Issue Attention on the European Court of Justice: A Text-Mining Approach

Arthur Dyevre* and Nicolas Lampach†

We use natural language processing techniques to explore issue attention in the decisions of the European Court of Justice. Analysing the entire universe of ECJ decisions up to 2015, we map issue prevalence across time, procedures and litigant type. We find evidence that the more inclusive annulment and referral procedures are associated with greater issue heterogeneity whereas less inclusive infringement procedure displays greater issue cohesiveness as well as greater issue stability over time. The saliency of internal market and environmental issues in infringement proceedings is, we argue, a direct consequence of the European Commission's tight grip over the Court's agenda. Interestingly, while preliminary rulings also emphasise internal market themes, they devote far more attention to employment, social and immigration questions. We compare our computer-based topic models with manual case annotations from the Court's own public database and a citation network constructed from a study of family reunion cases. We conclude that text-mining methods represent a powerful and reliable tool to summarize large collections of judicial opinions and to investigate judicial agenda formation in the European context.

Keywords: European Court of Justice, Judicial agenda, Litigation, Text mining, Topic model

JEL Classification: C38, C46, J68, K19

Acknowledgment The authors gratefully acknowledge financial support from European Research Council Grant 638154 (EUTHORITY). We are grateful to Monika Glavina, Russell Neudorf 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: nicolas.lampach@kuleuven.be

Issue Attention on the European Court of Justice:

A Text-Mining Approach

We use natural language processing techniques to explore issue attention in the decisions of the European Court of Justice. Analysing the entire universe of ECJ decisions up to 2015, we map issue prevalence across time, procedures and litigant type. We find evidence that the more inclusive annulment and referral procedures are associated with greater issue heterogeneity whereas less inclusive infringement procedure displays greater issue cohesiveness as well as greater issue stability over time. The saliency of internal market and environmental issues in infringement proceedings is, we argue, a direct consequence of the European Commission's tight grip over the Court's agenda. Interestingly, while preliminary rulings also emphasise internal market themes, they devote far more attention to employment, social and immigration questions. We compare our computer-based topic models with manual case annotations from the Court's own public database and a citation network constructed from a study of family reunion cases. We conclude that text-mining methods represent a powerful and reliable tool to summarize large collections of judicial opinions and to investigate judicial agenda formation in the European context.

Keywords: European Court of Justice, Judicial agenda, Litigation, Text mining, Topic model

JEL Classification: C38, C46, J68, K19

1 Introduction

Understanding how a policy-making institution sets its agenda is crucial to understanding what and how it makes policies. This holds for courts as well as for legislatures and other public institutions. Judicial agenda-setting, though, looks a comparatively understudied topic, at least outside the institutional context peculiar to the US Supreme Court (Perry, 2009; Kastellec and Lax, 2008). This neglect may in part be explained by the common assumption that courts cannot act ex officio. That is, in order to make policy pronouncements, courts must wait for a litigant to bring a case. A German saying encapsulates this assumption: "*Wo da kein Kläger, da kein Richter*"–or literally: "where there is no litigant, there is no judge." So, since judges have little or no control over their agenda, judicial agenda-setting seems a pointless research question, except perhaps in the rare instances where, as do the justices on the US Supreme Court, they enjoy full docket discretion.[1] Another reason for the apparent lack of scholarly interest stems from the methodological challenge involved in analysing large amounts of textual information. It is not untypical for a court to decide hundreds, sometimes thousands, of cases every year. Moreover, the issues courts get to decide are articulated in complex, unstructured opinions that easily span dozens of pages. Manually curating the relevant information from such a mass of documents, therefore, is a task beyond the resources of most research projects. In practice, because the hypotheses they seek to test often apply to a subset of cases or require identification of the policy domain, judicial scholars have relied on case annotations from legal databases and case reports. In the European Union (EU) context,

[1] For a discussion of case selection on the US Supreme Court see Perry (2009), Caldeira et al. (1999) and Kastellec and Lax (2008).

for example, researchers investigating path-dependency and policy spill-overs in EU law litigation (Schmidt, 2012; Sweet and Brunell, 1998; Stone Sweet, 2004) or clustering effects referral activity (Kelemen and Pavone, 2016) along with legal scholars focusing on more doctrinal matters (e.g. Ravasi, 2017, 11) often turn to case annotations from the official case report and "subject matter" categories from the Court of Justice's official database, CURIA, for information on issue attention. Similarly, quantitative studies of the European Court of Human Rights (ECHR), such as (Lupu and Voeten, 2012), have relied on keywords from the HUDOC database–the ECHR's official online case law repository. This strategy, however, is problematic. Indeed, little is known about the annotation process underpinning these case reports and legal databases. They provide no documentation detailing the operationalisation and implementation of their classification scheme. In fact, there are good reasons to believe that the annotation process and labelling of cases do not obey any systematic procedure.[2] These problems cast doubt on the validity of empirical findings dependent on such case annotations.[3] Besides, even if we were ready to assume these annotations are valid, they may still prove irrelevant. Indeed, as legislation, precedents and society change, so too do litigation patterns. Issues once prominent on a court's docket may recede into insignificance as new questions take centre stage. Likewise, research interests evolve. New research programmes and emerging theories may require alternative issue categorizations. This can render even the most rigorous classification

[2]The coding protocol for the Court of Justice Cases dataset compiled by Clifford Carrubba advises against using issue area codes from the ECJ case books and refers to correspondence with the Court as indication that the choice of terms is not based on a well-defined coding scheme from the Court. See http://polisci.emory.edu/home/people/carrubba_ecjd/ECJ_Access_Data_Codebook.pdf (Accessed 5 July 2018).

[3]Alternatively, judicial scholars have used full-text keyword searches to establish the issue area of cases (see e.g. Conant (2006)). The downside of this strategy, though, is that it presupposes that the researchers knows all the relevant keywords a priori, an assumption that is often unrealistic.

schemes obsolete. The much-praised and widely used US Supreme Court Database, for example, applies a classification scheme that may better capture issue saliency during the Warren and Burger Court than for the current Supreme Court (Shapiro, 2008, 494).[4] As time goes by, therefore, it becomes less relevant to the evolving agenda of both the Court and judicial scholars.

By offering a solution to the methodological challenge of summarizing information from thousands of judicial opinions, computerised text-mining techniques promise to advance the study of judicial agenda formation. These techniques have already been applied to a varied set of classification and information-extraction tasks in connection with the judicial process. Evans et al. (2007) apply naive Bayesian classifiers together with Wordscores to estimate the policy positions of party and amicus curia briefs. Corley et al. (2011) and Corley (2008) use plagiarism software to investigate the influence of, respectively, lower court opinions and party briefs on the content of US Supreme Court opinions. Sulea et al. (2017) use the content of French Court of Cassation opinions to predict case disposition and legal area. Aletras et al. (2016) apply a similar approach to predict European Court of Human Rights (ECHR) decisions. Mochales and Moens (2011) discuss approaches to argumentation-mining and argument-detection and present an application to ECHR opinions. Carlson et al. (2015) investigate the writing style of US Supreme Court Justices through an examination of the incidence of function words. Rice (2014) trains an ensemble classifier to measure lower court issue attention. Livermore et al. (2017) employ topic modelling to compare the evolution of federal appeal court writing style with US Supreme

[4]Recent experiments and recoding exercises have cast doubt on the validity of the Database's issue codes, see Harvey and Woodruff (2011). Scholars have also deplored the Database's tendency to restrict issue code to one per case (Edelman and Chen, 2007; Shapiro, 2008).

Court opinions. Carter et al. (2016) use the same method to study the decisions of the High Court of Australia.

This paper applies topic modelling and corpus comparison techniques to explore agenda formation on the European Court of Justice (ECJ). The trajectory of legal integration in Europe along with the ECJ's institutional architecture make the Court's rulings a particularly interesting case to assess how text-mining methods might contribute to the study of judicial agenda-setting. European integration has been accompanied by a rapid expansion of the policy remit of EU institutions. The impact of this evolution on litigation and the ECJ's agenda, though, has been mediated by a highly differentiated procedural setup. Access rules to the ECJ is governed by separate procedures, which contemplate different objects and empower distinct sets of litigants. We collect the entire universe of ECJ rulings up to 2015 and use various topic models along with corpus dissimilarity analysis to measure, visualize and compare issue prevalence across time, procedures and litigant characteristics. We find evidence that the more inclusive procedures–that is, the preliminary ruling mechanism and the annulment procedure–are associated with greater issue diversity whereas the less inclusive infringement procedure is associated with both greater issue stability and greater issue cohesiveness. Infringement cases tend to focus on internal market and environmental issues, reflecting the European Commission's enforcement priorities. Preliminary rulings address a broader range of topics, from customs and consumer protection to public procurement, employment, welfare benefits, driving licences and immigration. Annulment cases feature a medley of economic and administrative issues, including competition, trademark, agriculture, EU external relations and EU staff

disputes. Annulment cases filed by national governments mostly relate to the agricultural guarantee fund, state subsidies to domestic industries, border crossing and food imports. Private litigants use the annulment procedure to challenge EU regulatory decisions on competition, dumping and trademarks, while cases brought by EU institutions tend to centre on external relations, environmental measures and competition issues.

While showing how different methods and modelling approaches can be used to represent the Court's case law, our analysis documents several important agenda shifts. We find that commercial policy, social policy, free movement of goods and constitutional issues have declined in relative importance in referral proceedings. So too has the regulation of the coal and steel sector in the context of the annulment procedure. By contrast, taxation, labour, telecommunications sector regulations and, to a lesser degree, immigration and citizenship issues have attracted more litigation. Taxation, in particular, has become more salient in both preliminary and infringement decisions.

The results of our automated case classification do not contradict the perception of legal and judicial scholars. However, we argue that they provide a richer and more precise measure of issue prevalence than impressionistic statements based on a handful of cases or even official case report annotations. Topic models, in particular, produce meaningful categorizations, which, in addition, can occasionally reveal unexpected themes or cast light on patterns overlooked by human case annotators. To demonstrate the validity of our text-mining approach, we relate our machine-generated topics to manually collected cases and case citations from a study of family reunion rulings. Not only do manually collected family reunion rulings correlate with the expected machine-generated topic, but

the more embedded rulings–those that are more central to the citation network–also feature a greater proportion of that very topic. Finally, we compare our automated case classification to a selection of area codes from the Court's official database. The pattern of correlations between manual and automated case categorizations suggests that the two classification methods overlap to a great extent but differ sufficiently to make computer-assisted case classification an attractive alternative. We conclude that natural language processing techniques offer a powerful and reliable approach to investigate agenda formation in the European judicial context.

2 Judicial Agenda Formation: Theory

2.1 Standing, Docket Control and Issue Attention

In democratic regimes, the rule is that legislatures freely set their agenda.[5] Not only can legislators decide *what* and *when* to decide. They can also decide *not to decide*. In sum, legislators enjoy positive agenda control–the ability to get an issue onto the legislative agenda–as well as negative agenda control–the ability to keep issues off the legislative agenda. On the face things at least, judicial institutions appear far more constrained in regard to both positive and negative agenda control. First, courts are normally dependent on the initiative of other actors. Unless a litigant brings issue A to the courts, the courts will not address A.[6] Second, if a litigant brings A, the courts may have the obligation to

[5] This rule knows, of course, a number of exceptions. In parliamentary systems, legislative procedures may afford the executive branch substantive control over the agenda of the legislature (Huber, 1992). In the EU, the treaties severely restrict the agenda-setting powers of the European Parliament, the democratic world's largest supranational assembly.

[6] That policy questions are raised within the context of dispute resolution severely restricts the scope for issue creation. This is because policy determinations are expected to bear some relevance to the dispute being resolved (Cameron and

address it, even in instances where they would deem it preferable to evade or delay A's resolution.

Where restrictive rules of standing combine with a mandatory docket, the litigants who have standing to initiate action may, indeed, be in position to exert tight control over the judicial agenda. Yet where these constraints are looser judges may have greater influence over litigation dynamics and, thereby, over the range of issues that land on their docket. Indeed, the combination of broad rules of standing with a discretionary docket, as with the US Supreme Court, may grant judges' wide-ranging positive and negative agenda-setting powers, to the point of making them comparable to those of a legislature.

While docket management rules determine the judges' ability to remove cases from their docket, standing rules establish both *who* is entitled to lodge a suit and *what* constitutes an admissible cause of action. Admissible causes of action may embrace challenges to the legality of executive or legislative acts, personal harm or criminal enforcement, while authorised case initiators may include individuals, private organisations, other courts and, under international regimes, sovereign states (Alter, 2006). Many combinations are possible and the spectrum of institutional variation is large. At one end of the spectrum are the courts where virtually all persons and organizations justifying a broadly defined cause of action enjoy standing.[7] At the other end are the courts whose access is restricted to state parties or which allow actions to be brought only on a very narrow set of grounds.[8] Some

Kornhauser, 2013). Judges cannot easily escape this institutional constraint. For a discussion of issue creation on the US Supreme Court see Epstein et al. (1996).

[7] The *actio popularis*, which allows any person to bring action in the interest of the public, can be viewed as an extreme variant of such an institutional arrangement. Such a procedure existed in Hungary between 1990 and 2011 before the Constitutional Court. The Hungarian *actio popularis* led to an explosion in constitutional litigation and allowed the Hungarian Constitutional Court to become one of the world's most activist judicial constitutional tribunals during this period (see Scheppele, 2005).

[8] Only national governments can bring disputes before the International Court of Justice. The same holds for the WTO

international regimes allow individuals to file their claims with a commission, which then decides what cases to pass on to the international tribunal.[9] Other regimes have different standing rules depending on the cause of action and remedy. With regard to constitutional courts, for example, the power to initiate abstract review is typically restricted to legislators whereas concrete review can only be initiated by ordinary courts.

Standing rules directly affect positive agenda control. Broad standing rules have been linked to more powerful courts. Alter (2001) has argued that restrictive access rules placed severe constraints on the ability of the French Constitutional Council to influence the process of European integration whereas, thanks to multiple access channels and a broader pool of potential case initiators, the German Constitutional Court had repeated opportunities to reopen policy deals struck in Brussels or Luxembourg. The literature on international courts, in particular, associate private actor access with greater judicial influence and effectiveness (Keohane et al., 2000; Alter, 2006, 2001, 2012; Conant, 2006). This view rests on the recognition that public and private litigants differ in their number as well as in interests and motives for litigation. Where only state parties can bring cases, the pool of authorised litigants will never exceed the number of independent states. Where, as with abstract review cases on constitutional courts, only parliamentary groupings have standing to initiate proceedings, the pool of potential litigants will be even smaller. Allowing for private actor access, on the other hand, may increase the number of potential litigants by several orders of magnitude. Second, being more numerous, private

Dispute Settlement Body. Prior to 1974, only the president of the republic, the prime minister and the presidents of the two parliamentary assemblies had access to the French Constitutional Council. Moreover, they could only refer bills up to their promulgation into law and, before 1971, only on the grounds that the legislature had encroached on executive prerogatives (Stone, 1992).

[9] The Inter-American Court of Human Rights and the European Court of Human Rights prior to 1998 illustrate such an arrangement.

litigants are also more heterogeneous. Corporations, NGOs, minorities, rich and poor individuals have distinct interests and are, therefore, unlikely to be driven by the same motives. Where standing rules permit it, judges can exploit this motivational heterogeneity to achieve greater control over their agenda. By interpreting laws and developing legal doctrines in ways that incentivise specific subsets of litigants, judges can effectively orient litigation in the direction they deem most desirable (Baird, 2004; Baird and Jacobi, 2009). To the extent that they can be litigated in courts, changes in legislative policies may also have a more far-reaching impact on issue attention when rules of standing are more inclusive.

From these considerations two broad conclusion follow. The first is that, when docket discretion is limited or inexistent, the range of issues addressed in court rulings is likely to closely mirror the litigation agendas of authorised litigants. The second is that more inclusive rules of standing are more likely to produce shifts in issue attention as new doctrines spur fresh litigants to enter the legal process.

2.2 ECJ: Procedural Differentiation and Litigation Dynamics

The ECJ operates under what is, in principle, a mandatory docket regime. Cases can be through three distinct procedures, which correspond to distinct causes of action and empower distinct classes of litigants. First, the preliminary ruling mechanism established by Article 267 of the Treaty on the Functioning of the European Union (TFEU) permits–and, in the case of last instance courts, mandates–national judges to refer questions to the ECJ regarding the application and interpretation of EU law. Formally, domestic

judges are the gatekeepers of the procedure. Yet, because private litigants have the opportunity to try and persuade domestic judges to submit references, this channel is commonly viewed as giving private actors semi-direct access to the ECJ (Börzel, 2006; Alter, 2006; Keohane et al., 2000). The ability to dismiss references that are "manifestly inadmissible" or that pertain to "settled case law" affords the ECJ a limited measure of negative agenda control (Craig and Burca, 2015, 484).[10] The second access channel to the European Court is the infringement procedure laid down in Article 258 TFEU. The object of an infringement action is always the conduct of a member state alleged to have contravened EU law. The power to initiate infringement proceedings lies with the European Commission, which Article 258 elevates to the status of chief EU prosecutor. By virtue of Article 259 TFEU each member state is also formally vested with the right to initiate infringement proceedings against another member state. Private persons have the possibility to file complaints over breaches of EU law with the Commission–a practice encouraged by the Commission, which uses private complaints as a tool to detect violations (Harlow and Rawlings, 2006; Smith, 2008). However, the decision to bring a case before the Court is one over which the Commission always has the ultimate say. Finally, a case can reach the Court via the annulment procedure (Article 263 TFEU) when the object of the case is the abrogation of an EU act. Member states, the European Parliament, the European Commission or the Council have standing to bring annulment against any EU act. The Court of Auditors, the European Central Bank and the Committee of the Regions can also bring actions for the purpose of protecting their prerogatives. Private

[10] There has been no empirical research of note on issue suppression on the ECJ. We point it out as an interesting area for future research.

litigants and EU civil servants enjoy limited standing under this procedure. According to the ECJ's reading of Article 263 TFEU, private plaintiffs are entitled to bring annulment proceedings only when "directly" and "individually" affected by the EU measure at issue (Craig and Burca, 2015, 515). This precludes challenges to general EU legislative acts such as EU directives, but covers individual decisions as well as appeals against decisions of the General Court–the lower EU court.

Aside from empowering different classes of litigants, the three procedures just sketched out correspond to different judicial roles (Alter, 2006). In terms of their respective object, or cause of action, they present some degree but not full overlap. EU acts can be challenged under both the annulment and preliminary ruling procedure, but not under the infringement procedure. Likewise, national legislation and practices can be challenged under both the infringement and preliminary ruling procedure, but not under the annulment procedure. Thus cross-procedural variations in relative issue prevalence can arise from differences in case initiators as well as from differences in permissible causes of action. The linkage between procedure, litigation patterns and issue attention is one that is widely acknowledged in EU studies. In some instances, the effect of the identity of the case initiator on case selection is obvious. National governments have initiated hundreds of annulment cases (Adam et al., 2015) but only six infringements cases (Craig and Burca, 2015, 454). Similarly, both legal scholars and political scientists recognize that most ECJ landmark decisions–including the constitutional rulings that have articulated the doctrines of supremacy and direct–have come out of the preliminary ruling system rather than out of infringement or annulment proceedings (Craig and Burca, 2015; Sweet

and Brunell, 1998). The preliminary ruling mechanism is, for that reason, regarded as central to the legal integration process as well as to the ECJ's emergence as one of the world's mightiest judicial bodies (Craig and Burca, 2015; Stone Sweet, 2004; Alter, 2006).

Generally speaking, because of their greater inclusiveness, we should expect the preliminary ruling procedure and, to a lesser extent, the annulment procedure to be associated with greater issue diversity. By contrast, we should expect the less inclusive infringement procedure to exhibit greater issue cohesiveness as well as greater issue stability over time. The preliminary ruling mechanism enables the ECJ to use its case law to socialize new domestic judges and litigants into the development of EU law. The annulment procedure offers the Court similar opportunities to leverage motivational heterogeneity to shape its docket. The same does not apply for the infringement procedure. True, the Commission's preferences and priorities may change over time. But it is reasonable to assume that, if at all, this will occur at a slower pace than the pace at which domestic judges and litigants respond to the Court of Justice's doctrinal signals. The institutional configuration of the infringement procedure, together with the member states' reluctance to use Article 259, gives the European Commission's tight control over the issues that might seek their way to the Court's docket via this channel. For that reason, we should expect issue attention in infringement rulings to closely mirror the Commission's policy agenda, notably its internal market agenda (Tallberg, 1999; Börzel, 2003).

3 Methodology: Automated Content Analysis

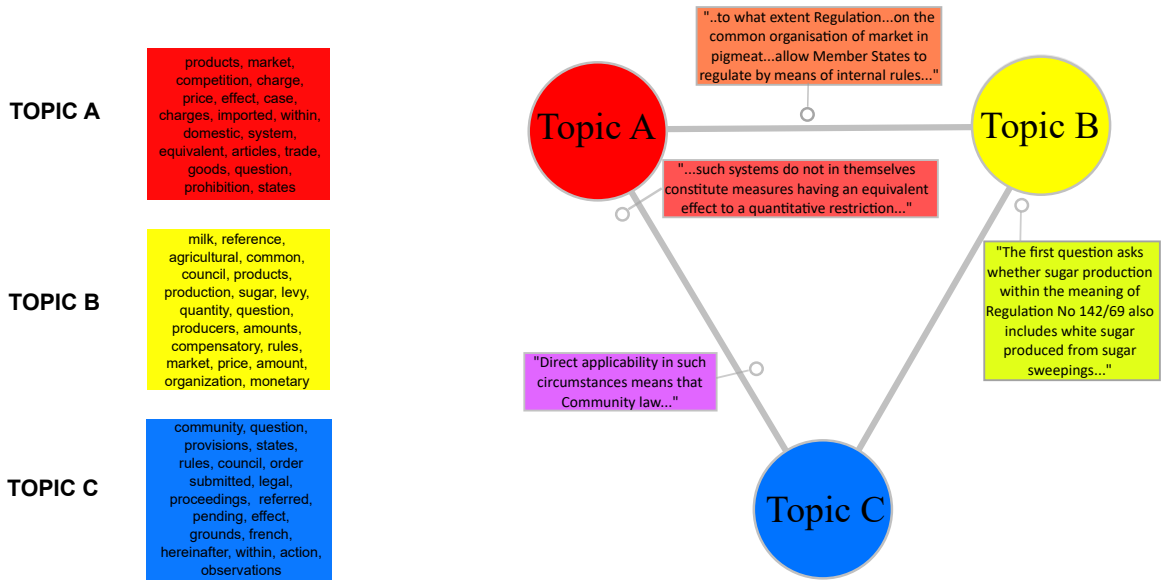
The present Section introduces the natural language processing techniques we employ to investigate issue prevalence on the European Court of Justice. Rather than treating all the rulings as a single corpus, our approach treats each procedure as forming a distinct corpus of judicial texts. This has several advantages. First, this averts the problem posed by procedural language. Owing to the object of the procedure, every infringement ruling features the words "state", "fulfil", "obligation" and "failure". Likewise, "court", "refer", "preliminary", "proceedings", "main", "ruling" and "question" occur in virtually all preliminary rulings. The same holds for the word "appeal", "decision", "contested", "measure" and "adopt" in annulment decisions. These words are procedurally distinctive but substantively uninteresting from the perspective of the study of judicial agenda formation. Treating procedures as separate corpora prevents these words from distorting the results of our modelling exercise. Second, it allows us to document how issues cluster around broader themes within procedures. Third, because EU scholars tend to study each procedure in isolation, this approach makes it easier to relate our results to the literature. However, for our corpus comparison exercise, which aims to supplement the results of our topic models regarding divergence in issue attention among procedures, we remove the procedure-distinctive procedural words and conduct the analysis on the resulting corpora.

3.1 Probabilistic Topic Modelling

Probabilistic topic modelling is a suite of methods developed for the purpose of discovering and annotating large archives of documents with thematic information (Blei, 2012).

The simplest topic modelling technique is based on *latent Dirichlet allocation* (LDA). The basic intuition behind LDA is that documents addressing the same topic are likely to contain similar words. Words such as "import", "products" and "trade" will appear more often in decisions about barriers to cross-border trade whereas "milk", "quantity" and "amounts" will be more common in decisions about farming regulations. A judicial decision will typically concern more than one topic (Shapiro, 2008; Edelman and Chen, 2007). A decision about trade may also raise procedural questions or a case may pertain to the importation of dairy products. LDA builds on these assumptions to model topics as a cluster of words and documents as a mixture of topics. More specifically, it posits a latent space in which topics are represented as distributions over words and documents as distributions over topics (Chang et al., 2009; Blei, 2012; Blei et al., 2003). The number of topics K is set by the researcher–although heuristics, such as perplexity and semantic exclusivity, have been proposed (Wallach et al., 2009; Chang et al., 2009; Mimno et al., 2011). The K parameter determines how fine-grained a summary of the corpus the final topic model will provide. Varying the number of topics thus effectively allows the analyst to zoom in and out to find specific or broader themes. As opposed to the number of topics, the words defining the topics are not chosen by the researcher but emerge from the analysis of the documents. Figure 1 illustrates the intuition behind probabilistic topic modelling with three topics and four opinions. Topic A, B and C are represented by the words with which they are most closely associated, as they typically are in the output of a topic model. Obviously, if a decision mostly addresses barriers to trade in manufactured goods, it will have a higher probability of featuring Topic A. So its position in the latent

topic space will be closer to that topic. A decision focusing on the regulation of the sugar market, by contrast, will have a high probability of featuring Topic B. Accordingly, its position in the latent topic space will be closer to B than to A or B. A decision which discusses trade in manufactured goods but elaborates extensively on constitutional issues such as the direct applicability of supranational rules will have high probabilities over both Topic A and C. In the topic space it will occupy a position somewhere between these two topics. Now, to find the topics and assign the documents to the topics, LDA effectively treats the observed data, the word counts in the documents, as arising from a hidden generative process. The goal of the estimation is to infer this hidden topic structure from the observed documents.



(a) Topics (b) Latent topic space

Figure 1: Illustration (adapted from Chang et al. (2009)) of the latent space of a topic model for judicial opinions with topics represented as distributions over words and documents as distributions over these topics. On the left are three topics, with their most characteristic words, from a topic model with $K = 25$ topics of Article 267 preliminary rulings. On the right is a simplex showing the distribution of the three topics associated with four opinions. The distance between document and topic reflects the topic proportion in the document, with shorter distance denoting greater topic proportion.

The proportion of topics in documents is treated as a random variable drawn from a Dirichlet prior distribution–hence the method's name. Formally, estimating an LDA model requires computing the following posterior distribution:

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D}) = \frac{p(\beta_{1:K}, \theta_{1:D}, z_{1:D}|w_{1:D})}{p(w_{1:D})} \quad . \tag{1}$$

where $\beta_{1:K}$ are the topics and β_k is the distribution over words for topic k; θ_d is the topic proportion for document d (and $\theta_{d,k}$ is the proportion of topic k in document d); z_d is the topic assignments for words in document d and w_d are the observed words for doc-

ument d. The numerator on the right-hand side of the equation is the joint distribution of the hidden and observed variables while the denominator is the marginal probability of the words observed in the corpus. Because the problem thus posed is computationally intractable, probabilistic topic models use algorithmic methods to approximate the posterior distribution, such as Gibbs samplings and variational inference (Arora et al., 2013; Blei, 2012).

Our analysis applies the simple LDA model implemented in the *lda* package for R to construct a synchronic summary of the ECJ's caselaw. To analyse agenda shifts through time, however, we turn to dynamic topic modelling. Whereas the plain LDA model assumes that the order of documents within the corpus does not matter, dynamic topic models allow topics to change over time by representing topics as sequences of distributions. We use the structural topic model developed by Roberts et al. (2016) and implemented in the **stm** package for R. Building off from the Correlated Topic Model (Blei et al., 2007), the model assumes a logistic normal prior instead of a Dirichlet prior for topic proportion. Time formally enters the document-generating process as a covariate interacting with topic prevalence:

$$\theta_{1:D}|t_{1:D}\gamma, \Sigma \sim LogisticNormal(\mu = t_{1:D}\gamma, \Sigma). \tag{2}$$

where t_d is the year in which document d was issued; γ is a $p \times (K-1)$ matrix of coefficients for topic proportion and Σ is a $(K-1) \times (K-1)$ covariance matrix. As implemented in the **stm** package, the posterior distribution for this dynamic topic model is computed via variational Expectation Maximization.

The variety of actors granted standing under the annulment procedure allows us to assess whether topic prevalence within this particular procedure is systematically related to the type of case initiator. To do so, we construct an author-topic model (Rosen-Zvi et al., 2004; Blei, 2012) by replacing the time covariate in equation 2 with a plaintiff covariate. Plaintiffs are in effect treated as if they were co-authors of the cases they initiated. We distinguish seven categories of plaintiffs: (1) member state, (2) private persons (including corporations), (3) European Commission, (4) Council, (5) European Parliament, (6) EU staff and (7) others. We use the same **stm** implementation to compute the model.

3.2 Between-Corpus Comparison

We complement our document-level topic modelling with a corpus-level analysis of issue dissimilarity.[11] A first corpus-level measure of dissimilarity considers what terms are distinctive of a corpus assuming that all corpora are drawn from the same population. Let $p_{w,c}$ be the rate at which word w occurs in corpus c and C be the number of corpora. How distinctive of corpus c word w is is determined by the extent do which its count in c deviates from expectation:

$$p_{w,c} - \frac{\sum_{c=1} p_{w,c}}{C}. \tag{3}$$

We compute this measure for all words appearing in our four text bodies–referrals, infringements, annulments and appeals–to mark off the set of words most distinctive of each

[11] Compared to topic modelling, corpus comparison is a less developed research field. For a survey of the literature see Kilgarriff (2001) and Remus and Bank (2012).

collection of decisions.

While (3) should normally provide a reliable indication of the issues on which the corpora diverge most, it may be distorted by the high incidence of rare words in a small subset of documents. So, as alternative measure of corpus dissimilarity, we draw repeated, equal-sized random samples from each document collection and conduct χ^2 and Mann-Whitney U tests on pairwise sample comparisons (Kilgarriff, 2001). We use the value of these statistics averaged over pairwise random samples to identify the most distinctive terms.

4 Results

We used R scripts to scrap the text of all published ECJ decisions, up to 31 December 2015, from the EUR-Lex website.[12] The scraped documents add up to 11 725 rulings spanning hundreds of thousands of pages of text. Figure 2 illustrates their distribution across the three procedures through time. Because references submitted by national courts may be withdrawn or formally dismissed, they do not always result in a preliminary ruling. While it might, in principle, be interesting to explore issue attention in references and possible discrepancies with preliminary rulings, existing EU law databases often fail to provide the text of references, whether or not these have resulted in a preliminary ruling. For that reason, our analysis ignores references and considers only preliminary rulings. Although there are fewer preliminary rulings than references, preliminary rulings still far outweigh annulment and infringement decisions in our aggregate corpus.

[12] More recent years were still incomplete at the time of writing.

The raw documents were subject to the conventional pre-processing steps of removing punctuation, numbers, html tags, functional words and words appearing in less than five decisions before converting each corpus into a document-term matrix. To prevent our computer-assisted document classification from being influenced by case report annotations, we also removed the issue codes from the judgements. As indicated above, for the between-corpus comparison we removed the procedure-distinctive procedural terms, as these would otherwise automatically come out as most distinctive of each procedure.[13]

To set the number of topics, we relied primarily on interpretability, although we also report the results of a topic model where we select K on the basis of perplexity. We found that $K = 15$ resulted in easily interpretable models for infringement and annulment rulings, which form smaller corpora. For preliminary rulings, which form a much larger corpus, $K = 25$ was found to strike a good balance between interpretability and specificity. The topics generated by the topic models were labelled manually by the research team. Labels were chosen after looking both at the words most characteristic of the topic and at the decision displaying the highest prevalence of the topic according to the corresponding model.

4.1 Issue Prevalence Within and Between Procedures: Static Analysis

Shown in Figure 3 are the unlabelled topics of a $K = 25$ LDA model of preliminary rulings. For each topic the ten words with the highest β value are reported. Considering

[13] The complete list is report in the Supplementary Materials.

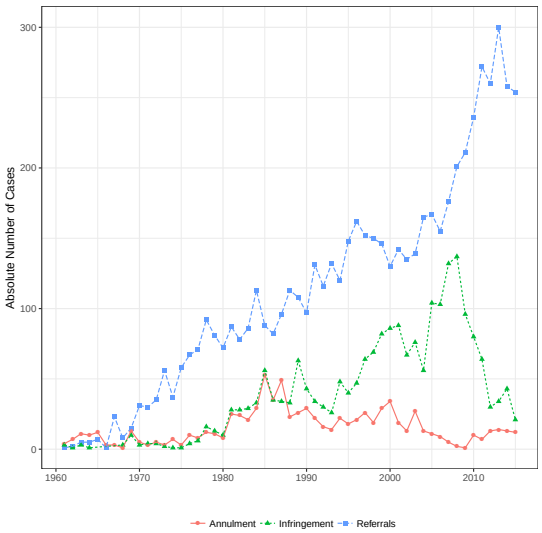


Figure 2: ECJ decisions by procedure 1961-2015. Annulment decisions include appeals against General Court rulings.

only these words, most of the topic appear readily interpretable to anyone familiar with EU law. Intuitively, with words such as "television", "copyright" and "communication", topic 1 seems to be about intellectual property in the context of television broadcasting. Topic 9 appears to be about public procurement; topic 11 about criminal matters and topic 22 about social security. Inspecting the opinion with the highest proportion of the topic–the highest θ_k–can help in interpreting the terms that define it. As an illustration, consider topic 14. The decision with the highest θ_{14} value (0.95) turns out to be *Test Claimants in the FII Group Litigation v Commissioner of Inland Revenue*, a 2012 Grand Chamber ruling on a reference from the High Court of Justice of England and Wales. We quote the first paragraphs of the facts section of the decision:

> The High Court of Justice of England and Wales, Chancery Division, seeks, first, to obtain clarification regarding paragraph 56 of the judgment in Test Claimants in the FII Group Litigation and point 1 of its operative part. It

> recalls that the Court of Justice held, in paragraphs 48 to 53, 57 and 60 of that judgment, that national legislation which applies the exemption method to nationally-sourced dividends and the imputation method to foreign-sourced dividends is not contrary to Articles 49 TFEU and 63 TFEU, provided that the tax rate applied to foreign-sourced dividends is not higher than the rate applied to nationally-sourced dividends and that the tax credit is at least equal to the amount paid in the Member State of the company making the distribution, up to the limit of the tax charged in the Member State of the company receiving the dividends.[14]

The decision is clearly about corporate taxation in the context of the internal market, which is what the terms characterising topic 14 suggest too.

Labelled topics can be visualized as a network, as depicted in Figure 4, where node size is proportionate to the issue's overall prevalence in the corpus (measured as $\frac{\sum_{d=1} \theta_{d,k}}{D}$) and edge thickness denotes the degree of correlation among topics (as calculated from their β distributions over the vocabulary). Visualizing a topic model this way facilitates the identification of issues clustering around broader themes. In Figure 4 policy, customs, taxation and labour market regulations come out as the most prevalent issues. On closer examination, though, two distinct clusters of issues become apparent. In the lower left part of the network are core internal market issues: labelling, goods, import, export, trademark, customs, tariffs, corporate taxation and services. In the right region are the welfare, law and order and administrative law issues: labour, social security, residence and

[14] Judgment of 13 November 2012, C-35/11, para. 21.

immigration, criminal matters, public procurement and recognition of foreign judgements. These clusters are also recognisable when the Court's rulings are represented by a more fine-grained model. Figure 5 depicts a model with $K = 97$ topics–which is the model that minimizes perplexity.[15] Here we find core internal market issues in the lower part of the network, whereas welfare issues are clustered in the left region (with social policy, pension, labour, working time, citizenship, refugee, recognition of foreign judgments, education and parental leave). The saliency of the internal market theme dovetails well with the notion that "dull tax cases, consumer protection actions, common customs tariff classification disputes, trans-border enforcement of small civil claims, companies' shareholders quarrels and so on" represent a large chunk of the ECJ's caseload (Bobek, 2013a). Yet preliminary rulings are about more than just market regulations. As both Figure 4 and 5 illustrate, welfare benefits and migrants' and workers' rights form another important theme (Caporaso and Tarrow, 2009; Conant, 2006).

[15] Minimizing perplexity means that the optimal topic model is the one that best predicts the content of held-out documents.

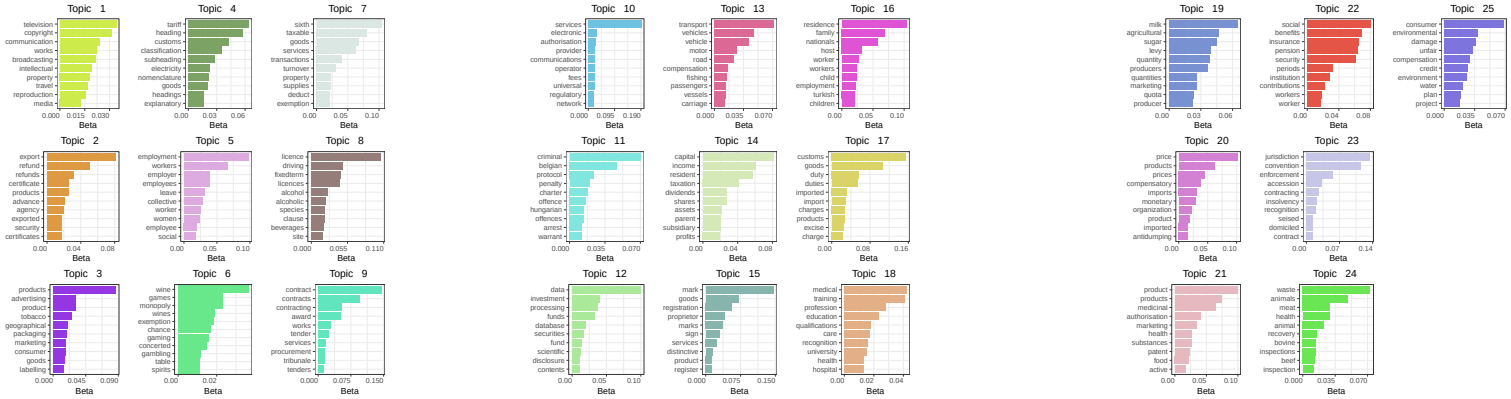


Figure 3: Unlabelled topics from a $K = 25$ topic model (latent Dirichlet allocation) fit to preliminary rulings. Topics are represented by their ten most specific terms. Terms with larger β values are more characteristic of the topic.

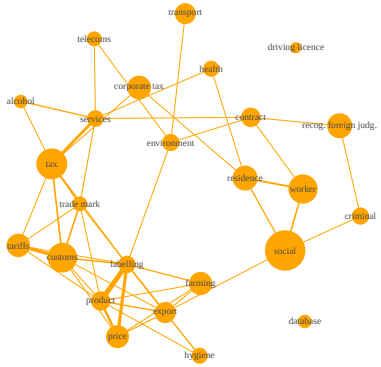


Figure 4: Topic model from Figure 3 with labelled topics represented as a network. Node size denotes topic prevalence across the entire set of documents. Edge thickness denotes correlation among topics.

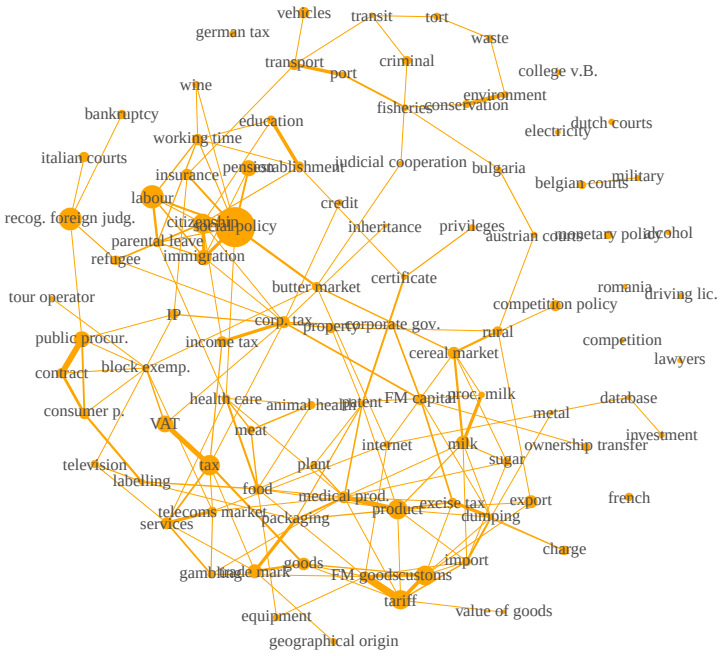


Figure 5: Topic model (LDA) with 97 labelled topics represented as a network. Number of topics is chosen so as to minimize perplexity. Node size denotes overall topic prevalence across. Edge thickness denotes correlation among topics.

How do preliminary rulings compare to the other procedures? Figure 6 illustrates topic models of infringement and annulment proceedings. The infringement model features one topic, "monarchies", which simply captures language in cases brought by or against member states with a constitutional monarchy.[16] This one Overall though, the

[16]The case with the highest $\theta_{monarchies}$ is *Kingdom of Spain v United Kingdom of Great Britain and Northern Ireland* over Gibraltar, one of the rare infringement actions initiated by a national government. Unsurprisingly, less than a handful of decisions have a $\theta_{monarchies}$ larger than 0.5.

topics are readily interpretable and meaningful. Internal market and environmental topics clearly dominate proceedings. Internal market issues include free movement of goods, workers, services and capital along with tariffs, competition and public procurement. Issues pertaining to the harmonization of hygiene regulations and the liberalization of the telecoms and passenger rail market are part of the same internal market theme. Environmental issues encompass conservation and waste water treatment. The saliency of the internal market theme is consistent with previous research linking the rise in infringement actions in the 1980s and 1990s to the European Commission's internal market agenda (Tallberg, 1999). That environmental and internal market cases account for the bulk of infringement cases has also been suggested by legal scholars (see Craig and Burca, 2015, 430). Compared to referral proceedings, infringement cases appear more cohesive and less diverse. Striking is the low attention to welfare, residence and social rights, which are relatively prominent in preliminary rulings. This appears consonant with the view that the referral procedure affords individuals, particularly the economically and politically disadvantaged, greater influence over the European Court's agenda (Alter, 2006; Keohane et al., 2000; Caporaso and Tarrow, 2009; Cichowski, 2007; Conant, 2006). Nor do infringement actions show much attention for law and order issues or the recognition of foreign judgments and driving licences. The greater emphasis on environmental matters, though, suggests an interesting exception. It has been argued that in many EU member states the combination of weak civil society with rules of standing restricting access to domestic courts to environmental NGOs inhibit private enforcement of EU environmental regulations, leaving infringement actions as the sole viable avenue to address

endemic non-compliance (Börzel, 2006; Vanhala, 2018). On core internal market issues, on the other hand, referral and infringement cases show a significant overlap. Here at least, private enforcement may be an effective substitute or complement for prosecution by Commission officials (Kelemen, 2012; Börzel, 2006; Tallberg, 1999). Taxation, too, is a domain where the priorities of the Commission and the interests of private litigants largely overlap (Genschel and Jachtenfuchs, 2011). A plausible explanation is that private businesses and wealthy individuals have strong incentives to invoke EU law to reduce their tax burden (Genschel and Jachtenfuchs, 2011, 303) while the Commission sees tax harmonisation as a building block of the internal market (Genschel et al., 2011; Radaelli and Kraemer, 2008).

Annulment cases (Figure 6b) address various EU policies: anti-dumping measures, the regulation of fisheries, the common agricultural policy (CAP), steel production (under the European Coal and Steel Community), the internal market (tariffs and hazardous substances) and the common foreign and security policy (CFSP). Several topics capture aspects of the EU competition policy: the application of competition rules to distribution systems and parent companies and the fines imposed by the European Commission, whose magnitude parties often challenge. As with the referral procedure, greater inclusiveness is associated with greater issue heterogeneity.

Corpus-level analysis confirms the asymmetry in issue attention between referral and infringement proceedings regarding welfare and environmental protection. Figure 7 displays a comparison wordcloud, where the size of the words is proportionate to procedural distinctiveness. We see that the terms “employ”, “benefits”, “pension” and “social” occur

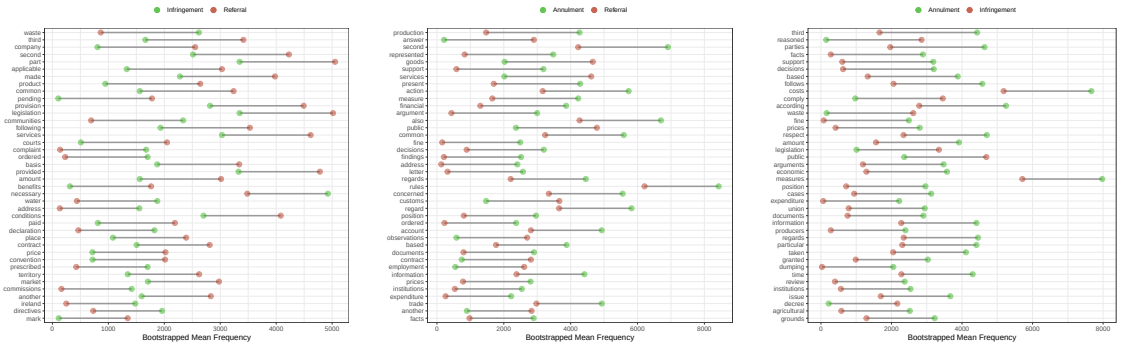
disproportionately more in preliminary rulings whereas "environment", "waste", "water", "pollution", "protect" and "bird" are more characteristic of infringement decisions. Member states are often dragged before the Court of Justice for failing to transpose EU directives adequately (Tallberg, 1999; Craig and Burca, 2015, 444). This is attested by the words "transpose", "transposition" and "implementation". Figure 8 shows the most dissimilar words in pairwise comparison of bootstrapped samples of the three corpora.Here again, we see that "waste", "water" and "directives" are more characteristic of infringement proceedings whereas "benefits" and "employment" are more distinctive of preliminary rulings. Attesting to the large number of competition cases brought under the annulment procedure is the word "fine" (compared to both infringement and preliminary rulings). The difference in the average frequency of the word "agriculture" in panel 8c confirms that the common agricultural policy is more frequently litigated under this procedure. Both the χ^2 and Mann-Whitney U test point to significant pairwise corpus dissimilarities. Averaged over the 10000 bootstrapped samples, p-values for these statistics are $p < 0.001$ in all three pairwise comparisons.

(a) Infringement procedure.

(b) Annulment procedure.

Figure 6: Topic models (latent Dirichlet allocation) with labelled topics for infringement and annulment. Node size denotes overall topic prevalence across the entire set of decisions. Edge thickness denotes correlation among topics.

Figure 7: Corpus dissimilarity: the plot illustrates the words most distinctive of each procedure.



(a) (b) (c)

Figure 8: Corpus dissimilarity: plot shows the most distinctive terms in pairwise comparisons of the procedures. Words are sorted by bootstrapped mean difference.

Figure 9 shows topic proportion across seven categories of plaintiffs in annulment proceedings. By showing how issue emphasis varies across litigant types, it lends additional support to the inclusiveness-diversity hypothesis. National governments mostly bring cases pertaining to the agricultural guarantee fund, state subsidies to domestic industries, border crossing and oversea territories. Private litigants use the annulment procedure to challenge EU regulatory decisions on competition, dumping and trademarks. Until the ECSC was absorbed into the European Union, private litigants were also more likely to bring cases in connection with the regulation of coal and steel industries. EU institutions tend to litigate the same issues. Yet the European Parliament is disproportionately likely to litigate border crossing issues while the cases in which the European Commission is the plaintiff are more likely to be about state aid. EU civil servants, meanwhile, mostly bring employment disputes.

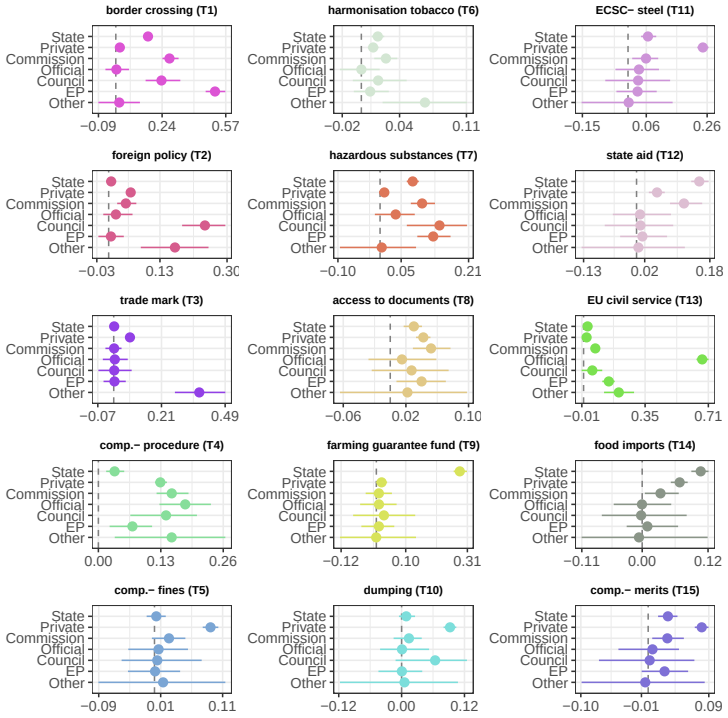


Figure 9: Author-topic model of annulment cases ($K = 15$). Plot shows topic proportion for seven categories of case initiators. Horizontal bar denotes 95 per cent confidence interval.

4.2 Dynamic Analysis

The foregoing textual analysis is static in that it models ECJ decisions without consideration of the date at which they were issued. The number of ECJ rulings has grown dramatically since the inception of the integration process in the 1960s. In consequence, later years are associated with larger sets of decisions. Hence the models presented furnish a more accurate picture of issue attention for the 2000s than for the 1970s. This should not necessarily be viewed as a problem. A static model focusing on more recent cases may reveal interesting insights. The mutual recognition of driving licences–topic 8 in Figure

3–is an issue that seems to have landed only recently on the Court's docket (see Janssens, 2013, 101) but which few judicial scholars may spontaneously associate with the ECJ. In that sense, different models simply constitute different ways to look at the Court's case law. Note, however, that none of the topics in Figure 3 directly relates to the doctrines and principles of supremacy, direct effect, judicial empowerment and state liability, which have been central to the constitutionalisation of EU law (Sweet and Brunell, 1998; Weiler, 1991). Because these questions received more attention in the early phase of the integration process, when ECJ rulings were few and far between, a plain LDA model is unlikely to detect them. For this, we must turn to dynamic topic modelling.

Figure 10 illustrates a dynamic topic model of preliminary rulings. The topic labelled "Constitutional" is defined by terms that include "community" (as in "Community law"), "provisions" and "effect" (as in "direct effect"). Most characteristic of this topic is the decision *Eunomia di Porro v Ministry of Education of the Italian Republic* of 26 October 1971. We quote the grounds of judgment section of the ruling in full:

> By decision of 6 April 1971, received at the Court registry on 15 April 1971, the President of the Tribunale di Torino referred to the Court under Article 177 of the Treaty Establishing the European Economic Community two questions on the interpretation of Article 16 of the EEC Treaty.
>
> The decision making the order for reference shows that the national court is dealing with a request for the refund of sums paid on the export of a work of art to another member state by way of the tax on the export of articles of an artistic, historic, archeological or ehtnographic interest, which was introduced

by the Italian Law No. 1089 of 1 June 1939.

As the Court of Justice found in its judgment 10 December 1968 in 7/68, this tax constitutes a charge having en effect equivalent to customs duties on exports and is governed by Article 16 of the Treaty.

In the first question the Court is asked to rule whether Article 16 constitutes a legal rule which is immediately applicable and which produces direct effects within the territory of the Italian State as from 1 January 1962. Should the answer to the first question be in the affirmative the Court is requested to rule whether, as from that date, this rule has created individual rights in relation to the Italian State which the courts must protect. As these two questions are closely connected they must considered together.

According to Article 9 of the EEC Treaty, the Community is to be based upon a customs union which is to involve in particular the prohibition between Member States of customs duties and all charges having equivalent effect. Under Article 16 of the Treaty Member States are to abolish between themselves customs duties on exports and charges having equivalent effect by the end of the first stage at the latest.

Articles 9 and 16 taken together involve, at the latest at the end of the first stage, with regard to all charges having an effect equivalent to customs duties on exports, a clear and precise prohibition on exacting the said charges, which is not subject to any reservation for the states to subject its implementation to a positive act of national law or to an intervention by the institutions of the

> Community. It lends itself, by its very nature, to producing direct effects in the legal relations between Member States and those subject to their jurisdiction.
>
> Therefore, from the end of the first stage, that is, from 1 January 1962, these provisions have conferred on individuals rights which the national courts must protect and which must prevail over conflicting provisions of national law even if the member state has delayed in repealing such provisions.

Though the decision is hardly a landmark ruling, it addresses the direct effect of Article 16 of the Rome Treaty. As with rulings addressing constitutional issues in general, though, it is not entirely given over to the constitutional topic. Rather, the discussion of direct effect is a prelude to the resolution of the more substantive question, which is whether a tax on the export of artworks constitutes a non-tariff restriction on trade. Accordingly, the model classifies the ruling as 70 per cent constitutional ($\theta_{constitutional} = 0.7$), twelve per cent about trade in goods ($\theta_{FM\,goods} = 0.12$) ten percent about non-tariff barriers ($\theta_{non\,tariff\,barriers} = 0.1$).[17] The classification of other landmark constitutional rulings also makes intuitive sense. *Simmenthal*, a case arising from a dispute over meat imports in which the ECJ held that every domestic court had the power to set aside domestic legislation contrary to EU law, is classified as 63 per cent constitutional. *Costa v ENEL*, in which the ECJ first spelled out the principle of supremacy of EU law, is 48 per cent constitutional. *Factortame I*, the first time an act of the British Parliament was declared contrary to EU law, shows a similar topic proportion (47 percent). So too does *Van Gend en Loos*, the 1963 ruling that gave birth to the doctrine of direct effect (43 per

[17] Five topics account for the remaining 10 per cent.

cent constitutional). At 30 per cent the topic proportion for *Francovich*, in which the ECJ established the principle of state liability for failure to transpose EU directives, is somewhat lower, although still plausible.

Figure 10 indicates that, relative to other topics, attention to constitutional questions has declined dramatically over time. Constitutional goes from most important topic in the 1960s to marginal one after the 1980s. This evolution is in line with the dominant narratives of the European integration process in EU studies. These narratives describe the 1960s and early 1970s as the "foundational" period during which the ECJ "constitutionalised" EU law (Weiler, 1991; Sweet and Brunell, 1998). The foundational phase laid the ground for the subsequent expansion of EU law litigation to new domains (Alter, 2001; Sweet and Brunell, 1998).

Next to constitutional principles, trade, particularly non-tariff barriers, represented the most important item on the Court's agenda in the early period. Then, as the scope of EU treaties and legislation expanded, so too did the Court's agenda. Trade and non-tariff barriers became less salient as new issues grew in importance: corporate taxation, trademarks, consumer protection, citizenship, environmental protection, labour regulation, telecommunications, patents, intellectual property, establishment freedom (notably the mutual recognition of professional certifications), asylum, EU funding, public procurement and VAT. This accords with the spill-over dynamics hypothesized by neo-functionalist accounts of European integration (Sweet and Brunell, 1998; Stone Sweet and Sandholtz, 1997). Surprisingly, attention to social policy questions (benefits, pensions) appear to have declined over time, although some questions ordinarily falling under the

social policy heading may be captured by topic 12 (labour).[18]

Compared to referrals, issue prevalence in infringement cases looks more stable through time (Figure 11), as we expected from the procedure lower degree of inclusiveness. Save for the taxation of spirits and agriculture and forestry, change in issue attention has been less abrupt and more gradual. Several topics, though, show short-term blips in issue attention. These may result from the Commission going after multiple member states at the same point in time after the deadline for transposition of EU legislation has expired.

[18] This intuition is supported by our comparison with issue codes from the official case report in Figure 14b.

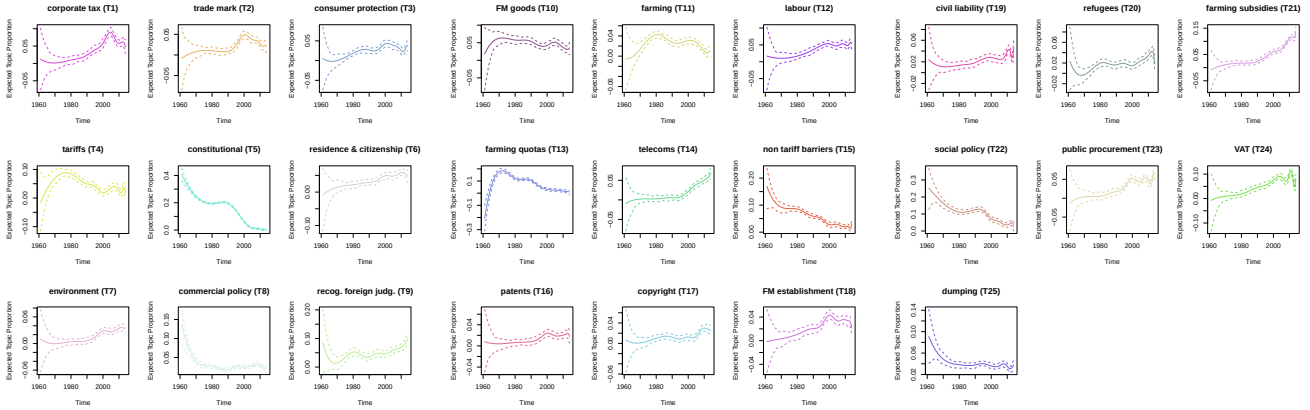


Figure 10: Topic proportion by year from dynamic topic model ($K = 25$) of preliminary rulings. Dashed lines show 95 per cent confidence interval.

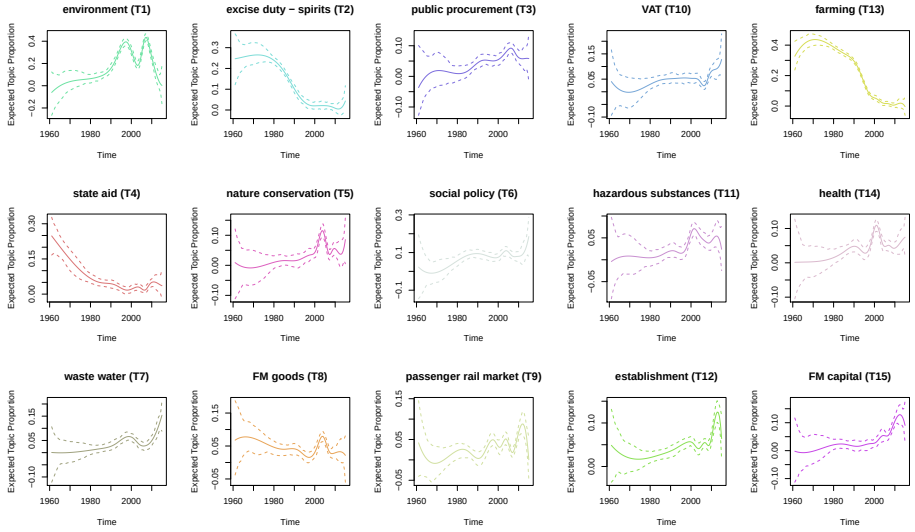


Figure 11: Topic proportion by year from dynamic topic model ($K = 15$) of infringement decisions.

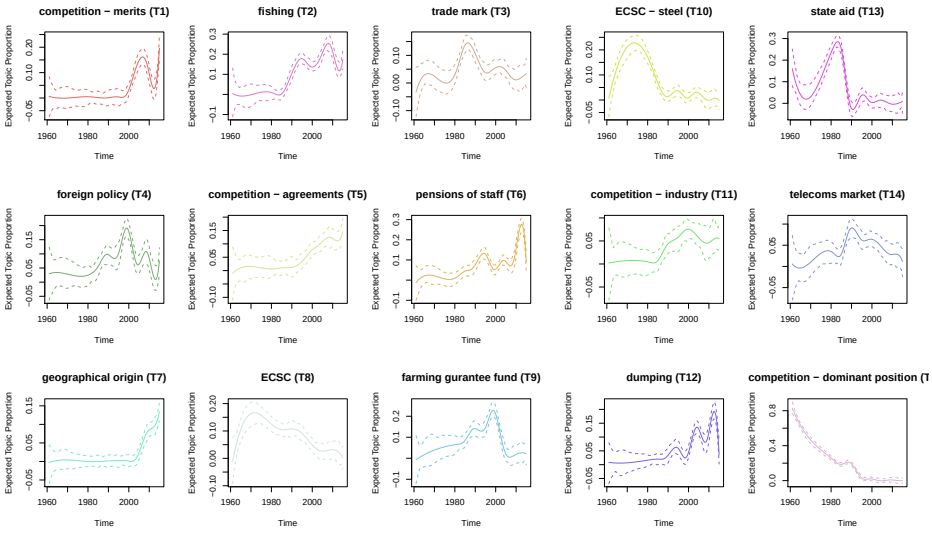


Figure 12: Topic proportion by year from dynamic topic model ($K = 15$) of annulment decisions.

It is for annulment cases that topic proportion shows the greatest temporal variation, as illustrated in Figure 12. All fifteen topics show sharp variations, whether it is trade mark, foreign policy, dumping or competition-related issues. This, again, supports the hypothesis that broader access rules and greater litigant heterogeneity have the effect of lowering issue stability.

The results reported in this Section suggest that, in addition to occasionally highlighting unexpected themes and patterns, computer-assisted textual analysis methods offer a richer, more precise and more comprehensive picture of judicial issue attention than statements inspired by fragmented data (e.g. Craig and Burca, 2015, 430), official case annotations (Schmidt, 2012; Sweet and Brunell, 1998) or mere impressions (Bobek, 2013b). In the next two sections, we demonstrate that topic models produce valid, meaningful case categorizations and that computerised case classification is a strong alternative over classification schemes based on human annotations.

5 Validation: Family Reunion Cases

The methods applied in the present study rest on simplifying assumptions about the data-generating process. Specifically, they rely on the bag-of-words approach, which means that complex rulings are reduced to vectors of word counts without consideration for syntax and context. In consequence, validation is recommended (Grimmer and Stewart, 2013). To that end, we relate two of our topic models to manually collected decisions on family reunion issues. This set of decisions was assembled in the context of a longitudinal study of the ECJ's case law on family reunion (De Somer, 2019). Cases included in the dataset are cases that (1) explicitly refer to EU legislation on family reunion, migrant worker, asylum or citizenship, and (2) involve at least one third-country national.[19] In total 67 cases, covering four decades, were selected. Of the 67 decisions, 58 are preliminary rulings, eight are infringement and one is an annulment action.

Generally speaking, if we consider only preliminary rulings, family reunion cases should be strongly correlated with topic 16 in our static topic model (Figure 3) and topic 6 (residence & citizenship) from our dynamic topic model of preliminary rulings (Figure 10). Moreover, we should expect decisions that are more central to the Court's family reunion jurisprudence to exhibit both a higher θ_{16} (static topic model) and a higher θ_{T6} (dynamic topic model). To measure the degree of embeddedness of decisions in the Court's case law, we use cross-citations to construct an indicator of jurisprudential centrality. Our indicator is simply the sum of in-degree and out-degree citations. It reflects the intuition that decisions which either cite more decisions or are more cited by other decisions in the

[19] We refer to De Somer (2019) for detail of the operationalisation of these two criteria.

citation network (or both) are likely to share more doctrinal language. We hypothesize that more embedded decisions will exhibit a higher θ for the relevant topic. Figure 13 illustrates the network of case citations with larger nodes denoting greater jurisprudential embeddedness.

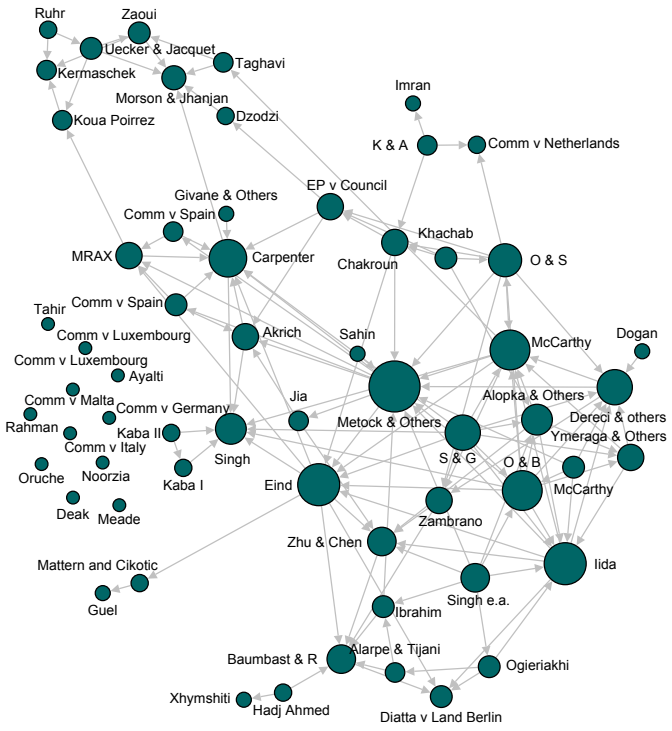
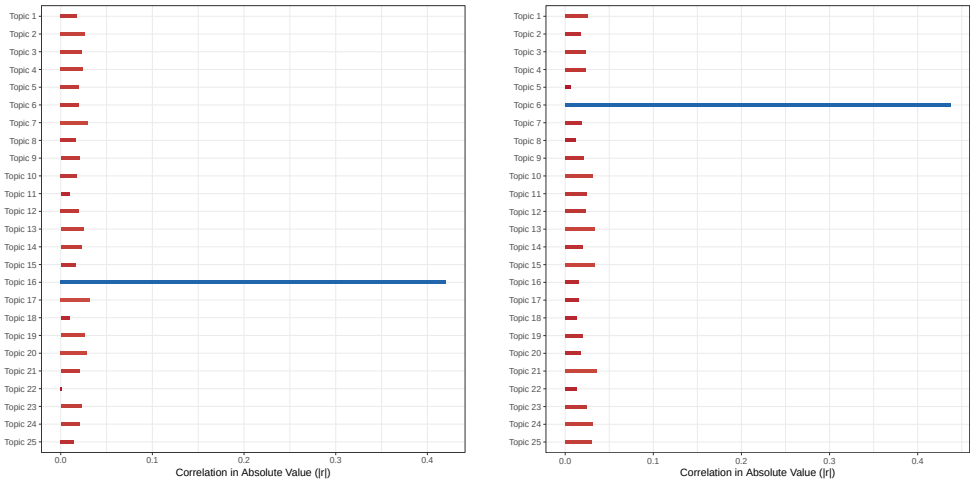


Figure 13: Case citation network of manually collected family reunion cases.

As shown in Figure 14, manually collected family reunion cases are strongly related to the expected topics–namely topic 16 (residence) for the static LDA model and topic 16 (residence and citizenship) for the dynamic model. This is strong evidence that the family reunion cases are correctly classified by our topic models.



(a) (b)

Figure 14: Correlation between manually collected family reunion cases and topics from static (14a) and dynamic (14b) topic models with loess curve. Topic 16 in Figure 14a we labelled as "residence" (see Figure 3 and 4). Topic 6 in Figure is labelled as "residence & citizenship" (see Figure 10). Only preliminary rulings are considered.

Next, Figure 15 depicts the relationship between θ_{16} (from LDA model) and our measure of jurisprudential embeddedness. As expected, higher jurisprudential centrality correlates with higher topic proportion. Cases that cite more or get cited more (or both) show a greater proportion of topic 16. Since case citation is driven by law-finding and law creation rather than by fact-finding, this suggests that topic 16 is capturing legal doctrines as well as factual regularities. Interestingly, the fitted loess curve in Figure 15 follows a convex shape. This implies that jurisprudential centrality is associated with θ_{16} up to a certain level, presumably because the topic also captures factual regularities.

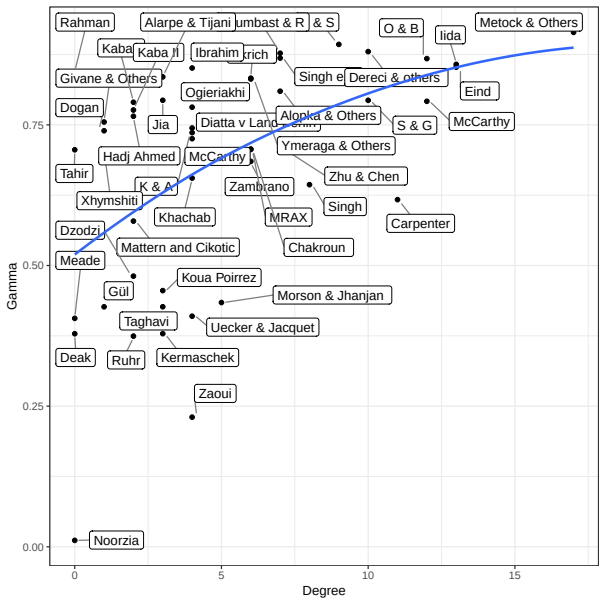


Figure 15: Correlation between θ_{16} from static LDA model (Figure 3) and jurisprudential embeddedness of manually collected family reunion cases.

6 Official Case Annotations vs Computerised Classification

We now compare our automated case classification with subject matter codes from CURIA, the official case law database of the European Court of Justice. The database provides little in the way of documentation. No description of the annotation procedure is provided. Nor are the subject matter categories defined, although the database suggests that they correspond to the legal bases covered by the rulings. In spite of these limitations,though, these case annotations are a popular tool for case law search, among jurists (Ravasi, 2017) as well as social scientists (Schmidt, 2012; Sweet and Brunell, 1998;

Stone Sweet, 2004; Kelemen and Pavone, 2016).[20]

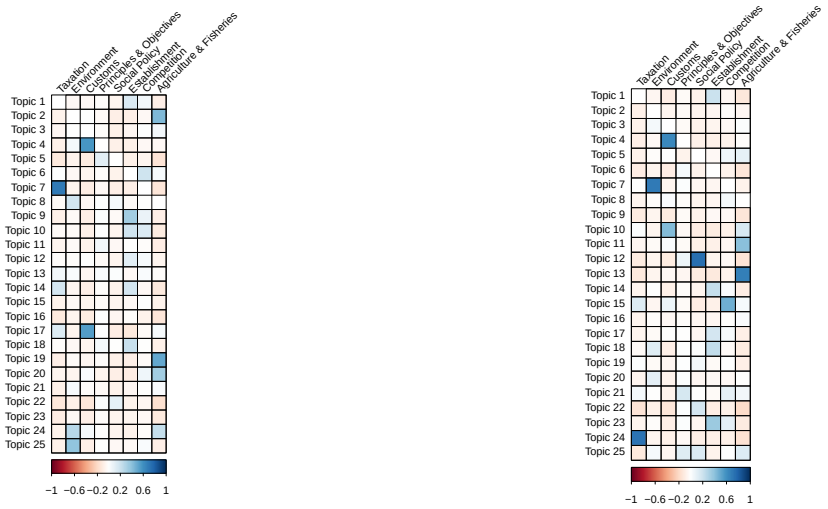
We should expect our computer-aided classification to relate to a significant degree to CURIA case tags. However, we should not expect a perfect match. First, whereas manual classification schemes such as the one under consideration have been applied to cases that did not exist at the time they were developed, our topic models leverage information from all the decisions to organize the corpus. Second, as mentioned in the introductory section, we have reasons to suspect that the Court's internal annotation process does not obey rigorous rules. Moreover, we should expect the correlation to vary depending on the case tag as well as the topic model. For example, one CURIA tag is "Principles, objectives and tasks of the treaties", which we take to denote constitutional issues. As no topic in our static topic model of preliminary rulings (Figure 3) really captures constitutional questions, we should not expect a high correlation between this annotation and any of the 25 topics. Conversely, we should expect the case tag "Customs" to be strongly correlated with topic 4.

Figure 16 shows the correlation between a range of frequent CURIA tags and the 25 topics from both our simple static LDA model and our dynamic topic model of preliminary rulings. The pattern of correlation coefficients is, for the most part, in line with our expectations. "Taxation" strongly correlates with topic 7 (tax) and, to a lesser extent, with topic 14 (capital) in the simple LDA while it strongly correlates with topic 24 (VAT) in the dynamic topic model. "Customs" is strongly correlated with topic 4 (tariffs) and 17

[20] Parallel to the CURIA subject matter codes, the Court has developed a "systematic classification scheme", which is far more detailed. However, the original classification scheme was discontinued following the entry into force of the Lisbon Treaty and replaced by a new classification scheme so that case law prior to 2010 is classified according to one scheme and case law after that date according to another.

(customs) in the LDA model as well as with topic 4 (tariffs) and, albeit to a lesser extent, with topic 10 (free movement of goods) in the dynamic model. Likewise, "Environment" correlates with topic 24 (hygiene) and 25 (environment) in the static topic model and topic 7 (environment) in the dynamic model.

Interestingly, "Principles and Objectives" is not correlated with any of the topics. This is especially surprising for the dynamic model as we expected that tag to relate to topic 5 (constitutional). This could possibly reflect a a lack of rigour on the part of human case annotators. However, even if the Court's internal annotation process is error-prone, we would not expect zero correlation. So a better explanation is that because constitutional issues are usually secondary to the resolution of more substantive questions human case annotators tend to classify cases chiefly on the basis of what appears to be the main substantive issue. The upshot is that the constitutional dimension of cases is often neglected. This is an illustration of text-mining methods outperforming a popular case classification scheme.



(a) Static

(b) Dynamic

Figure 16: Correlation between issue codes from official case report and topics from static (panel 16a) and dynamic (panel 16b) topic models. Plots show how official case report annotations correlate with topic proportion.

7 Conclusion

We showed how natural language processing techniques can be applied to investigate issue attention in a large corpus of ECJ decisions. Topic modelling and corpus comparison tools can be used to represent topic, themes and trends across procedures, time and litigants. They produce maps of issue attention that are rich, precise and comprehensive. In the ECJ context, they help cast a wider light on the link between procedural inclusiveness and judicial agenda-setting. That who has standing to initiate legal action matters does not come as a surprise. But it is a proposition that text-mining methods make easier to document. Our validation exercise demonstrates that computer-based case classification produces meaningful, reliable and legally relevant categories. Finally, comparison with a popular case annotation scheme suggests that text-mining methods represent a reliable

and powerful alternative for the purpose of investigating issue attention in legal texts.

References

Adam, C., M. W. Bauer, and M. Hartlapp (2015, March). It's Not Always about Winning: Domestic Politics and Legal Success in EU Annulment Litigation. *JCMS: Journal of Common Market Studies 53*(2), 185–200.

Aletras, N., D. Tsarapatsanis, D. Preoţiuc-Pietro, and V. Lampos (2016). Predicting judicial decisions of the european court of human rights: A natural language processing perspective. *PeerJ Computer Science 2*, e93.

Alter, K. J. (2001). *Establishing the supremacy of European law: The making of an international rule of law in Europe*. Oxford University Press Oxford.

Alter, K. J. (2006). Private litigants and the new international courts. *Comparative Political Studies 39*(1), 22–49.

Alter, K. J. (2012). The global spread of european style international courts. *West European Politics 35*(1), 135–154.

Arora, S., R. Ge, Y. Halpern, D. Mimno, A. Moitra, D. Sontag, Y. Wu, and M. Zhu (2013). A practical algorithm for topic modeling with provable guarantees. In *International Conference on Machine Learning*, pp. 280–288.

Baird, V. and T. Jacobi (2009). Judicial agenda setting through signaling and strategic litigant responses. *Wash. UJL & Pol'y 29*, 215.

Baird, V. A. (2004). The effect of politically salient decisions on the us supreme court's agenda. *The Journal of Politics 66*(3), 755–772.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM 55*(4), 77–84.

Blei, D. M., J. D. Lafferty, et al. (2007). A correlated topic model of science. *The Annals of Applied Statistics 1*(1), 17–35.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of machine Learning research 3*(Jan), 993–1022.

Bobek, M. (2013a). Of feasibility and silent elephants: The legitimacy of the court of justice in the eyes of national courts. In M. Adams, H. De Waele, J. Meeusen, and G. Straetman (Eds.), *Judging Europe's Judges: The Legitimacy of the Case Law of the European Court of Justice*, pp. 197. Hart Publishing.

Bobek, M. (2013b). Of feasibility and silent elephants: The legitimacy of the court of justice through the eyes of national courts. In M. Adams, H. de Waele, J. Meeusen, and G. Straetmans (Eds.), *Judging Europe's judges: The legitimacy of the case law of the European Court of Justice*, pp. 197. Hart Publishing.

Börzel, T. A. (2003). Guarding the treaty: the compliance strategies of the european commission. In T. A. Börzel and R. A. Cichowski (Eds.), *The state of the European Union*, Volume 6, pp. 197–220. Oxford University Press.

Börzel, T. A. (2006). Participation through law enforcement: the case of the european union. *Comparative Political Studies 39*(1), 128–152.

Caldeira, G. A., J. R. Wright, and C. J. W. Zorn (1999, October). Sophisticated voting and gate-keeping in the supreme court. *JLEO 15*(3), 549–572.

Cameron, C. M. and L. A. Kornhauser (2013). Bargaining on Appellate Courts.

Caporaso, J. A. and S. Tarrow (2009). Polanyi in brussels: Supranational institutions and the transnational embedding of markets. *International Organization 63*(4), 593–620.

Carlson, K., M. A. Livermore, and D. Rockmore (2015). A quantitative analysis of writing style on the us supreme court. *Wash. UL Rev. 93*, 1461.

Carter, D. J., J. Brown, and A. Rahmani (2016). Reading the high court at a distance: Topic modelling the legal subject matter and judicial activity of the high court of australia, 1903-2015. *UNSWLJ 39*, 1300.

Chang, J., S. Gerrish, C. Wang, J. L. Boyd-Graber, and D. M. Blei (2009). Reading tea leaves: How humans interpret topic models. In *Advances in neural information processing systems*, pp. 288–296.

Cichowski, R. (2007). *The European Court and Civil Society: Litigation, Mobilization and Governance*. Cambridge University Press.

Conant, L. (2006). Individuals, courts, and the development of european social rights. *Comparative Political Studies 39*(1), 76–100.

Corley, P. C. (2008, September). The Supreme Court and Opinion Content The Influence of Parties' Briefs. *Political Research Quarterly 61*(3), 468–478.

Corley, P. C., P. M. J. Collins, and B. Calvin (2011). Lower Court Influence on U.S. Supreme Court Opinion Content. *The Journal of Politics 73*(01), 31–44.

Craig, P. and G. d. Burca (2015, November). *EU Law: Text, Cases, and Materials* (5 edition ed.). Oxford ; New York: Oxford University Press.

De Somer, M. (2019). *The Court's Role in Processes of European Integration*. Springer.

Edelman, P. H. and J. Chen (2007). The most dangerous justice rides into the sunset. *Const. Comment. 24*, 199.

Epstein, L., J. A. Segal, and T. Johnson (1996). The Claim of Issue Creation on the US Supreme Court. *American Political Science Review*, 845–852.

Evans, M., W. McIntosh, J. Lin, and C. Cates (2007). Recounting the Courts? Applying Automated Content Analysis to Enhance Empirical Legal Research. *Journal of Empirical Legal Studies 4*(4), 1007–1039.

Genschel, P. and M. Jachtenfuchs (2011). How the european union constrains the state: Multilevel governance of taxation. *European Journal of Political Research 50*(3), 293–314.

Genschel, P., A. Kemmerling, and E. Seils (2011). Accelerating downhill: How the eu shapes corporate tax competition in the single market. *JCMS: Journal of Common Market Studies 49*(3), 585–606.

Grimmer, J. and B. M. Stewart (2013, January). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*.

Harlow, C. and R. Rawlings (2006). Accountability and law enforcement: The centralised eu infringement procedure'(2006). *European Law Review 31*, 447.

Harvey, A. and M. J. Woodruff (2011). Confirmation bias in the united states supreme court judicial database. *The Journal of Law, Economics, & Organization 29*(2), 414–460.

Huber, J. D. (1992). Restrictive Legislative Procedures in France and the United States. *The American Political Science Review 86*(3), 675–687. ArticleType: research-article / Full publication date: Sep., 1992 / Copyright © 1992 American Political Science Association.

Janssens, C. (2013). *The principle of mutual recognition in EU law*. OUP Oxford.

Kastellec, J. P. and J. R. Lax (2008). Case Selection and the Study of Judicial Politics. *Journal of Empirical Legal Studies 5*(3), 407–446.

Kelemen, D. R. and T. Pavone (2016). Mapping european law. *Journal of European Public Policy 23*(8), 1118–1138.

Kelemen, R. D. (2012). Eurolegalism and democracy. *JCMS: Journal of Common Market Studies 50*(s1), 55–71.

Keohane, R. O., A. Moravcsik, and A.-M. Slaughter (2000). Legalized Dispute Resolution: Interstate and Transnational. *International Organization 54*(03), 457–488.

Kilgarriff, A. (2001). Comparing corpora. *International journal of corpus linguistics 6*(1), 97–133.

Livermore, M. A., A. B. Riddell, and D. N. Rockmore (2017). The supreme court and the judicial genre. *Ariz. L. Rev. 59*, 837.

Lupu, Y. and E. Voeten (2012). Precedent in international courts: a network analysis of case citations by the European Court of Human Rights. *British Journal of Political Science 42*(2), 413–439.

Mimno, D., H. M. Wallach, E. Talley, M. Leenders, and A. McCallum (2011). Optimizing semantic coherence in topic models. In *Proceedings of the conference on empirical methods in natural language processing*, pp. 262–272. Association for Computational Linguistics.

Mochales, R. and M.-F. Moens (2011). Argumentation mining. *Artificial Intelligence and Law 19*(1), 1–22.

Perry, H. W. (2009). *Deciding to decide: agenda setting in the United States Supreme Court*. Harvard University Press.

Radaelli, C. M. and U. S. Kraemer (2008). Governance areas in eu direct tax policy. *JCMS: Journal of Common Market Studies 46*(2), 315–336.

Ravasi, E. (2017). *Human rights protection by the ECtHR and the ECJ: a comparative analysis in light of the equivalency doctrine*. Brill.

Remus, R. and M. Bank (2012). Textual characteristics of different-sized corpora. In *The 5th Workshop on Building and Using Comparable Corpora*, pp. 148.

Rice, D. (2014). The impact of supreme court activity on the judicial agenda. *Law & Society Review 48*(1), 63–90.

Roberts, M. E., B. M. Stewart, and E. M. Airoldi (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association 111*(515), 988–1003.

Rosen-Zvi, M., T. Griffiths, M. Steyvers, and P. Smyth (2004). The author-topic model for authors and documents. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, pp. 487–494. AUAI Press.

Scheppele, K. L. (2005). Democracy by judiciary: Or, why courts can be more democratic than parliaments. *Rethinking the rule of law after communism*, 25–60.

Schmidt, S. K. (2012). Who cares about nationality? the path-dependent case law of the ecj from goods to citizens. *Journal of European Public Policy 19*(1), 8–24.

Shapiro, C. (2008). Coding complexity: Bringing law to the empirical analysis of the supreme court. *Hastings LJ 60*, 477.

Smith, M. (2008). Enforcement, monitoring, verification, outsourcing: the decline and decline of the infringement process. *European Law Review 33*(6), 777–802.

Stone, A. (1992). *The birth of Judicial Politics in France: The Constitutional Council in Comparative Perspective*. Oxford University Press, USA.

Stone Sweet, A. (2004). *The Judicial Construction of Europe*. Oxford [u.a.]: Oxford University Press.

Stone Sweet, A. and W. Sandholtz (1997). European integration and supranational governance. *Journal of European Public Policy 4*(3), 297–317.

Sulea, O.-M., M. Zampieri, M. Vela, and J. van Genabith (2017). Predicting the law area and decisions of french supreme court cases. *arXiv preprint arXiv:1708.01681* .

Sweet, A. S. and T. L. Brunell (1998). Constructing a supranational constitution: Dispute resolution and governance in the european community. *American Political Science Review 92*(1), 63–81.

Tallberg, J. (1999). *Making states comply: the European Commission, the European Court of Justice & the enforcement of the internal market*, Volume 109. Lund University.

Vanhala, L. (2018). Is legal mobilization for the birds? legal opportunity structures and environmental nongovernmental organizations in the united kingdom, france, finland, and italy. *Comparative Political Studies 51*(3), 380–412.

Wallach, H. M., I. Murray, R. Salakhutdinov, and D. Mimno (2009). Evaluation methods for topic models. In *Proceedings of the 26th annual international conference on machine learning*, pp. 1105–1112. ACM.

Weiler, J. H. H. (1991). The Transformation of Europe. *Yale L.J.*, 2403–2483.