

Chilling or Learning? The Effect of Negative Feedback on Inter-judicial Cooperation in Nonhierarchical Referral Regimes

Arthur Dyevre*, Monika Glavina†, Nicolas Lampach‡

Abstract

We exploit the nonhierarchical nature of the European Union legal system to investigate the effect of negative feedback on inter-court cooperation. We argue that, in the context of a nonhierarchical referral system, formal dismissals expose shirking which the principal, the referral court, has no formal power to curb. Yet, we find that when referring courts have experienced a formal dismissal, they are not only more likely to resubmit but are also more likely to see their references accepted. This effect increases with the number of formal dismissals previously experienced. Our findings suggest that referring courts expect significant benefits from cooperation which the referral court is able to leverage to ameliorate the quality of the referring judges' work.

Keywords: court systems; team model; agency; docket control; European Court of Justice; EU legal system

JEL Classification: K10, K40, N34, C11, C26

Word count: 10,437

Acknowledgment The authors gratefully acknowledge financial support from European Research Council Grant 638154 (EUTHORITY). We are grateful to Tomas Adamec, Angelina Atanasova, Gilian Bens, Frauke Petra Hein, David Ketch, Russell Neudorff and Anna Maria Tonikidou for invaluable research assistance.

* *Corresponding author:* KU Leuven Faculty of Law, Centre for Legal Theory and Empirical Jurisprudence, 45 Tiensestraat, Leuven, Belgium; E-mail: arthur.dyevre@law.kuleuven.be; Phone: +32 16325387

†, KU Leuven Faculty of Law, Centre for Legal Theory and Empirical Jurisprudence; E-mail: monika.glavina@kuleuven.be

‡, KU Leuven Faculty of Law, Centre for Legal Theory and Empirical Jurisprudence; E-mail: nicolas.lampach@kuleuven.be

1 Introduction

Judicial systems are commonly characterized in hierarchical terms. This is manifest in the language typically used to distinguish among the various rungs of courts comprising a country’s judiciary, with “lower”, “higher”, “appellate” and “supreme” figuring among the most frequent descriptors to designate the different echelons of the judicial pyramid. Yet court systems differ in important respects from archetypes of hierarchical organization such as the army, the modern corporation or even the (non-judicial) civil service. Higher court judges neither contract nor select their putative judicial subordinates.¹ They also lack most of the ex post control tools that usually come with a hierarchical management structure. They cannot promote, demote or fire lower court magistrates; alter their compensation package; or modify their conditions of employment. This seemingly leaves the power to reverse lower court decision as the sole lever of control available to higher court judges.

If the threat of reversal is the incentive that is supposed to drive inter-court cooperation, then court systems which, like the European Union (EU), lack an appeal procedure might seem to find themselves in a precarious situation. The EU legal system is structured around a *nonhierarchical* referral mechanism, which has three distinctive characteristics.² First, cases are referred to the European Court of Justice (ECJ) at the behest of domestic judges rather than at the behest of litigants. The referring courts, thus, act as gatekeepers of the procedure. Second, although preliminary rulings are, in principle, binding for the referring judges, the ECJ does not have the power to reverse the decisions of referring courts. Finally, referrals pertain exclusively to points of law and exclude factual deter-

¹In some countries, though, supreme court judges may exert control over the appointment of lower court magistrates via oversight bodies such high councils of the judiciary (Garoupa and Ginsburg, 2009).

²By “nonhierarchical” we mean *structurally* nonhierarchical. When it comes to the flow of communication, these legal regimes retain an element of hierarchy. Information goes up from referring courts while directions go down from the referral court.

minations. The task of implementing preliminary rulings is, thus, left to the referring courts.³

Now the puzzle that these court systems raise with regard to inter-court collaboration looks especially acute when judicial preferences happen to diverge. When judicial preferences are perfectly aligned, referring and referral judges see eye to eye, which facilitates cooperation. But what is the prospect for cooperation when judicial preferences are not perfectly aligned and the referral court signals to referring judges that it is not happy with the references they have submitted?

This is a quintessential principal-agent problem and one that nonhierarchical referral regimes, which by definition lack an appeal procedure, are not formally structured to address. The decision to refer seems to presuppose some degree of preference alignment between referring and referral court. But what happens if the referring courts shirk and submit poorly drafted and insufficiently researched work? As low quality references make their job more difficult, referral court judges may signal their dissatisfaction by refusing to review poor-quality references in the hope that negative feedback will induce referring judges to work harder. In the best case scenario, this would spur a “learning” effect, ameliorating subsequent cooperation. However, it is conceivable that the negative feedback may, in fact, discourage subsequent cooperation if referring courts, refusing to invest greater efforts in polishing references, stop submitting altogether. This possibility

³Note that the EU is not the only exemplar of such a nonhierarchical referral regime. Similar institutional arrangements have been put in place in several international legal regimes in Africa and Latin America (Alter and Helfer, 2010; Alter, 2012). Interlocutory procedures in countries with specialized constitutional courts such as Belgium, France, Italy, Taiwan and Luxembourg (Dyevre, 2013) also fit the description. Some countries with specialised constitutional courts, such as Germany, know similar referral procedures, yet the constitutional court retains the power to reverse the decisions of referring courts in case these are appealed. Likewise, referral procedures, in the form of certified questions and interlocutory appeals, exist in the United States both at the federal and at the state level as well as between federal and state courts, yet with the possible exception of certified questions addressed by federal courts to state courts, these procedures leave the door open for litigants to file an appeal and seek a reversal of the referring court’s decision on the merits. In that sense, inter-court interactions retain a hierarchical character.

has been raised by EU law scholars who have warned that the ECJ's recourse to procedural dismissals may have a "chilling effect" on referring judges (Bobek, 2008; Vink et al., 2009).

The present study tests these two competing hypotheses in the EU context. We begin by theorising the conditions under which learning and chilling will be elicited. We hypothesise that the reaction of referring judges to formal dismissal depends on the strength of their team motive as well as on their prior beliefs about the referral court's quality standards, the cost of writing high-quality references and the reputational risk they associate with a formal dismissal.

We devise separate empirical tests for the two hypotheses and evaluate them using data on referral activity in the EU legal system. To evaluate the chilling hypothesis, we examine the behaviour in year t of courts that have received feedback in year $t - 1$ either in the form of a ruling (positive feedback) or in the form of a formal dismissal (negative feedback). While inexperienced courts are more likely to face a formal dismissal, we find that exposure to a formal dismissal does not inhibit resubmission. In fact, we observe the opposite. Exposure to negative feedback results in higher resubmission rates. This result is robust to alternative operationalisations of the feedback variable and holds for experienced as well as first-time submitters. To assess the learning hypothesis, we use the ECJ's acceptance decision on resubmission as proxy for quality and look at the effect of prior negative feedback. We implement the test in a Delayed Lag Model (DLM) to account for the stronger effect of more recent feedback moments. We find that previous exposure to negative feedback is associated with lower rejection rates, with more recent exposure having a stronger effect.

To the extent that these results point to a strong team motive, we claim that they lend support to the team model. High writing costs constitute a disincentive for cooperation,

but the motivation to secure a referral court ruling is sufficiently strong to outweigh this agency cost.

Our analysis is organised as follows. Section 2 situates the study of nonhierarchical court systems within the literature on judicial organisations. Section 3 sets out our theoretical framework and hypotheses. Section 4 introduces our data. Section 5 tests the chilling hypothesis and Section 6 the learning hypothesis. Section 7 concludes.

2 Judicial Organisations

The effort to understand how layers of courts work together has spurred the emergence of two distinct models of adjudication ([Kastellec, 2016](#); [Kornhauser, 1999](#)). The agency model takes the classic principal-agent framework as point of departure and seeks to explain how higher court judges enforce their preferences on judges sitting at the lower echelons of the judicial pyramid. As higher court judges tend to lack most of the control tools usually associated with a hierarchical management structure – such as promotion, monetary rewards, sanctions or threats of termination – agency models typically assume that inter-court cooperation is sustained by the threat of reversal ([Kastellec, 2016](#)).⁴ In contrast, the team model posits that judges across layers of courts share the same objectives and explains inter-judicial cooperation by the desire to maximize the number of “correct” outcomes. To the extent that reversals occur, they are viewed as reflecting informational differences rather than misaligned preferences or shirking ([Kornhauser, 1994](#); [Kastellec, 2016](#)).

Team and agency models of adjudication ask different questions about distinct aspects of inter-court relations ([Kastellec, 2016](#); [Cameron and Kornhauser, 2006](#)). The team per-

⁴Whether judges do actually mind being reversed is itself a matter of contention. See [Kim \(2011\)](#) and [Nash and Pardo \(2013\)](#).

spective asks how rational judges pursuing the same goal should want to divide resources and adjudicatory tasks. It posits *preference alignment*. Accordingly, explaining adjudication comes down to uncovering the mechanisms and cooperative strategies allowing judges to maximize the number of correct outcomes. The agency perspective, on the other hand, asks how lower court judges can be controlled by their superiors assuming *preference misalignment*. Accordingly, agency accounts concentrate on the mechanisms presumed to enable judicial superiors to hold in check their judicial subordinates. Both perspectives are useful to understand how courts interact in non-hierarchical regimes and how these regimes differ from hierarchical judiciaries.

Obviously, in the real world, instances of both converging and diverging judicial preferences exist. In fact, legal systems are often designed in ways that reveal both team and agency concerns. The division of labour that underpins most national judiciaries, with lower courts concentrating on fact-finding and appellate courts on law-finding and law creation, is viewed as reflecting a team conception of adjudication (Kornhauser, 1994). On the other hand, mechanisms such as docket discretion (Cameron et al., 2000; Lax, 2003) and the appellate review of factual determinations in civil law jurisdictions (Kornhauser, 1994) seem to owe their existence to the desire to mitigate agency problems. The power of reversal afforded to higher courts, meanwhile, can serve both to tackle agency costs and to facilitate efficient team cooperation (Kastellec, 2016; Kornhauser, 1999; Shavell, 1995).

The most fundamental difference between hierarchical and nonhierarchical judiciaries can be similarly characterised in terms of differing team and agency concerns. Hierarchical judiciaries are structured by an appeal procedure. Litigants decide whether to appeal lower court decisions. Appellate courts review the appealed cases and then reverse the

decisions with which they disagree. In such a setting, losing litigants have the strongest incentive to file an appeal when they believe the appellate court will reach a different outcome on account of either informational or preference divergence. The resulting litigant selection effect (Cameron and Kornhauser, 2006; Shavell, 1995) means that appeal procedures can serve to address preference misalignment as well as to maximize correct outcomes.

Consider now a nonhierarchical referral regime. Instead of being vested in litigants, the right to submit cases to the referral court belongs to referring courts. These courts act as gatekeepers of the procedure. What drives the process whereby cases are selected for referral, therefore, is not litigant selection but *judicial self-selection*. The importance of this structure-induced difference cannot be overstated because the incentives of judges and litigants differ fundamentally. Whereas litigants have the strongest incentive to appeal decisions with which they expect higher court judges to disagree, judges have the strongest incentive to refer cases when they expect the referral court to render a decision they will like. Unlike an appeal, therefore, a referral presupposes that referral and referring courts share the same goal – at least the referring court must initially hold this belief. Nonhierarchical referral regimes thus are designed to operate according to a team conception of adjudication. They do not provide mechanisms to curb agency losses. As we argue in the next Section, this institutional characteristic means that negative feedback can conceivably be detrimental to inter-court cooperation, except if the referring courts' team motive is strong.

3 Theoretical Framework

Our theoretical framework seeks to capture how a referral court's feedback affects the subsequent referral behaviour of a referring court. We are primarily concerned with the behaviour of the referring court rather than with the referral court's choice of docket control policy. In particular, our analysis is not concerned with how formal dismissals can be used to structure a court's docket by signalling the referral court's policy priorities. Instead, our analysis relies primarily on the quality-vetting function of docket control to assess the impact of formal dismissals on subsequent inter-court collaboration.

Our theoretical argument assumes that the referral court has a preference for high-quality submissions over low-quality submissions and is, for that reason, willing to reject at least some low-quality submissions with non-zero probability. For a court like the ECJ, accepting a reference has both costs and benefits. As issuing a ruling requires that a reference be filed and accepted, the referral judges' influence on law-finding and law-creation depends on the number of references received and accepted. Yet accepting more references entails additional work, especially when references are badly drafted or poorly researched. Low quality references make it harder for the referral judges to predict how their doctrinal pronouncements will apply to the facts of the case or the area of law that gave rise to the referral.⁵ Similarly, when references raise questions that appear to fall outside the referral court's jurisdiction, there might be costs associated with their reformulation. Acceptance may entail additional costs such as translation (e.g. in international regimes).

For judges with limited resources but who care about leisure time as well as prestige and policy impact, acceptance is generally characterised by decreasing marginal benefit

⁵Rules of procedure in both appeal and referral regimes attest the importance judges attach to formal criteria of quality in the documents – whether briefs or references – submitted for review. US Supreme Court Rules of Procedure, for example, specify what information must be included in filed briefs, the number of pages and even the size of the paper.

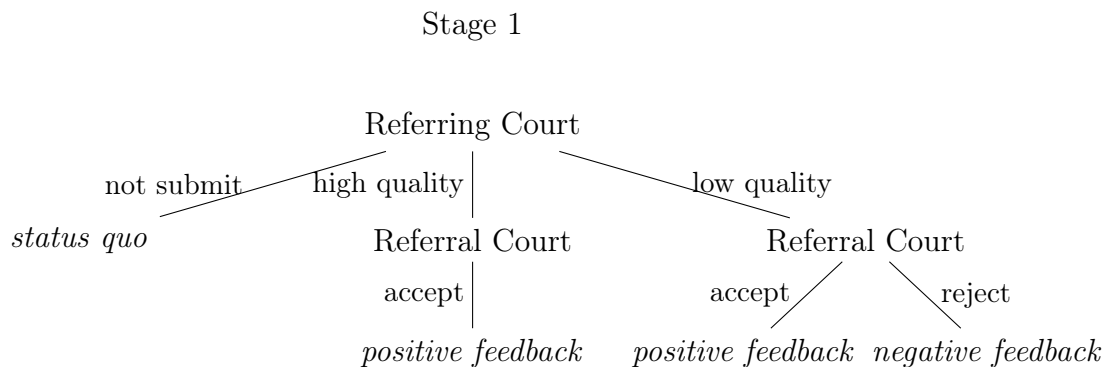
but increasing marginal costs. If its caseload is small or nonexistent, a referral court will want to accept most, if not all, references, including low quality ones. As the overall number of referrals rises, though, the referral court will tend to find a stricter docket policy more attractive.⁶

3.1 Sequence of Inter-court Interaction

Suppose a two-stage interaction sequence between a Referral and a Referring Court, as in Figure 1. In Stage 1, the Referring Court must decide whether to submit a high quality reference, a low quality reference, or not to submit at all. If it submits a case, the Referral Court either accepts or rejects the reference. The Referring Court observes the Referral Court's decision and accrues the corresponding payoff. In Stage 2, the Referring Court must decide again whether to submit a high quality reference, a low quality reference or not to refer. In this informal framework, we further suppose that the Referring Court has incomplete information over the Referral Court's acceptance criteria. Inexperienced judges may even ignore the possibility that a submission might be rejected.⁷ After observing the Referral Court's decision in Stage 1, the Referring Court revises its prior beliefs about the Referral Court's true docket policy. As we shall see, the assumption of incomplete information is important here to understand when and why chilling and learning might occur.

⁶Before getting to the stage where referral rates have reached such a level that the referral court may even consider rejecting some good references (which we believe rarely occurs in the real world), reducing the number of low quality and irrelevant references is likely to be the first priority.

⁷The Referring Court's beliefs about the Referral Court's docket policy can be represented as uncertainty over the Referral Court's true type, *strict* or *lax*. The strict type follows a strict docket policy and rejects low quality references with high probability, whereas the lax type is believed to follow a lax docket policy and to reject low quality references with low probability.



Stage 2 after negative feedback in Stage 1

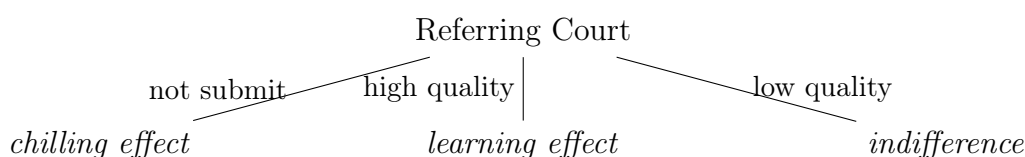


Figure 1: Sequence of interaction between Referring and Referral Court over two stages.

3.2 Case Selection Assumptions

We assume that the Referral Court has a preference for high quality submissions. This reflects the intuition that referral courts treat the choice of docket policy as an optimization problem given the overall caseload and aggregate response of referring courts. Similar to a high-ranking academic journal, the Referral Court can afford to be more demanding when submitters are numerous and, thanks to the journal’s prestige, unlikely to be put off by a rejection. The analogy with the process by which academic journals select papers for publication further suggests that to the degree that strategizing occurs, it is likely to be restricted to prominent authors. In a legal context like the EU legal order, this implies that strategic behaviour is more plausible when domestic apex courts – the legal equivalent of star academics – are involved. Our empirical design controls for this possibility by conditioning on peak courts or via the inclusion of court fixed effects.

3.3 Referring Court's Preferences

Whenever the Referring Court submits a reference and the Referral Court accepts it, the Referring Court accrues the benefit of cooperation. We assume that the Referring Court always prefers an acceptance resulting in a ruling over a rejection yielding no ruling. In the real world, referral court rulings may occasionally fail to meet the referring courts' expectations to the point that they would have preferred no ruling at all. However, what matters for the decision to submit is the expected gain from a ruling, not the actual gain.

What benefit the Referring Court expects from cooperation determines the intensity of its team motive. We do not seek to investigate the specific nature of the gain referring judges may expect from cooperation. Instead, we restrict ourselves to evaluating its intensity. We note that team models of adjudication are compatible with various understandings of the nature of this motive (Kornhauser, 1994). Referrals address the law creation and law-finding dimensions of adjudication and referring judges may draw satisfaction from securing a legally sound outcome. Alternatively though, they may seek one closer to their policy preferences or one fairer to litigants than the status quo. In other words, the benefit of cooperation can be understood in terms of the legal model, the attitudinal model or of richer, hybrid variants of these theories of judicial decision making (Stephenson, 2009; Ash and MacLeod, 2015; Epstein et al., 2013).

Preparing, writing and submitting a reference takes time and intellectual effort. Writing a high quality reference, though, requires more work than does a low quality one. A high quality reference requires a referring judge to research the case law and to write her reference carefully in accordance with Referral Court's Rules of Procedure.⁸ As we shall see, both learning and learning effects presuppose that referring judges attach some non-

⁸Survey data and qualitative interviews suggest that, in the EU context, the effort lower court judges associate with the preparation of a high quality reference is as far from negligible (Jaremba, 2012).

trivial costs to writing high quality references. This is why, ignoring or underestimating the possibility of rejection, they initially decide to submit low quality work.

Another consideration that may weigh on the Referring Court's behaviour is reputation. Because referral procedures are transparent and their outcome is public, a social stigma may be attached to the referring courts that have experienced a procedural dismissal.⁹ Depending on its perceived magnitude, this reputational risk may affect the Referring Court's decision making calculus both in Stage 1 and in Stage 2 when contemplating a new submission.

3.4 Chilling and Learning

Of central interest to the present study is the Referring Court's response to the Referral Court's decision in Stage 2. Assuming that the Referral Court always accepts high quality reference and accepts at least some low quality ones, it is easy to see that if the Referring Court submitted a high quality reference and this was accepted, its beliefs about the Referral Court's docket policy should remain unchanged. After observing a rejection, by contrast, the Referring Court should update its beliefs, leading it to revise the utility it expects from referring a low quality reference. In so doing, belief revision can thus induce a change of behaviour in Stage 2. Chilling occurs when the Referring Court, after referring a low quality reference in Stage 1, observes a rejection and then refrains from referring in Stage 2 because the costs associated with writing a high quality reference outweigh the benefits of cooperation. A chilling effect also presupposes that the higher expected

⁹Interestingly, certification processes to which a docket control mechanism within a referral environment can be compared, such as the certification of papers by academic journals, entry-level examinations or bond rating by certified bond rating agencies, tend to be opaque. [Farhi et al. \(2013\)](#) argue that the institutional rationale for opacity reflects the applicants' individual preference for nontransparency as applicants do not want others to know about their rejections. In the judicial context, opacity, while possibly advantageous for courts taken individually, would conflict with prevailing norms mandating transparency of judicial outcomes.

risk of experiencing another rejection when submitting another low quality reference will discourage this option, either because the expected benefit of cooperation has been revised down or the reputational risk has been revised up.

Learning occurs when the Referring Court submits a low quality reference in Stage 1 but decides, upon observing a rejection, to submit a high quality reference in Stage 2. Learning presupposes that the expected benefit of cooperation compensates the effort cost attached to writing a high quality reference. It also presupposes that the expected payoff of submitting a high quality reference is larger than going for another low quality reference in Stage 2.

As mentioned previously, a rejection reveals shirking on the part of the Referring Court. What explains the difference between learning and chilling is that the *learning* judge sees the benefit of inter-judicial cooperation as large enough to justify extra work whereas the *chilled* judge does not. With regard to learning, though, it is important to take into account the time elapsed between Stage 1 and Stage 2. Indeed, if a long period of time elapses between the exposure to negative in Stage 1 and resubmission in Stage 2, the Referring Court's knowledge may become obsolete, either because the Referral Court has re-adjusted its docket policy in reaction to the evolution of submission numbers or because of skills atrophy and receding institutional memory – e.g. as a consequence of change in judicial personnel. As with academic journals, a paper that would have been recognised as a high quality submission ten years ago may no longer meet the evolving standards of editors and reviewers. So, to the degree that learning occurs, we should expect time to reduce its impact.

In sum, theory suggests that, depending on the relative strength of the referring court's team motive, negative feedback may induce either a chilling or a learning effect. In the

next Sections we test the chilling and learning hypotheses using data on referral activity and formal dismissals in the European Union legal system.

4 Data: Referring Courts and Procedural Dismissals in the EU Legal System

Our raw dataset consists of all references submitted by domestic courts in the period 1961-2014. For a subset of references, we were able to build on a dataset originally compiled by [Stone Sweet and Brunell \(1998\)](#). References were coded either manually or by scraping the data directly from the EUR-Lex website¹⁰. Our dataset codes the following information: member state of origin; case number; date on which the Court of Justice received the reference; date on which the Court of Justice issued its decision (whether order or ruling); name of the referring court; and position of the referring court in the domestic judicial hierarchy (first instance, intermediate, peak court). We use this core dataset to construct separate datasets and empirical tests for the chilling and learning effects.

The referral system put in place by Article 267 of the Treaty on the Functioning of the European Union (TFEU) permits courts in all EU member states to file requests for a preliminary ruling with the ECJ. Article 267 leaves it to the discretion of domestic court judges to decide whether to file a request in cases where EU law is material to the dispute. The only exception are courts of last resort, which have a duty to request a preliminary reference when questions regarding the interpretation of EU law are raised before them. EU law, however, provides no mechanism to enforce this obligation and the refusal to refer a legal question cannot be appealed to the ECJ.

An Article 267 reference can be disposed of in three ways. As illustrated in Figure 2,

¹⁰www.eur-lex.europa.eu

the first, and by far the most common, is by a preliminary ruling of the Court of Justice – a “judgment” in the Court’s nomenclature. Next come removals. References are removed from the ECJ’s docket at the behest of the referring court. This typically occurs when the facts that initially gave rise to the referral have changed or the parties have settled. Finally, a reference may be disposed of by a procedural dismissal, known as a “reasoned order” in EU law parlance.

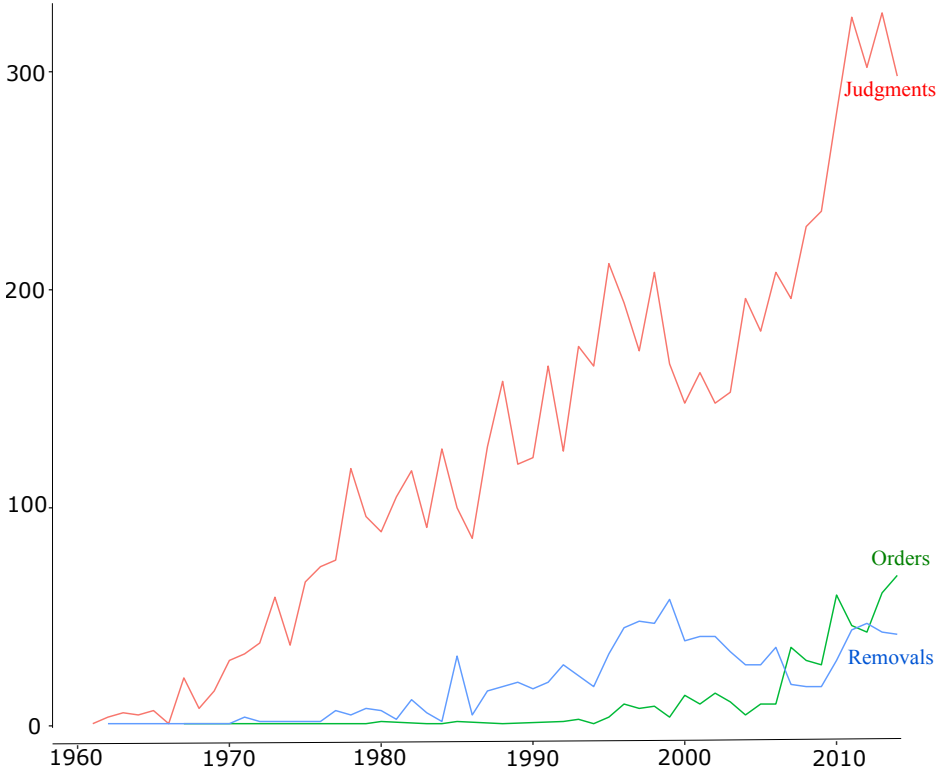


Figure 2: Judgments, removals and orders, 1961-2014.

In early years, when referral activity was still low and, therefore, the net marginal benefit of accepting references relatively high, the ECJ largely refrained from assessing the procedural validity of the questions referred during this early period (Broberg and Fenger, 2014). With the expansion of EU law and the concomitant rise in referral activity, however, the marginal reward of reviewing extra references decreased. Costs rose as enlargement required that references be translated in more languages while more complex

questions reached the Court's docket (including a rise in VAT and corporate taxation cases, see [Dyevre and Lampach 2020](#)). These developments eventually spurred the ECJ to reconsider its docket policy. Between 1961 and 1995, only 14 references were disposed of by an order. By the end of 2014 the same figure had risen to 429. [Figure 2](#) suggests that the mid-2000s marked a major inflection point in the Court's docket policy. The "Orders" line rises sharply from that point onward.

This evolution was confirmed by subsequent changes to the Court's Rules of Procedure. In 2012, formal minimum requirements were introduced for the admissibility of references. Aside from the requirement to provide a summary of the subject matter along with the facts of the domestic dispute giving rise to reference, referring courts are asked to report the language of national provisions material to the case, the relevant national court decisions and a statement of the reasons which led the referring court to ask for an interpretation of EU law.

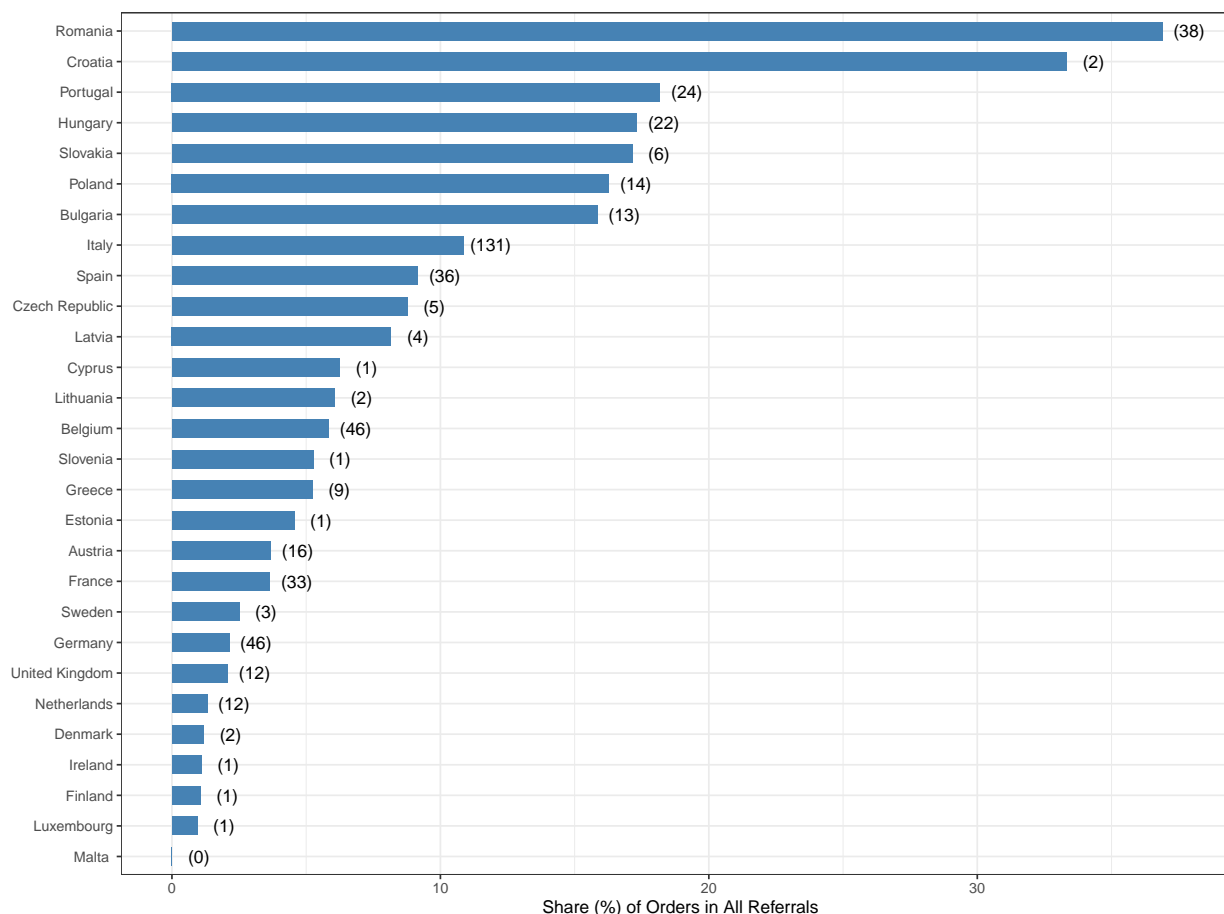


Figure 3: Share of Article 267 references resulting in a formal dismissal (reasoned order) per member state, 1961-2014. Absolute number of dismissals is in parentheses.

Legal scholars see significant disparities in the quality of references submitted by domestic courts (Galetta, 2014; Bobek, 2014; Kornezov, 2015). This is consistent with the cross-national patterns reported in Figure 3. Italian courts, notably, have been singled out for the poor quality of their references (Galetta, 2014). This perception is mirrored in the large number of Italian references dismissed by the ECJ (131). Courts in the member states that joined after 1995, particularly in Central and Eastern Europe, have also seen a substantial proportion of their references dismissed by reasoned order.

First submitters are between three and four times more likely to experience formal dismissals than more experienced courts (14.33 against 3.86 per cent). This pattern is consistent with our incomplete information assumption. Inexperienced courts are unaware

or underestimate the quality requirements imposed by the ECJ. This pattern also rules out the possibility that the ECJ may be more lenient towards first submitters as part of a deliberate strategy to socialize domestic courts into the preliminary ruling system.

Accompanying every order is some brief, perfunctory reasoning summarizing the reasons that led the ECJ to dismiss the reference. They fall into four categories: (1) lack of jurisdiction (either *rationae materiae* or *rationae temporis*); (2) reference manifestly inadmissible; (3) settled case law; and (4) other. These reasons indict the work submitted by the domestic court to a varying degree ([Kornezov, 2015](#)). (2) typically signals a badly drafted reference or one lacking necessary information such as the factual and regulatory context of the questions asked. (3) signals poor knowledge of ECJ case law. (4) includes a handful of references dismissed due to lack of a case or controversy. Category (3), followed by category (1) and (2), is the most frequent justification for procedural dismissals (Figure 4).

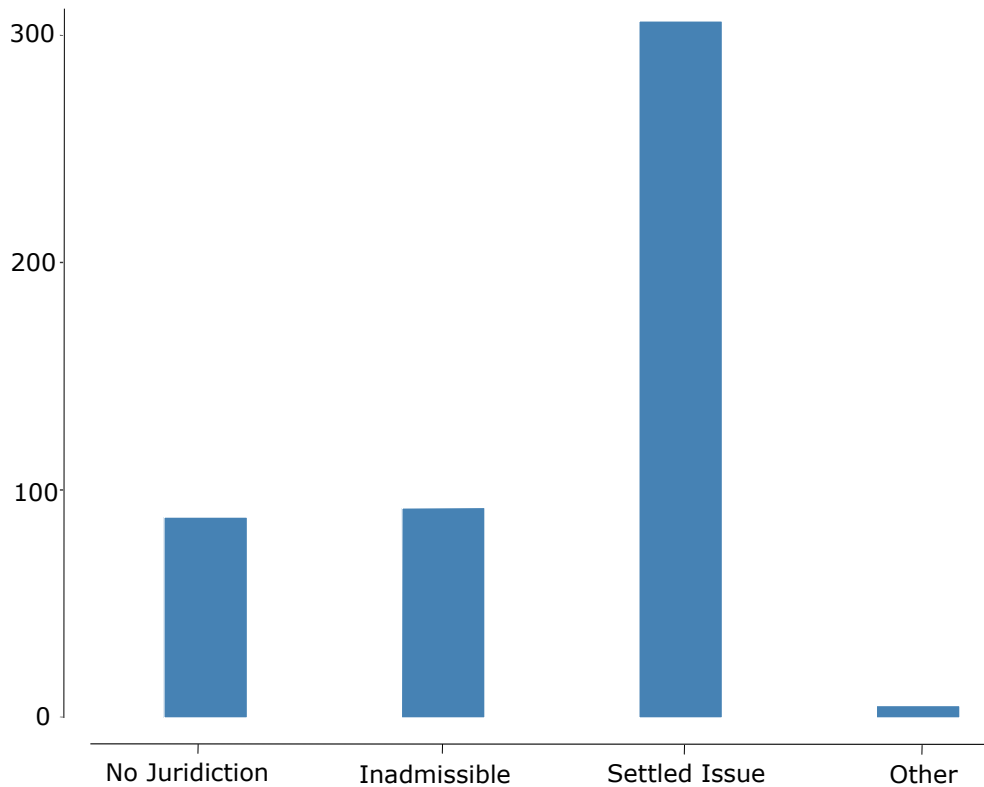


Figure 4: Orders by ground for dismissal, 1961-2014.

The nonhierarchical structure of the preliminary ruling system together with the use of reasoned orders by the Court of Justice provide a fitting environment to investigate the strength of the team motive of referring courts. In the next two sections, we present the data and empirical tests we use to assess the chilling and learning hypotheses.

5 Assessment of the Chilling Hypothesis

5.1 Empirical Test

We conduct three tests to assess the existence of a chilling effect.

Our first test operationalises the dependent variable as the number of preliminary references submitted by the court i of member state j in year t .¹¹ If their team motive

¹¹We use the date on which the Court of Justice received the reference as submission year. Though

is weak, we expect that the courts that have experienced a procedural dismissal in year $t - 1$ will be less likely to submit a reference in year t than if their reference had been accepted.

For this first test, our primary independent variable is a lag ($t - 1$) measuring the number of dismissal experienced in the previous year. Here the set of observations ($N=22,831$) includes all courts that have submitted at least once.

Second, our theoretical framework implies that, if judges were exposed to negative feedback in year $t - 1$, the choice between referring and not referring in the following period will depend, in part, on the post-feedback reassessment of the reputational risk. Thus, as a second test for the chilling effect, we operationalize our independent variable as an ordinal measure taking the value 4 if the reference is dismissed as manifestly inadmissible, 3 when dismissed on the ground that the question has already been settled, 2 when dismissed for lack of jurisdiction, 1 when rejected for other reasons, and 0 if ended by a preliminary ruling. This ordering reflects the degree of stigma legal scholars associate with each of these outcomes (Kornezov, 2015). Consistent with our theoretical framework, we hypothesize that if chilling is caused by reputational concerns, courts that have experienced a more negative evaluation will be more reluctant to resubmit than courts that have experienced a less negative one.

Finally, our third test restricts the analysis to first submitters. Because first submitters, by definition, do not have a history of interactions with the ECJ, they have less information about the ECJ's docket policy. To the extent that inexperienced submitters are initially less confident about the ECJ's true docket policy, far-reaching belief revision after experiencing a dismissal may make them more prone to chilling.

technically not the same date as the date on which the domestic court issued the reference, in practice reception follows issuance by only a couple of days.

5.2 Covariates

We control for covariates of referral activity considered in previous studies: familiarity with EU law (Hornuf and Voigt, 2015; Lampach and Dyevre, 2020), monism (Carrubba and Murrah, 2005), support for EU membership (Carrubba and Murrah, 2005), intra-EU trade (Stone Sweet and Brunell, 1998; Carrubba and Murrah, 2005), GDP (Vink et al., 2009) and enlargement (Wind et al., 2009; Lampach and Dyevre, 2020). We also control for peak court. Compared to courts lower down the domestic judicial hierarchy, peak court judges typically enjoy greater resources to process a comparatively smaller caseload. Moreover, as previously mentioned, last instance courts do not only have a duty to raise a preliminary reference when EU law is material to the dispute at hand (Dyevre et al., 2020) but the ECJ is also more likely to behave strategically when deciding whether to accept a reference submitted by such a court. Our dummy for peak court takes the value 1 if the court is identified as a peak court and 0 otherwise. Information on the definition, operationalization and data sources for our controls is provided in the Appendix.

5.3 Modelling Resubmission Behaviour

Our modelling exercise presents several challenges. First, court-level resubmission rates, our dependent variable for the chilling hypothesis, can only take non-negative integer values. Second, because lower courts are very sporadic in their use of the preliminary ruling mechanism and many never have a second go at the Article 267 procedure after their first submission, our dependent variable includes many zeros. Its distribution is best approximated by a zero-inflated negative binomial distribution. Third, while one may be tempted to apply a frequentist, fixed effects approach, this estimation strategy runs into intractable convergence problems.

To address these issues, we estimate the number of references Y submitted by court i in country j in year t conditional on lagged ECJ dismissal decision d_{ijt-1} and covariates x_{ijt} using Bayesian multilevel modelling. Multilevel models offer less protection against omitted variable bias. Yet because they discard less variation and are able to detect thresholds effects via the random intercepts, they often have greater power than fixed effects models. In that sense, if our multilevel model does not detect a chilling effect, we can be more confident that this effect is not present. The model can be expressed as:

$$p(Y_{ijt} = y_{ijt} | d_{ijt-1}, x_{ijt}, \beta_t, \psi_t) = \begin{cases} p_{ijt} + (1 - p_{ijt}) \left(\frac{\psi_t}{\mu_{ijt}\psi_t} \right)^{\psi_t} & \text{if } y_{ijt} = 0; \\ (1 - p_{ijt}) \frac{\Gamma(y_{ijt} + \psi_t)}{y_{ijt}! \Gamma(\psi_t)} \left[\frac{\mu_{ijt}}{\mu_{ijt} + \psi_t} \right]^{y_{ijt}} \left[\frac{\psi_t}{\mu_{ijt} + \psi_t} \right]^{\psi_t} & \text{if } y_{ijt} = 1, 2, \dots \end{cases} \quad (1)$$

where ψ^{-1} is a dispersion parameter with $\psi > 0$; $\Gamma(\cdot)$ is the gamma function; $E(Y_{ijt} | x_{ijt}, \beta_t, \psi_t) = \mu_{ijt}$; $\text{Var}(Y_{ijt} | x_{ijt}, \beta_t, \psi_t) = \mu_{ijt} \left(1 + \frac{\mu_{ijt}}{\psi_t} \right)$; and $0 \leq p_{ijt} \leq 1$. Note that the distribution has two components. One for $y_{ijt} = 0$ and one for $y_{ijt} = 1, 2, \dots$. The log-likelihood function for the vector of parameters $\theta = (\psi, \beta^T, z^T)^T$ has the form

$$l(\theta) = \sum_{i:y_{ijt}=0} l_1(\psi_t, d_{ijt-1}^T \beta, x_{ijt}^T \beta, z_{ijt}^T \gamma) + \sum_{i:y_{ijt}>0} l_2(\psi_t, d_{ijt-1}^T \beta, x_{ijt}^T \beta, z_{ijt}^T \gamma) \quad (2)$$

As with the distribution, $l(\theta)$ has two components. We estimate (2) via Markov Chain Monte Carlo (MCMC) to obtain the posterior inference $\pi(\theta)$ for the parameters β, γ and ψ .

We put conservative priors – which assign lower probability to implausibly large values – on our model parameters.¹² Accordingly, we assume that $\beta_{lk} = \underbrace{G_{kl}}_{1 \times k} \underbrace{\eta_k}_{k \times 1} + \alpha_{ijk}$ and $\gamma_{lq} = \underbrace{S_{ql}}_{1 \times q} \underbrace{\eta_q}_{q \times 1} + \alpha_{ijq}$ where we specify multivariate normal priors on η_k and η_q , and α_{ijk} as well as α_{ijq} follow half-student prior ($\alpha_{ijk} = t(\nu, 0, s), \alpha_{ijq} = t(\nu, 0, s)$) with $\nu = 3$ and $s = 10$.

¹²Note that flat, non-informational priors raise convergence issues.

The dispersion parameter ψ follows a gamma distribution satisfying $\psi \sim \text{Gamma}(a, b)$ where $a = 0.01$ and $b = 0.01$ and the zero-inflation probability of referrals follows a beta prior with shape parameters $\alpha = 1$ and $\beta = 1$.

5.4 Empirical Results

Our Bayesian estimation is obtained by running 20000 stationary iterations, burning the first 4000 as warm up iterates. Figure 1 reports the estimated mean posteriors of the coefficients with the 95% credible intervals for all three tests of the chilling effect.

Here $Order(t - 1)$, the ECJ's decision in year $t - 1$, is our main explanatory variable. Remember that for Test 1 $Order(t - 1)$ is operationalized as a dichotomous measure (order/judgment) whereas for Test 2 it is operationalized as an ordinal measure reflecting the intensity of the negative feedback. We can readily see that the mean of its posterior distribution is positive for Test 1 and 2. Interestingly, for Test 2 the mean posterior of the coefficient for $Order(t - 1)$ is not only positive, but the lower bound of the 95% confidence interval does not include 0, implying that a more indicting feedback has a significant as well as positive impact on resubmission rates. The mean coefficient of 0.112 means that a one-point increase in the intensity of the negative feedback results in $e * 0.112 = 1.119$ increase in the number of references submitted in the following year.

Test 3 restricts the analysis to first submitters and, as with Test 1, operationalizes $Order(t - 1)$ as a dichotomous measure order/judgment. The posterior mean is slightly negative, but the bounds of the 95% confidence interval include 0. As explained above, while our identification strategy for this set of empirical tests offers less guaranty against omitted variables bias, multilevel models tend to exhibit lower Type II error. This gives us more confidence to conclude that a chilling effect does not exist.

Substantively, these results indicate that formal dismissals do not inhibit resubmission. If anything, the opposite seems to be true: courts that have experienced an order are more likely to give the procedure another try than courts that have secured a judgment. In sum, the evidence does not support the existence of a chilling effect. Referring courts do not opt out of cooperation when formal dismissals expose shirking.

The lack of evidence for the chilling hypothesis is not enough to infer the existence of a learning effect. Learning, as we defined it, is not merely a matter of resubmitting. Rather, it is a matter of resubmitting *better* work. Referring courts may simply keep submitting the same low quality references in the hope that the ECJ will accept their submissions at the next attempt. So we now move on to examine how the ECJ evaluates submissions from the courts that have already submitted at least one reference and we compare the performances of the courts that have experienced a formal dismissal to those that have not.

Table 1: The effect of formal dismissals on subsequent resubmission estimated by Bayesian hierarchical zero-inflated negative binomial regression.

Dependent variable: Log of references in year t									
	Test 1			Test 2			Test 3		
	Post.mean	l-95% CI	u-95% CI	Post.mean	l-95% CI	u-95% CI	Post.mean	l-95% CI	u-95% CI
Intercept	-3.514	-3.874	-3.163	-3.510	-3.869	-3.153	-1.090	-1.442	-0.731
Order ($t - 1$)	0.227	-0.035	0.503	0.112	0.005	0.223	-0.002	-0.277	0.280
Peak Court	2.010	1.622	2.405	2.013	1.613	2.400	0.872	0.449	1.303
Monism	0.088	-0.145	0.322	0.085	-0.148	0.324	0.024	-0.202	0.250
EU Support	-0.691	-1.071	-0.310	-0.692	-1.075	-0.304	1.338	0.894	1.778
Intra-EU Trade	0.001	0.001	0.001	0.001	0.001	0.001	0.000	-0.001	0.000
GDP per Capita	0.163	-1.777	2.145	0.172	-1.780	2.142	-0.419	-2.350	1.519
New Member	1.631	1.243	2.019	1.631	1.242	2.019	0.133	-0.248	0.519
Familiarity EU Law	0.026	0.018	0.034	0.026	0.018	0.034	-0.016	-0.025	-0.007
σ_{ij}	1.240	1.160	1.330	1.240	1.160	1.33	0.840	0.760	0.940
ψ	0.930	0.680	1.230	0.940	0.690	1.240	0.930	0.780	1.110
z_{ij}	0.230	0.040	0.310	0.240	0.150	0.310	0.020	0.000	0.060
WAIC	17051.95			17689.89			8554.46		
$R^2_{gtmm(m)}$	0.083			0.083			0.151		
$R^2_{gtmm(c)}$	0.541			0.541			0.576		
Number of iterations	16000			16000			16000		
Number of court levels	729			729			640		
Number of country levels	27			27			27		
Original sample size	22502			22502			6118		

Note: Unlike for Test 1 and 3, for Test 2 $Order(t - 1)$ is operationalised as an ordinal measure reflecting the intensity of the negative feedback. Test 3 restricts the analysis to first submitters.

6 Assessment of the Learning Hypothesis

6.1 Testing Approach

Our analytical framework entails that if the team motive is strong and writing costs are neither trivial nor too high, the Referring Court will respond to the negative feedback by submitting a high quality reference.

Establishing a learning effect necessitates an evaluation of the quality of the referring courts' submission. Instead of developing an external measure of quality, we assume that the ECJ's decision at the acceptance stage reflects the reference's underlying quality. In other words, we assume that the ECJ is more likely to reject a poor quality reference

regardless of the domestic court’s referral history.¹³

Consistent with our theoretical framework, if a negative feedback induces a learning effect, we expect courts that have experienced a procedural dismissal to improve the quality of their references, making it less likely that they will experience a procedural dismissal at the next attempt. Accordingly, we define our dependent variable as a dichotomous measure taking the value 1 if the reference submitted in year t is disposed of by an order and 0 if disposed of by a judgment. For this test, we consider the impact of ECJ admissibility decision over all years following feedback rather than just the year after the ECJ decision. This is because we assume that the effect of a negative feedback may last beyond the year following exposure (Schwartz, 2000; Huang et al., 2005; Gasparrini, 2011). However, because knowledge and skills tend to atrophy when unused and institutional memory is likely to weaken over time, we expect the exposure effect to be larger when the exposure is more recent and weaker when the exposure is more distant in time. The universe of observations for this test are all the courts that have submitted at least two preliminary references (N=1824).

As robustness test, we restrict the analysis to those procedural dismissals that solely state formal shortcomings as reason for rejecting the submission, that is those held manifestly inadmissible (N=1768). Two considerations motivate this additional test. First, because they refer to the substance rather than to the form of the submission, the two other grounds invoked by the ECJ to reject cases, “settled case law” and “no jurisdiction”, may in some cases be interpreted as signalling the type of questions ECJ judges would like to see domestic courts referring. When so, procedural dismissals may serve a docket construction strategy which may potentially confound the learning mechanism our analysis aims to uncover. Second, some courts – particularly in jurisdictions where the authority of precedents is less established – may deliberately submit questions known to be considered settled case law simply to secure an authoritative statement that the case law applies to the case at hand.

¹³Observations by prominent ECJ members are consistent with this assumption. The ECJ’s President was about the impact of formal dismissals on the behaviour of national courts on [masked for review process]. While acknowledging the possibility of a chilling effect, he gave explicit examples of courts that improved the quality of their work after experiencing a formal dismissal. [record of interview masked for the review process].

6.2 Covariates

We control for four factors, which we believe, may affect the ECJ’s decision to dismiss a reference. The first is the hierarchical position of the referring court. For reasons pointed to above, a higher resource/caseload ratio implies that courts higher up in the judicial hierarchy have a greater capacity to produce high-quality references, which may make them less prone to formal dismissals. Our second control is familiarity with EU law, for which length of membership in the EU is a proxy (Lampach and Dyevre, 2020). Judges in member states that have been in the EU for a longer period of time, are more likely to have received EU law training, to have presided over disputes involving EU law arguments and to know the ECJ’s case law. The third factor is the ECJ’s own workload. A larger caseload may induce the ECJ to apply its docket policy with more severity. We measure caseload as the number of pending cases aggregating all procedures through which cases can be brought before the ECJ (Articles 267, 263, 259 and 258 TFEU).¹⁴

6.3 Modelling Delayed Effect of Negative Feedback on Odds of Acceptance

We assess the learning hypothesis using distributed lag modelling. Distributed lag models (DLM) are designed to capture delayed effects (Huang et al., 2005; Schwartz, 2000). The linear form of our basic DLM can be expressed as:

$$\log\left(\frac{\pi_{ijt}}{1 - \pi_{ijt}}\right) = \beta_0 + \sum_{s=0}^q \gamma_s x_{ijt-s} + \sum_{r=1}^R \nu_r \alpha_{rijt} \quad (3)$$

where γ_s is the lagged weight. Lagged weights collectively define the lag distribution. Here the lagged weights specify that the exposure linearly affects the outcome over time. The term α_{rijt} include other predictors with linear effects defined by the related parameters ν_r .

As explained above, owing to knowledge and skills atrophy or receding institutional memory, we expect the exposure effect to be larger when the exposure is more recent and weaker when the exposure is more distant in time. The literature on DLMS has proposed ways to relax the assumption of linearity (Gasparrini, 2011; Roberts and Martin, 2007).

¹⁴A filed case is coded as pending in year t if not disposed of that same year.

Delayed effects can be modelled by a polynomial distributed lag restricting the lagged coefficients to fit a cubic function. Relaxing the linearity assumption, equation (3) can be rewritten as:

$$\log\left(\frac{\pi_{ijt}}{(1-\pi_{ijt})}\right) = \beta_0 + \sum_{s=0}^q (\kappa_0 + \kappa_1 s + \kappa_2 s^2 + \kappa_3 s^3)(x_{ijt-s}) + \sum_{r=1}^R \nu_r \alpha_{rijt} \quad (4)$$

where κ_0 , κ_1 , κ_2 and κ_3 are the parameters of the cubic function describing the lagged weights. The number of parameters is chosen on the basis of a visual inspection of the data (see Figure 6 and 5 below).

In addition to conditional logit models with court fixed effects, we also report the results of multilevel variants of these models in the Supplementary Materials. As Bayesian DLM is still in its infancy and technical hurdles remain,¹⁵ we estimate the parameters of these models via frequentist EM-ML.

6.4 Results

A visual inspection of the raw data reveals a pattern that is, at least *prima facie*, consistent with the learning hypothesis. Specifically, recent exposure to a formal dismissal is associated with a lower rejection rate for the next submission. A court is less likely to see its reference rejected when it experienced its last formal dismissal the year before than when it experienced it eight years earlier. This pattern is most evident when we consider courts in member states that joined before 2004 (Figure 5). But it is also apparent when all courts are considered together (Figure 6). For courts with the most recent exposure (same year or one year before the ECJ’s acceptance decision), the proportion of dismissed references is zero or close to zero.

Figure 6 and 5 indicate greater heteroscedasticity when the last exposure to a formal dismissal goes back more than five years, with lots of peaks and troughs. This is consonant with our assumption that receding institutional memory – caused by judges leaving the court without passing on their experience to their followers – or knowledge atrophy – caused by judges failing to practice referral writing – weakens the effect of learning over time.

The difference in the shape of the LOESS curve between Figure 6 and Figure 5 results

¹⁵Notably with regard to the prior settings on the basis matrix (relationship and related space), see Huang et al. (2005). For a discussion of Bayesian DLM see Rushworth (2018).

from courts in Eastern and Central Europe experiencing high rejection rates (see Figure 3 above) as the number of submissions started picking up in these countries following accession to the European Union. Many of these Eastern and Central European courts were, it would appear, shirkers. However, these countries joined at a time when the ECJ shifted to a stricter docket policy and the plot implies that courts in these countries, too, have improved the quality of their work as a consequence of the ECJ's negative feedback. Courts that resubmitted shortly after experiencing a dismissal have seen a sharp drop in the probability to suffer another rejection. Those who waited more to resubmit, by contrast, have been less successful.

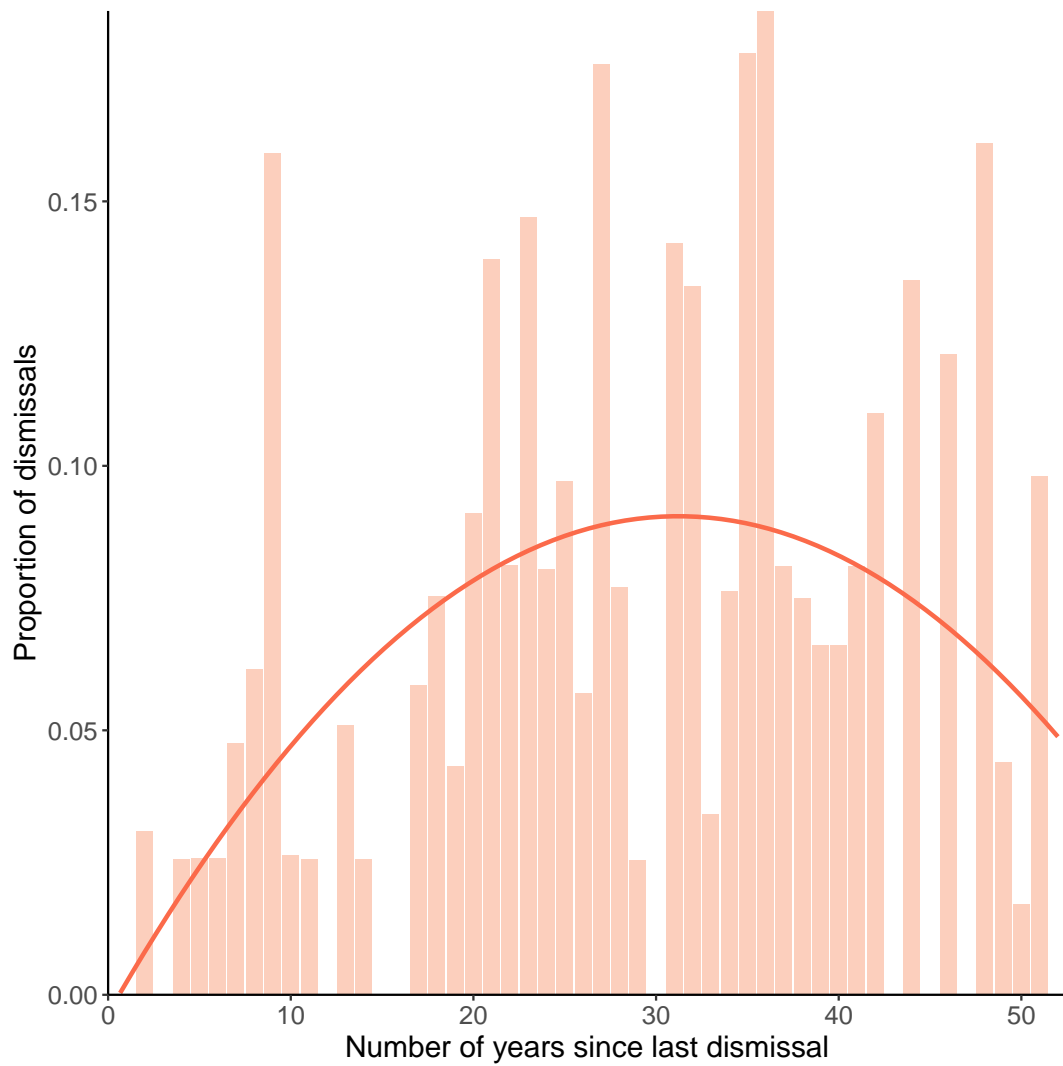


Figure 5: Proportion of rejected submissions by years elapsed since experiencing last negative feedback, EU-15.

Note: Plot shows the fraction of dismissed references as a function of the year in which the last formal dismissal was experienced. Courts in member states that joined after 1995 are excluded.

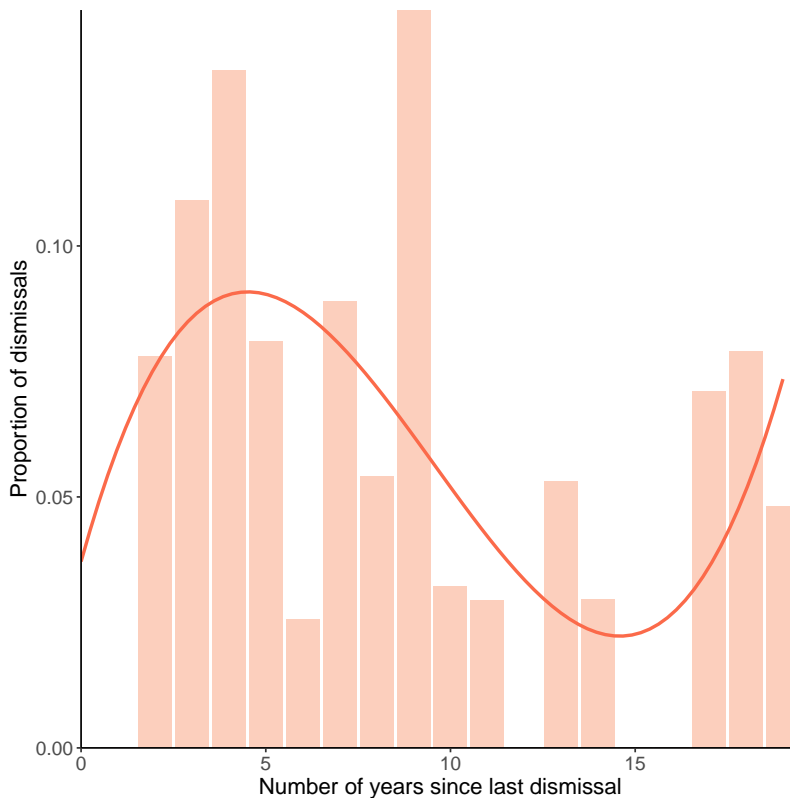


Figure 6: Proportion of rejected submissions and date of previous rejection, EU-28.

Note: Plot shows the proportion of dismissed references as a function of the year in which the last formal dismissal was experienced.

The pattern depicted in Figure 6 suggests that a non-linear model should better capture the relationship between independent and response variable. This intuition is confirmed by the results of our DLM estimation in Table 2.¹⁶ Table 2 reports the results of four model specifications based on the fixed effect conditional logistic regression estimation. The first two models correspond to our first test of the learning hypothesis. We consider all formal dismissals, regardless of the justification provided by the ECJ.

Model (1a) applies a linear form. Here our independent variable of interest is *Distributed Delayed Order* ($s > 0$). The log-odds coefficient is negative and strongly significant. This entails that the ECJ is less likely to reject a reference when the referring court has experienced a negative feedback on a previous occasion. In line with expectations, *Workload* has a significant, positive effect on the ECJ’s propensity to dismiss references. *Familiarity with EU Law* is negative and non-significant.

¹⁶Similar results are obtained from a hierarchical distributed lag model reported in Table ?? in Supplementary Materials.

Model (1b) estimates the same model but applying a non-linear form. As we can see, both *Distributed Delayed Order* ($s^2 > 0$) and *Distributed Delayed Order* ($s^3 > 0$), which account for the non-linear part of the non-linear delayed effect function, are significant.

In terms of effect size, the linear model yield relatively small coefficients. A log odds-ratio of -0.040 translates into the odds of facing rejection being about 4% lower for courts that have previously experienced a dismissal. The non-linear model (1b) shows a much larger effect size, which is consistent with our assumption that the effect of a negative feedback will disappear over time due to changes in judicial personnel or knowledge atrophy.¹⁷ The log odd ratio for $s > 0$, -29.08 , indicates that the effect of exposure to negative feedback on the probability that the ECJ will accept the next submission is rather large for recent exposures, amounting to a reduction of up to 25% in the odds a of a rejection. Log-likelihood, pseudo R^2 and accuracy imply that the non-linear form is a better fit.

As discussed above, “lack of jurisdiction” and “settled case law” may potentially be interpreted as signalling the ECJ’s preferences over the type of cases it wishes to hear rather than over the quality the references it wishes to receive. For that reason, models (2a-2b) run the same analysis but exclude the formal dismissals, which invoke these justifications. In other words, these models only consider dismissals on grounds of “manifest inadmissibility”, which cannot be interpreted as signalling issue preferences, nor used as agenda-setting tool.¹⁸ Importantly though, this restriction makes little difference regarding the substantive results, which remain consistent with the learning hypothesis. The results of our alternative multilevel model specifications which we report in the Supplementary Materials are also consistent with these findings.

We further compute the predicted values of lag-specific effects for the number of previous exposures to negative feedback and the $s + 1$ lags. We expect courts that have experienced more negative feedback in more recent periods to have learned more and thus to face a lower risk of rejection when resubmitting. Figure 7 displays the lag-specific effect of exposure to negative feedback on the likelihood to experience a dismissal in time

¹⁷The coefficients for $s > 0$, $s^2 > 0$ and $s^3 > 0$ for the non-linear model (1b) are very large and may suggest perfect separation, but this is merely an artefact of the empirical range of the terms in the polynomial function. Rescaling the polynomial terms yields coefficient values within more conventional bounds for logistic models.

¹⁸While this restriction only removes 56 observations from our dataset, due to the other restrictions for inclusion in the dataset (notably requirement of having submitted at least twice), this represents a large share of the courts that have experienced a negative feedback.

t for dismissals in $t - s$ for both fixed effects and random effects. The effect is expressed as relative risk (RR).¹⁹ It can be interpreted as a court previously exposed to negative feedback has a $(RR-1)*100\%$ higher chance of seeing its reference accepted compared to courts that were not exposed. $RR > 1$ means a positive association; $RR < 1$ a negative association. $RR = 1$ means no effect. The overall effect of negative feedback on the likelihood of experiencing formal dismissals seems robust across model specifications (see Panel 7a-7d). The decrease of the relative risk is larger for courts that have experienced more formal dismissals, implying that these courts improve the quality of their work faster than courts that have experienced fewer dismissals.

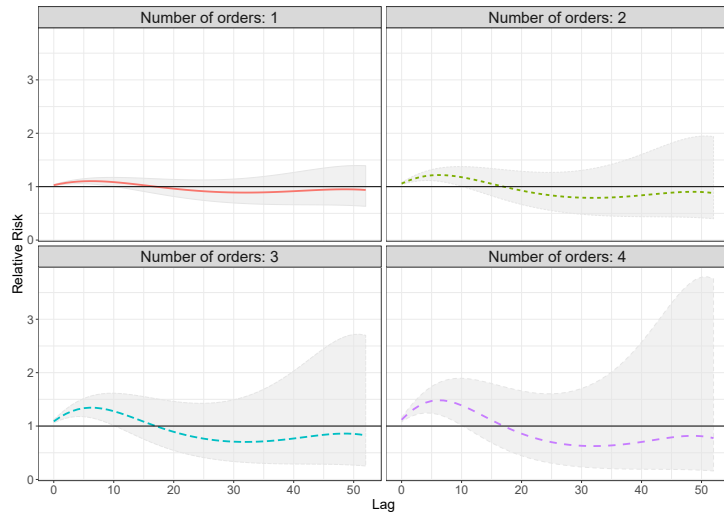
¹⁹Unlike the marginal effect, relative risk is more informative and concise when the quantity of interest is the risk of an exposed group compared to that of a control group (Hosmer et al., 2013).

Table 2: Fixed effects conditional logistic regression of ECJ admissibility decision on previous exposure to negative feedback

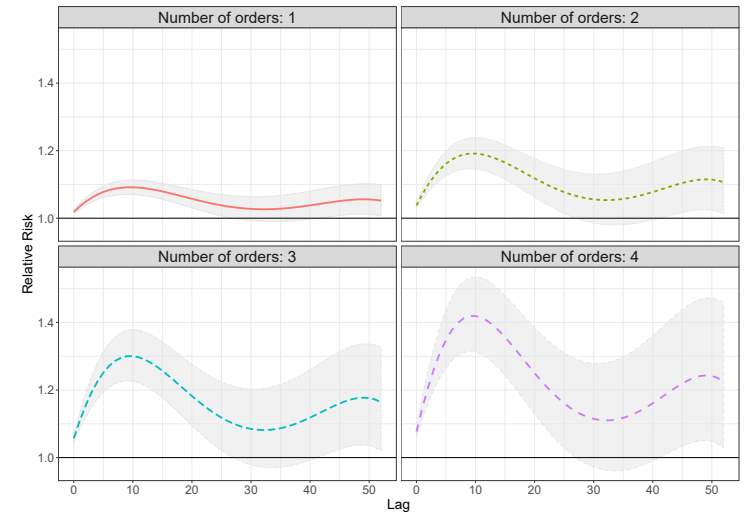
	<i>Dependent variable: ECJ admissibility decision in year t</i>			
	<i>Learning Test 1</i>		<i>Learning Test 2</i>	
	Linear Model (1a)	Non-linear Model (1b)	Linear Model (2a)	Non-linear Model (2b)
Delayed Order Effect ($s > 0$)	-0.040*** (0.007)	-29.084*** (4.259)	-0.025*** (0.008)	-6.531** (2.754)
Delayed Order Effect ($s^2 > 0$)		58.162*** (10.697)		11.062 (6.987)
Delayed Order Effect ($s^3 > 0$)		-32.570*** (7.169)		-5.745 (4.576)
ECJ Workload	0.009*** (0.003)	0.006 (0.004)	0.006* (0.004)	0.007* (0.004)
Familiarity with EU Law	-0.059 (0.039)	-0.037 (0.051)	0.017 (0.048)	0.028 (0.049)
Observations	1772	1772	1716	1716
Pseudo $R^2_{\text{Cox/Snell}}$	0.055	0.100	0.029	0.031
Accuracy	0.953	0.968	0.966	0.971
Log Likelihood	-139.375	-96.058	-99.019	-96.624

Note: *p<0.1; **p<0.05; ***p<0.01

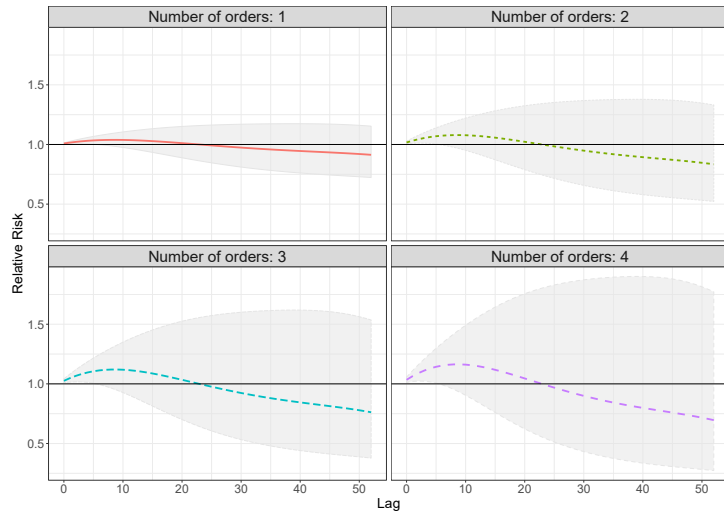
(a) Fixed effects: Learning Test 1



(b) Multilevel model: Learning Test 1



(c) Fixed effects: Learning Test 2



(d) Multilevel model: Learning Test 2

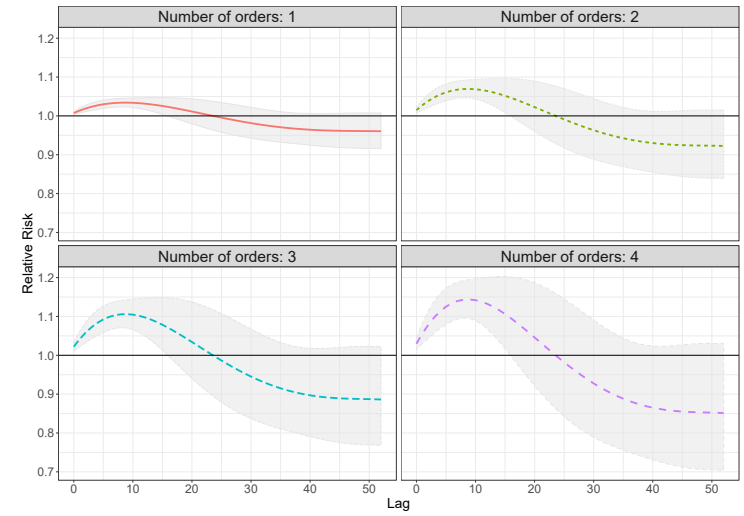


Figure 7: Relative risk of ECJ rejecting reference after previous exposure to rejection

7 Conclusion: Revise and Resubmit

How does negative feedback affect cooperation among courts in nonhierarchical referral regimes? We hypothesised that the strength of the referring courts' team motive would determine whether they drop out of cooperation or respond by improving the quality of their work. We found that the referring courts that have experienced a formal dismissal are not only more likely to resubmit, but that they are also more likely to see their references accepted when they do so, at least when the negative feedback is relatively recent. This effect of recent feedback also increases with repetitions.

The ECJ can, thus, effectively use formal dismissals to make referring courts work harder and improve the quality of their submissions. Referring courts behave like motivated academic authors submitting work to a well-established journal. Because they are highly motivated, referring courts take the editor's negative feedback as a "revise and resubmit" rather than as plain rejection. Docket control, thus, serves to raise the overall quality of submissions. This suggests that the EU court system, which lacks mechanisms to address agency losses, effectively works as a team where judges see mutual benefits in cooperation.

Our analysis presents some limitations which, we hope, future research will be able to address. First, our analysis says more about the *strength* of the team motive than about its overall *prevalence*. In nonhierarchical court systems, formal dismissals may influence self-selection and, over time, alter the profiles of the judges willing to join inter-judicial cooperation, leaving only the most motivated in the judicial team (the same may happen with high-ranking academic journals). Future research should aim to explore the possibility of temporal shifts in judicial self-selection and analyse how these relate to the institutional and personal attributes of courts and judges. Second, if the analogy with an academic journal is correct, one reason why negative feedback works in the EU may lie precisely in the fact that the ECJ is now the judicial equivalent of a high-ranked academic journal. Had the Court followed the same docket policy forty years earlier, when it was still a relatively obscure supranational body, domestic courts may not have deemed a resubmission worth the effort. Thus, the relationship between institutional standing and the timing of the introduction of docket control mechanisms may be another area worth further investigation.

References

- Alter, K. J. (2012). The Global Spread of European Style International Courts. *West European Politics* 35(1), 135–154.
- Alter, K. J. and L. R. Helfer (2010, October). Nature or Nurture? Judicial Lawmaking in the European Court of Justice and the Andean Tribunal of Justice. *International Organization* 64(04), 563–592.
- Ash, E. and W. B. MacLeod (2015). Intrinsic motivation in public service: Theory and evidence from state supreme courts. *The Journal of Law and Economics* 58(4), 863–913.
- Bobek, M. (2008). Learning to talk: Preliminary rulings, the courts of the new member States and the Court of Justice. *Common Market Law Review* 45(6), 1611–1643.
- Bobek, M. (2014). Talking Now? Preliminary Rulings in and from the New Member States. SSRN Scholarly Paper ID 2532852, Social Science Research Network, Rochester, NY.
- Broberg, M. and N. Fenger (2014). *Preliminary References to the European Court of Justice*. OUP Oxford.
- Cameron, C. M. and L. A. Kornhauser (2006). Appeals mechanisms, litigant selection, and the structure of judicial hierarchies. *Institutional Games and the US Supreme Court*, 173–204.
- Cameron, C. M., J. A. Segal, and D. Songer (2000). Strategic auditing in a political hierarchy: An informational model of the supreme court’s certiorari decisions. *American Political Science Review* 94(1), 101–116.
- Carrubba, C. J. and L. Murrah (2005). Legal Integration and Use of the Preliminary Ruling Process in the European Union. *International Organization* 59(2), 399–418.
- Dyevre, A. (2013). Filtered Constitutional Review and the Reconfiguration of Inter-Judicial Relations. *American Journal of Comparative Law* 61(4), 729–755.

- Dyevre, A., M. Glavina, and A. Atanasova (2020). Who refers most? institutional incentives and judicial participation in the preliminary ruling system. *Journal of European Public Policy* 27(6), 912–930.
- Dyevre, A. and N. Lampach (2020). Issue attention on international courts: Evidence from the european court of justice. *The Review of International Organizations*, 1–23.
- Epstein, L., W. M. Landes, and R. A. Posner (2013). *The Behavior of Federal Judges: a theoretical and empirical study of rational choice*. Harvard University Press.
- Farhi, E., J. Lerner, and J. Tirole (2013). Fear of rejection? tiered certification and transparency. *The RAND Journal of Economics* 44(4), 610–631.
- Galetta, D.-U. (2014, January). European Court of Justice and Preliminary Reference Procedure Today: National Judges, Please Behave! SSRN Scholarly Paper ID 2500746, Social Science Research Network, Rochester, NY.
- Garoupa, N. and T. Ginsburg (2009). Guarding the guardians: Judicial councils and judicial independence. *The American Journal of Comparative Law* 57(1), 103–134.
- Gasparrini, A. (2011). Distributed lag linear and non-linear models in r: the package dlmm. *Journal of statistical software* 43(8), 1.
- Hornuf, L. and S. Voigt (2015). Analyzing preliminary references as the powerbase of the european court of justice. *European Journal of Law and Economics* 39(2), 287–311.
- Hosmer, D. W., S. Lemeshow, and R. X. Sturdivant (2013). *Applied logistic regression*, Volume 398. John Wiley & Sons.
- Huang, Y., F. Dominici, and M. L. Bell (2005). Bayesian hierarchical distributed lag models for summer ozone exposure and cardio-respiratory mortality. *Environmetrics* 16(5), 547–562.
- Jaremba, U. (2012, October). Polish civil judges as European Union law judges: knowledge, experiences and attitudes.
- Kastellec, J. (2016). The judicial hierarchy: A review essay. *The Oxford Encyclopedia of Politics*, forthcoming.

- Kim, P. T. (2011). Beyond principal-agent theories: Law and the judicial hierarchy. *Nw. UL Rev.* 105, 535.
- Kornezov, A. (2015). When David Teaches EU Law to Goliath: A Generational Upheaval in the Making. In M. Bobek (Ed.), *Central European Judges Under the European Influence: The Transformative Power of the EU Revisited*, pp. 241–266. London: Bloomsbury Publishing,.
- Kornhauser, L. A. (1994). Adjudication by a resource-constrained team: Hierarchy and precedent in a judicial system. *S. Cal. L. Rev.* 68, 1605.
- Kornhauser, L. A. (1999). Appeal and supreme courts. *Encyclopedia of Law*.
- Lampach, N. and A. Dyevre (2020). Choosing for europe: judicial incentives and legal integration in the european union. *European Journal of Law and Economics* 50(1), 65–86.
- Lax, J. (2003). Certiorari and Compliance in the Judicial Hierarchy: Discretion, Reputation and the Rule of Four. *J. Theor. Politics* 15(1), 61–86.
- Nash, J. R. and R. I. Pardo (2013). Rethinking the principal-agent theory of judging. *Iowa L. Rev.* 99, 331.
- Roberts, S. and M. A. Martin (2007). A distributed lag approach to fitting non-linear dose–response models in particulate matter air pollution time series investigations. *Environmental research* 104(2), 193–200.
- Rushworth, A. (2018). Bayesian distributed lag models. *arXiv preprint arXiv:1801.06670*.
- Schwartz, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology* 11(3), 320–326.
- Shavell, S. (1995). The appeals process as a means of error correction. *The Journal of Legal Studies* 24(2), 379–426.
- Stephenson, M. C. (2009). Legal realism for economists. *The Journal of Economic Perspectives* 23(2), 191–211.

- Stone Sweet, A. and T. Brunell (1998). Constructing a Supranational Constitution: Dispute Resolution and Governance in the European Community. *American Political Science Review* 92, 63–81.
- Vink, M., M. Claes, and C. Arnold (2009). Explaining the use of preliminary references by domestic courts in EU member states: A mixed-method comparative analysis.
- Wind, M., D. S. Martinsen, and G. P. Rotger (2009). The Uneven Legal Push for Europe Questioning Variation when National Courts go to Europe. *European Union Politics* 10(1), 63–88.