies: The Impact of National Administrative Traditions." *Journal of Public Policy* 18(1):1–28.

Knill, Christoph and Andrea Lenschow. 1998. "Coping with Europe: the impact of British and German administrations on the implementation of EU environmental policy." *Journal of European Public Policy* 5(4):595–614.

König, Thomas and Lars Mäder. 2013. "Non-conformable, partial and conformable transposition: A competing risk analysis of the transposition process of directives in the EU15." *European Union Politics* 14(1):46–69.

König, Thomas and Lars Mäder. 2014. "The Strategic Nature of Compliance: An Empirical Evaluation of Law Implementation in the Central Monitoring System of the European Union." *American Journal of Political Science* 58(1):246–263.

Luetgert, Brooke and Tanja Dannwolf. 2009. "Mixing Methods A Nested Analysis of EU Member State Transposition Patterns." *European Union Politics* 10(3):307–334.

Mbaye, Heather A.D. 2001. "Why National States Comply with Supranational Law: Explaining Implementation Infringements in the European Union, 1972-1993." *European Union Politics* 2(3):259–281.

Sedelmeier, Ulrich. 2008. "After conditionality: post-accession compliance with EU law in East Central Europe." *Journal of European Public Policy* 15(6):806–825.

Steunenberg, Bernard. 2006. "Turning Swift Policy-making into Deadlock and Delay." *European Union Politics* 7(3):293–319.

Steunenberg, Bernard and Mark Rhinard. 2010. "The transposition of European law in EU member states: between process and politics." *European Political Science Review* 2(3):495–520.

Tallberg, Jonas. 2002. "Paths to Compliance: Enforcement, Management, and the European Union." *International Organization* 681(3):609–643.

Thomson, Robert. 2007. "Time to Comply: National Responses to Six EU Labour Market Directives Revisited." *West European Politics* 30(5):987–1008.

Zhelyazkova, Asya. 2012. "Complying with EU directives' requirements: the link between EU decision-making and the correct transposition of EU provisions." *Journal of European Public Policy* 20(5):702–721.