The Voices of European Law: Legislators, Judges and Law Professors

Arthur Dyevre* (corresponding author)
KU Leuven Centre for Empirical Jurisprudence
arthur.dyevre@kuleuven.be

Michal Ovádek
KU Leuven Centre for Empirical Jurisprudence

Monika Glavina
University of Antwerp

Abstract

EU legislators, CJEU judges and EU law scholars have produced streams of texts which determine both what EU law is and what it is perceived to be about. We explore what these distinct ‘voices’ tell us about the EU’s legal and policy priorities using a mega corpus compiling more than 200 000 legislative acts, 55 000 court rulings and opinions and 4000 articles from a leading EU law journal. Applying an unsupervised machine learning technique known as probabilistic topic modelling, we find that economic integration remains the focus of EU law but that scholars tend to emphasise rights issues more while ignoring certain topics, such as farming regulations, almost entirely. The relationship among these partly interdependent, partly autonomous voices, we suggest, can be conceptualised in terms of co-evolution. Legislation influences issue attention on the CJEU, which, in turn, influences what law professors choose to write about.

* Centre for Empirical Jurisprudence, KU Leuven. The authors acknowledge financial support from ERC Starting Grant 638154 (EUTHORITY).

1. Introduction

Anecdotal observation suggests that what legal scholars and law students know about EU law is principally shaped by what they read in textbooks, law review articles and, occasionally, the odd directive or regulation, when these happen to involve their main area of interest. Of course, nobody can read all the texts that EU legislators, judges and legal scholars produce. Over its lifetime, the EU has produced more than 200 000 legislative acts. The CJEU has issued more than 55 000 documents, while legal scholars have churned out tens of thousands of articles and case notes along with textbooks and commentaries. So, however voracious a reader, what one manages to read can only paint an incomplete and fragmented picture of the issues that EU law embraces and prioritises. Yet, despite the cognitive challenge posed by the sheer mass of texts, statements about the substantive focus of EU law or alleged inflections abound in the legal literature: the Maastricht Treaty marked an inflection from an economic ‘businessmen’s Europe’ to a ‘people’s Europe’¹; the Charter of Fundamental Rights is meant to have ushered in a more rights-centred vision of the European project²; the Lisbon Treaty is intended to have established a more robust and expansive conception of European citizenship,³ and so on. The importance of human rights in the jurisprudence of the CJEU has been the subject of conflicting assertions.⁴ Some authors have stressed the growing role of the CJEU ‘as a human rights adjudicator’⁵ that ‘has evolved from being a tribunal concerned primarily with economic matters to one with a much wider range of jurisdiction which is now explicitly tasked with

¹ Jennifer M Welsh, ‘A Peoples’ Europe? European Citizenship and European Identity’ (1993) 13 *Politics* 25; Richard C Eichenberg and Russell J Dalton, ‘Post-Maastricht Blues: The Transformation of Citizen Support for European Integration, 1973–2004’ (2007) 42 *Acta Politica* 128.

² Gráinne De Búrca, *After the EU Charter of Fundamental Rights: The Court of Justice as a human rights adjudicator?*, 20 *MAASTRICHT JOURNAL OF EUROPEAN AND COMPARATIVE LAW* 168–184 (2013); Gráinne de Búrca, *The Road not Taken: The European Union as a Global Human Rights Actor*, 105 *THE AMERICAN JOURNAL OF INTERNATIONAL LAW* 649 (2011); Sionaidh Douglas-Scott, *The European Union and Human Rights after the Treaty of Lisbon*, 11 *HUMAN RIGHTS LAW REVIEW* 645–682 (2011); Jörg Polakiewicz, *EU Law and the ECHR: Will EU Accession to the European Convention on Human Rights Square the Circle?*, SSRN (2013); Sara Sánchez Inglesias, *The Court and the Charter: The impact of the entry into force of the Lisbon Treaty in the ECJ’s approach to fundamental rights*, 49 *COMMON MARKET LAW REVIEW* 1565–1612 (2012).

³ Annette Schrauven, ‘European Union Citizenship in the Treaty of Lisbon: Any Change at All?’ (2007) 15 *Maastricht Journal of European and Comparative Law* 55.

⁴ De Búrca (n 2).

⁵ *ibid* 169.

enforcing human rights.’⁶ For many, the coming into force of the Lisbon Treaty marked the EU’s new age as a human rights actor,⁷ where the EU’s centre of gravity is no longer limited to the policy areas such as free trade, the single market and regulation⁸ but has expanded to the fields typically reserved for its sister organisation, the Council of Europe.⁹ Others have cast doubt on the ability of the CJEU to act as a human rights actor.¹⁰ Interestingly, some CJEU judges have explicitly declared their Court is not a ‘human rights court’.¹¹ While it has been claimed that the use of human rights discourse by the CJEU reflects less an aspiration to promote rights than a desire to strengthen the principles of supremacy, autonomy and legitimacy of EU law,¹² some see the recent case law as evidence that the CJEU is taking rights more seriously.¹³

Yet, given the unrepresentative character of the sample of EU legal materials on which these statements are likely to be based, it is not illegitimate to ask whether they actually provide an accurate depiction of reality. Have the Maastricht and Lisbon treaties really marked inflections in the focus of EU legislators and judges? Is EU law more about markets than it is about rights and citizens? To give but one example, the most popular EU law textbook¹⁴ devotes 17 per cent of its attention (as measured in pages) to human rights, citizenship, anti-discrimination law and criminal matters. But is this representative of the prominence of these topics in EU legislation and CJEU rulings? Does legal scholarship provide an accurate picture of the balance of issues that form the bulk of EU legislation and adjudication?

Before joining the CJEU as Advocate General, Michal Bobek criticised legal scholars for their overemphasis on constitutional cases:

⁶ *ibid* 171. Burca emphasises that a significant part of the EU law ‘corpus now covers areas such as immigration and asylum, security and privacy, alongside many of the more traditional fields of EU policy including competition and market regulation.’

⁷ Gráinne de Búrca (n 2) 649; Sánchez Inglesias (n 2) 1565.

⁸ Douglas-Scott (n 2) 645.

⁹ Gráinne de Búrca (n 2) 650.

¹⁰ S Douglas-Scott, ‘The European Union and Human Rights after the Treaty of Lisbon’ (2011) 11 *Human Rights Law Review* 645; T Lock, ‘Walking on a Tightrope: The Draft ECHR Accession Agreement and the Autonomy of the EU Legal Order’ (2011) 48 *Common Market Law Review* 1025; Polakiewicz (n 2); De Búrca (n 2) 172..

¹¹ Declaration by CJEU President Vassilios Skouris, FIDE Conference, Copenhagen, May 2014, cited in Julian Nowag, ‘EU Law, Constitutional Identity, and Human Dignity: A Toxic Mix? *Bundesverfassungsgericht: Mr R*’ (2016) 53 *Common Market Law Review* 1441, 1452.; Douglas-Scott (n 10) 649.

¹² Douglas-Scott (n 10) 649; Jason Coppel and Aidan Neill, ‘The European Court of Justice; Taking Rights Seriously?’ (1992) 29 *Common Market Law Review* 669, 672. That the CJEU’s use of human rights as a defence for the principle of supremacy, Coppel and O’Neill wrote, can be seen from the fact that the Court placed fundamental rights on the same conceptual level as the principle of economic freedom and the free movement of goods. See *ibid* 689. See Case C-222/86 *UNECTEFv. Heylens*, [1987] ECR 4098, 4117.

¹³ Coppel and Neill (n 12) 669.

¹⁴ Paul Craig and Grainne de Búrca, *EU Law: Text, Cases, and Materials* (Oxford University Press 2015).

‘The normal life of EU law’ is not defined by grand constitutional battles on the question of EU law supremacy over national law that reach the courts once in every ten years, but rather by thousands of dull tax cases, consumer protection actions, common customs tariff classification disputes, trans-border enforcement of small civil claims, companies’ shareholders quarrels and so on.’¹⁵

Similarly, in a seminal article with Hans-Wolfgang Micklitz, Rob van Gestel, who himself specialises in legislative drafting, has lamented the tendency of EU law scholarship to lapse into ‘case law journalism’.¹⁶ But do these assertions constitute a fair description of the focus of EU law scholarship? Do EU law academics overly concentrate on constitutional questions, thereby misleadingly suggesting that they represent the daily life of EU law?

Because of the aforementioned cognitive challenge, these questions may appear impossible to answer with much certainty. But imagine that we could find a way to summarize the contents of all legislative acts and CJEU judgments issued over the EU’s lifetime and compare this summary with the writings of EU law professors. This is, in short, what the present Essay attempts to do. We explore a mega EU law corpus, comprising the entire universe of legislative acts adopted since 1966, all CJEU judgments and opinions published since 1963 and all articles from the *Common Market Law Review* published since 1963 and up to 2018. We then map issue proportion in this mega corpus by applying a natural language processing technique known as probabilistic topic modelling. Based on unsupervised machine learning, probabilistic topic modelling aids us to annotate, classify and visualise variations in topic proportion across time and sources. Importantly, our methodological approach does not impose any *a priori* classification on the documents but lets the topics emerge from the analysis. This

¹⁵ Michal Bobek, ‘Of Feasibility and Silent Elephants: The Legitimacy of the Court of Justice Through the Eyes of National Courts.’ in Maurice Adams and others (eds), *Judging Europe’s Judges: The Legitimacy of Case Law of the European Court of Justice Examined* (Hart Publishing 2013) 200. In a similar vein, Hans Micklitz has argued that contrary to Karen Alter’s claim, it is not court-competition or the desire to get empowered *vis-à-vis* higher courts that drives courts’ participation in Article 267 TFEU proceedings but the desire to resolve technical disputes at the national level, see Hans-W Micklitz, *The Politics of Judicial Co-Operation in the EU: Sunday Trading, Equal Treatment and Good Faith* (Cambridge university press, 2005) 433; see Urszula Jaremba, ‘National Judges as EU Law Judges: The Polish Civil Law System’ (phdthesis, Martinus Nijhoff Publishers 2014) 341. Alter’s court competition thesis was originally articulated in Karen J Alter, ‘Explaining National Court Acceptance of European Court Jurisprudence: A Critical Evaluation of Theories of Legal Integration’ in Anne-Marie Slaughter (ed), *The European Court and National Courts: doctrine and jurisprudence: legal change in its social context*, Ed. by Anne-Marie Slaughter e.a. (Hart, 1998).

¹⁶ Rob van Gestel and Hans-Wolfgang Micklitz, ‘Why Methods Matter in European Legal Scholarship’ (2014) 20 *European Law Journal* 292, 298. The expression “case law journalism” was coined by Pierre Schlag, ‘Spam Jurisprudence, Air Law, and the Rank Anxiety of Nothing Happening (A Report on the State of the Art)’ 97 *The Georgetown Law Journal* 821. In his sharp critique of legal scholarship, Schlag writes that ‘the case-law journalist inserts the cases into some sort of normative narrative. Almost invariably, the case-law journalist ends on a cheery normative prescription for betterment of law, nation, or world.’ See *ibid* 823.

is important since our starting point is precisely that nobody can realistically claim to know the entire set of legal issues and policy areas addressed by legislation, court cases and law review articles. Our computer-aided text analysis documents several trends in addition to constraining issue prevalence across the three subsets of documents. First, economic and market integration has been and remains central to EU legislation and CJEU rulings. EU law is in that sense still predominantly about *homo economicus* rather than *homo politicus* or *homo juridicus*. Second, social policy issues (social benefits, pension rights, housing, etc) feature more prominently in CJEU cases than in legislation. Third, whereas farming regulation and the Common Agricultural Policy (CAP) account for a large share of legislative and judicial activity, they hardly receive any attention from legal scholars.¹⁷ Fourth, although attention to human rights has increased across the three categories of documents, human rights topics are significantly more prominent in the writings of academics than in the textual production of legislators and judges. These findings, we speculate, can be rationalised, for a significant part, by the distinct institutional incentives and constraints that determine the thematic orientation of the text producers. EU law text producers operate in separate, though interrelated, institutional environments that shape the co-evolution of their textual output. While the voices of EU law cannot be completely at variance, institutional conditions allow for differing patterns of emphasis. Legislators, judges and scholars can be thought of as distinct, yet partially interdependent voices that – through the production of texts – tell us what EU law is about. The relationship among these partly interdependent, partly autonomous voices, we suggest, can be conceptualised in terms of co-evolution. Specifically, in the context of EU law, co-evolution entails that legislation influences issue attention on the CJEU, which, in turn, affects what law professors choose to write about.

The remainder of the paper is structured as follows. First, we reflect on the text-generating process – the factors that have the potential to shape our legal authors' choice of subject. On the face of things, it would seem that one could not imagine more different legal writers than legislators, judges and law professors. Yet we argue that, although these legal

¹⁷ The work of Francis Snyder, however, deserves mention. See, for example, Francis G Snyder, *Law of the Common Agricultural Policy* (Sweet & Maxwell 1985); Francis G Snyder, 'The Special Legal Status of Agriculture: Assumptions and Contradictions in Economic Law', *New Directions in European Community Law* (1990); Francis Snyder, 'The Common Agricultural Policy in the Single European Market', *Academy of European Law, Collected Courses of the Academy of European Law*, vol 2 (1992); Francis Snyder, 'The Use of Legal Acts in EC Agricultural Policy', *Sources and Categories of European Union Law: A Comparative and Reform Perspective* (1996); Francis Snyder, 'CAP' [2012] *The Oxford Handbook of the European Union*; Francis G Snyder and A Mahiou, *La Sécurité Alimentaire/Food Security and Food Safety* (Leiden: Martinus Nijhoff for the Hague Academy of International Law 2006). See also A Burrell, 'The CAP: Looking Back, Looking Ahead' (2009) 31 *European Integration* 271.

authors are subject to differing institutional incentives and constraints, litigation often brings the output of EU legislators before the CJEU. Legal academics, meanwhile, are expected to cover judicial developments, at least as much as legislative ones. We introduce the concept of co-evolution to make sense of the dependencies among our three categories of EU law text producers. After that, we present our corpus and text-mining methodology. While describing our rich corpus, we explain the basic assumptions underpinning probabilistic topic modelling. We emphasise the unsupervised, let-the-data-speak character of the technique and its analytical leverage for the exploration of large collections of legal texts. Finally, we report and discuss our results. We do so mainly through visuals comparing topic prevalence across legislation, court cases and law review articles synchronically as well as over time. We conclude with some thoughts and suggestions on how to further refine our findings and how computer-aided text-mining methods ameliorate our understanding of the law.

2. The Production and Co-evolution of EU Law Texts

EU legislators, CJEU judges and academics who write about EU law operate in very different institutional environments. Depending on the policy domain and legislative procedure, the adoption of EU legislation may involve Commissioners, national government representatives, MEPs and various committees.¹⁸ Procedures have changed over time, as new treaties extended qualified majority voting and created new legal bases for EU rulemaking while vesting new powers in the European Parliament. Formally speaking, the European Commission has a general monopoly over legislative initiatives. Politically, though, national governments play a key role in defining the broad contours of the legislative agenda announced in European Council meetings.¹⁹ Together, the Commission and the European Council effectively act as the main legislative agenda-setters. Yet, thanks to the extension of its co-decision powers, the European Parliament has become a powerful veto-player able to block the passage of legislative proposals supported by the other institutions.²⁰ At the high level of generality assumed in this Essay, MPs,

¹⁸ Simon Hix and Bjorn Hoyland, *The Political System of the European Union* (Palgrave Macmillan 2011); Thomas Christiansen and Mathias Dobbels, 'Non-Legislative Rule Making after the Lisbon Treaty: Implementing the New System of Comitology and Delegated Acts' (2013) 19 *European Law Journal* 42.

¹⁹ Sebastiaan Princen, *Agenda-Setting in the European Union* (Springer 2009); Amie Kreppel and Buket Oztas, 'Leading the Band or Just Playing the Tune? Reassessing the Agenda-Setting Powers of the European Commission' (2017) 50 *Comparative Political Studies* 1118; Sebastiaan Princen, 'Agenda-Setting in the European Union: A Theoretical Exploration and Agenda for Research' (2007) 14 *Journal of European Public Policy* 21.

²⁰ For an elaboration of the European Parliament's ability to influence agenda setting under various decision rules see George Tsebelis, 'The Power of the European Parliament as a Conditional Agenda Setter' (1994) 88 *The American Political Science Review* 128; George Tsebelis, 'More on the European Parliament as a Conditional

Commissioners and national governments can be viewed as the main authors of EU legislation, although their respective influence over the legislative has fluctuated as a result of shifts in the prevailing balance of powers under the successive treaty regimes.

Unlike EU legislators such as the European Commission, CJEU judges cannot act *sua sponte*. They can only engage with an issue if a litigant files a suit. Case initiators act, in that sense, as functional co-authors of the judges' written output. It follows that litigation dynamics effectively determine the actual scope of judicial discourse. So examining who enjoys the right to initiate a case and under what conditions may already go some distance towards explaining what questions occupy the judges' agenda. The three main procedures governing access to the Court of Justice and the General Court come with distinct causes of action, remedies and standing rules. Infringement cases may be brought by national governments (who rarely do so) or the European Commission (who is meant to serve as chief European prosecutor). The right to bring annulment actions against EU acts belongs to EU institutions and member state governments. Private litigants may also bring annulment actions but only in restricted circumstances. Preliminary references are believed to offer private parties the most accessible channel to reach the CJEU,²¹ although the decision to refer ultimately lies with the domestic court hearing the case. Differences in procedural setup translate into different dockets, with judicial issue attention varying according to the priorities of the enfranchised case initiators.

Compared to judges and legislators, legal academics would seem to enjoy greater leeway in deciding what to write about. After all, the liberty to choose what one writes about arguably constitutes a cornerstone of academic freedom. However, several factors conspire to make academic writing more predictable than this principle, alone, would suggest. First, bar association regulations and law school curricula often determine what law has to be taught. Together with habit and tradition, these institutional factors influence which candidates are looked for and hired on the academic job market. We speculate that the topic on which legal academics choose to specialise is often influenced by the courses and instructors that they were exposed to during their studies and the specialisation of their doctoral supervisor, thereby inducing a great deal of thematic continuity in the focus (and methodology) of legal scholarship.

Agenda Setter: Response to the Moser' (1996) 90 The American Political Science Review 839; George Tsebelis and Geoffrey Garrett, 'The Institutional Foundations of Intergovernmentalism and Supranationalism in the European Union' (2001) 55 International Organization 357.

²¹ Karen J Alter, 'Private Litigants and the New International Courts' (2006) 39 Comparative Political Studies 22.

Second, despite attempts to revive the study of legislation,²² legal education and legal scholarship tend to concentrate on judicial activity – to such an extent that studying law is often implicitly equated with studying court rulings. Van Gestel and Micklitz’s “case law journalism” quip suggests that EU law scholarship is equally obsessed with court rulings.²³ Third, from the perspective of legal academia, judicial decisions themselves are not all born equal. Decisions that only resolve individual disputes without altering the stock of legal doctrines are of little interest to legal scholars, who typically restrict their doctrinal inquiry to the decisions that have the potential to become precedents. Finally, one should not underestimate the power of fashion in shaping the interests of academic writers. Some topics – say judicial independence in Central and Eastern European member states – simply appear more exciting than others – e.g. VAT.

While operating within distinct institutional frameworks, the three sites of EU law text production are nonetheless interrelated. CJEU judges are asked by litigants to review the validity of EU acts or to clarify their meaning. Adjudication, in that sense, responds to legislation, although always through the mediation of litigation. There might also be instances where EU legislators respond to CJEU rulings, although the evidence suggests that these cases are rare.²⁴ Legal scholarship has a greater degree of autonomy both *vis-à-vis* judges and *vis-à-vis* legislators. Yet what prompts an EU law scholar to write an article is often a salient CJEU ruling. Furthermore, a study by Harm Schepel and Rein Wesseling found a significant proportion of articles in the *Common Market Law Review*, the *Cahiers de droit européen*, and *Europarecht* to have authors affiliated with the European Commission, the CJEU or other EU institutions.²⁵ We would expect such institutional links to increase thematic overlap across legislative, judicial and academic discourse.

A helpful concept to capture the dynamic relationship between these three sources of EU law writings is that of *co-evolution*. Borrowed from biology, the term co-evolution designates a process by which two or more distinct groups evolve in tandem through reciprocal influence.²⁶

²² Ittai Bar-Siman-Tov, ‘The Global Revival of Legisprudence: A Comparative View on Legislation in Legal Education and Research’, *Conceptions and Misconceptions of Legislation* (Springer 2019).

²³ See van Gestel and Micklitz (n 16) 298.

²⁴ See the discussion on legislative overrides Olof Larsson & Daniel Naurin, *Judicial Independence and Political Uncertainty: How the Risk of Override Affects the Court of Justice of the EU*, 70 INTERNATIONAL ORGANIZATION 377–408 (2016); Gareth Davies, *Legislative control of the European Court of Justice*, 51 COMMON MARKET LAW REVIEW 1579–1607 (2014).

²⁵ Harm Schepel and Rein Wesseling, ‘The Legal Community: Judges, Lawyers, Officials and Clerks in the Writing of Europe’ (1997) 3 European Law Journal 165.

²⁶ On the notion of coevolution in the social and natural sciences see Miguel A Gual and Richard B Norgaard, ‘Bridging Ecological and Social Systems Coevolution: A Review and Proposal’ (2010) 69 Ecological Economics

Applied to EU law discourse, this definition suggests that legislative, judicial and academic writers constitute separate but interdependent groups. Even though the concept of convolution does not necessarily entail that reciprocal influences are symmetric in magnitude, it implies that one group cannot exist, or at least not in the same way, without the other. It is, indeed, difficult to imagine how EU scholarship would look like were the CJEU to stop issuing decisions. Assuredly, EU law scholarship would then lose its principal *raison d'être*. In like fashion, much of the decision making activity of CJEU judges would simply evaporate if the Commission, Parliament and Council ceased to produce legislation. Yet, because of legal scholars' focus on courts and relative neglect of legislation, we should expect considerable thematic overlap between judges and legislators, and judges and academics but less between academics and legislators.

3. EU Law as Data

a. EU Law Corpus

We used computer scripts written in the programming language *R* to extract legislative acts and CJEU documents from the EUR-Lex website. CLMR articles were obtained from the journal's webpage repository to which we had free access through our host institution. What motivated our choice for this particular law journal is that, although there are other important, established legal periodicals specialising in EU law, CMLR is, among the Anglophone publications, the one that goes furthest back in time. Whereas its editorial policy does, to some extent, differ from that of the *European Law Journal* and the *European Law Review* – it does not publish much interdisciplinary legal research, for example – it seems reasonable to assume that articles and case notes appearing in CMLR constitute a fairly representative sample of Anglophone EU law scholarship and of its evolution over time.²⁷

707; the term 'coevolution' was originally coined by Paul R Ehrlich and Peter H Raven, 'Butterflies and Plants: A Study in Coevolution' (1964) 18 *Evolution* 586.

²⁷ Our assumptions here are congruent with those of the aforementioned study by Harm Schepel and Rein Wesseling that also chose CMLR as representative of Anglophone EU law scholarship, see Schepel and Wesseling (n 25) 171.

As some documents – notably older legislative texts and CJEU decisions – were only available in PDF format, we applied optical character recognition to digitalise them. Although we collected the entire universe of published legislative and CJEU documents from 1953 onwards, our operational corpus – the one we use to perform our text-mining analyses – ignores legislative documents issued prior to 1966 and CJEU documents produced prior to 1963. The reason is that the very low number of legislative acts and judicial documents issued prior to these dates proved a source of modelling complications. Yet these documents constitute a tiny fraction of the corpus, meaning that this methodological choice can be regarded as innocuous. The three sets of documents, or sub-corpora, making up our operational corpus are summarised in Table 1. The first and smallest sub-corpus consists of 3809 CMLR articles and case notes. The second set of documents encompasses all Court of Justice, General Court and Civil Service Tribunal opinions and judgments along with all Advocate General opinions issued since 1963 and available at the time of writing. Figure 1 illustrates the distribution of these judicial documents over time. Whereas the number of Advocate General opinions has remained more or less constant since the 1970s, the number of Court of Justice and General Court rulings has experienced a sharp increase from 2005 onwards. The third and, by far, the largest sub-corpus comprises all the legislative acts passed by the European Union institutions between 1966 and 2017. Legislative acts include treaties, directives, regulations and decisions. Figure 2 shows the proportion of directives, regulations and decisions over time. Regulations clearly account for the bulk of legislative activity. Noteworthy, too, is the decline in legislative production since the turn of the millennium.

Table 1. Description of sub-corpora

Sub-corpus	Number of documents	Average number of words	Standard deviation	Min number of words	Max number of words
CMLR	3,809	6,991	6,175	28	84,741
CJEU	55,703	2,760	4,525	15	151,786
Legislation	194,655	1,748	7,085	6	694,435

Figure 1. Documents in judicial sub-corpus by year of issuance

Figure 2. Document types in legislative corpus by year of issuance

b. Methodology

Our main text-mining method is known as ‘probabilistic topic modelling’.²⁸ Probabilistic topic modelling has been developed for the purpose of discovering and annotating large archives of documents with thematic information. As with all-natural language processing techniques, probabilistic topic modelling relies on simplifying assumptions – in short, it does not parse words, sentence and contexts quite the way the human brain does. Yet, it has been shown to perform well on a wide range of documents, including legal ones.²⁹

Probabilistic topic modelling is an unsupervised machine learning method. That is to say, the document classification is not learned from human annotations but “discovered” by

²⁸ David M Blei, ‘Probabilistic Topic Models’ (2012) 55 Communications of the ACM 77.

²⁹ *ibid*; Michael A Livermore, A Riddell and Daniel Rockmore, ‘Agenda Formation and the US Supreme Court: A Topic Model Approach’ [2016] Arizona Law Review; David J Carter, James Brown and Adel Rahmani, ‘Reading the High Court at a Distance: Topic Modelling the Legal Subject Matter and Judicial Activity of the High Court of Australia, 1903-2015’ (2016) 39 UNSWLJ 1300.

the algorithm itself. The basic intuition behind the technique is that documents addressing the same topic are likely to contain similar words. Words such as ‘import’, ‘products’ and ‘tariffs’ will appear more often in documents about trade whereas ‘residence’, ‘citizenship’, ‘family’ and ‘visa’ will be more common in documents about immigration. A document – whether a legislative act, a ruling or a law review article – will frequently concern more than one topic. A ruling about a procedural or constitutional issue will also refer to the more substantive issue of the case – which may be milk quotas or wildlife conservation. Topic modelling builds on these assumptions to model topics as a cluster of words and documents as a mixture of topics. In more technical language, a topic model posits a latent space in which topics are represented as probabilistic distributions over words and documents as probabilistic distributions over topics. Once the researcher has set the desired number of topics, commonly denoted by the letter k , the algorithm seeks the topics (cluster of words) that best classifies the documents (mixtures of topics) in the corpus.

The k number of topics chosen to represent the content of a corpus is set by the researcher. This parameter is best viewed as determining how fine-grained a picture of the corpus the final topic model should provide.³⁰ A large k will result in a model with lots of very specific topics. A smaller k will produce a model with fewer topics but broader themes. For our analysis of EU law, we construct four topic models. One for the aggregate corpus and one for each of the three sub-corpora. For each of these models, we chose a number of topics that appeared to strike a reasonable balance between precision and tractability. As opposed to the number of topics, the words defining the topics are not chosen by the researcher but are produced by the model. For that reason, probabilistic topic modelling is known in natural language processing as an ‘unsupervised document classification’ technique.

Given the considerable temporal variation in document numbers (as illustrated in Figure 1 and 2), one might be concerned that a computerised text analysis would be skewed towards the period in which the largest number of documents has been produced. For the CJEU, for instance, we might fear that our topic model would essentially be a representation of topic

³⁰ Heuristics, such as perplexity and semantic exclusivity, have been proposed for choosing k values, see the discussion in Jonathan Chang and others, ‘Reading Tea Leaves: How Humans Interpret Topic Models’, *Advances in neural information processing systems* (2009); Hanna M Wallach and others, ‘Evaluation Methods for Topic Models’, *Proceedings of the 26th annual international conference on machine learning* (ACM 2009); David Mimno and others, ‘Optimizing Semantic Coherence in Topic Models’, *Proceedings of the conference on empirical methods in natural language processing* (Association for Computational Linguistics 2011). However, these heuristics do not only lack validity. Even if they were valid, there is no ‘correct’ k independently of the context and aim of the research.

prevalence after 2005. This is the reason why the topic model we utilise is dynamic. It assumes that thematic attention may change over time but gives each year the same weight.

Before running the algorithm we pre-processed the documents, as is common practice in natural language processing. Pre-processing involves removing words and symbols that provide no thematic information: punctuation, numbers and very frequent terms such as ‘the’, ‘but’, ‘with’. We also removed words which appear in a high proportion of documents, as they are not useful for discriminating between different topics and make the interpretation of the model more difficult. Characters were all converted to lowercase so as to ensure that ‘Legal’ and ‘legal’, for example, are treated as the same word.³¹ All these steps were performed using the *R* programming language.

4. Variations in Issue Attention in Legislative, Judicial and Academic Texts

This section presents and discusses the results of our topic modelling exercise.

a. Mega Topic Model of EU Law

We begin our textual exploration of EU law with our ‘mega’ topic model, which treats the three text collections as a single corpus. The model summarises this aggregate corpus in 63 topics. For this topic model, we restricted the analysis to 1972-2017 to facilitate interpretation.³²

Plotted in Figure 3 are the 63 topics with their four most characteristic words. It is important to bear in mind that these words have not been selected by the researchers but are produced by the algorithm. We did not know in advance what topics would emerge from the model. Interpreting the keyword-like terms defining a topic requires some basic knowledge of EU law and institutions. For example, the terms ‘questions’, ‘interpreted’, ‘referring’, ‘jurisdiction’ may not seem readily meaningful to a layperson but a jurist with some knowledge of EU law will instantly recognise words associated with the preliminary ruling procedure. In Figure 3, the position a topic occupies on the horizontal axis indicates its overall prevalence in the corpus. The most prevalent topic appears in the upper-right region of the plot. Its most characteristic terms are ‘wheat’, ‘flour’, ‘malt’ and ‘rye’. This topic accounts for more than 6

³¹ The final pre-processing step involved transforming the pre-processed corpus in a document-term matrix, in which each document is represented as a row and each term as a column. The matrix cell corresponding to a given document and a given term indicates how many times this particular word occurs in that particular document.

³² A topic model including documents issued over the period 1966-1971 produced topics more difficult to interpret. This problem results from the lower number of documents produced in this period. It could be addressed by a topic model with significantly a larger *k* parameter value but this would come in the way of tractability (a topic model with 200 topics or more would take a full article to report).

per cent of the overall thematic attention. The second most prevalent topic is defined by the terms ‘sugar’, ‘raw’, ‘white’ and ‘sucrose’. The fourth (‘fruit, vegetables, values, agriculture’) and fifth (‘rice, grain, broken, starch’) most-prevalent topics are also related to agriculture. The topic in the third position from the top (‘tender, invitation, tenders, intervention’) suggests that public procurement is another area of intense activity.

While we do find topics pertaining to judicial and constitutional issues (eg ‘constitutional, courts, fundamental, political’) the overwhelming majority of topics address either trade, market or farming activities. As we shall see when we examine the separate topic models for legislation, CJEU documents and CLMR contributions, the predominance of economic and farming issues is in large part the consequence of the legislators’ voice. Yet these themes are also, albeit to a lesser extent, central to the CJEU’s dispute settlement activity.

Figure 3. Mega topic model of EU law (1972-2017)

b. Legislative, Judicial and Academic Sub-corpora

For the three sub-corpora, we chose topic models with, respectively, 35 (legislation and court) and 45 (CMLR) topics. We first examine the model corresponding to the largest sub-corpus,

EU legislation. We find, again, a high proportion of agriculture-related topics, with again cereals ('wheat, flour, maize, malt'), sugar ('sugar, white, basic, exported') and rice ('rice, grain, broken, processed') coming out on top. Public procurement comes in the fourth position but is followed by yet another farming topic ('fruit, values, vegetables, nomenclature'). When not with agricultural commodities, other topics deal with competition (eg topic 'competition, business, services, share'), trade (eg topic 'industry, dumping, producers, investigation') and transportation ('transport, services, service, airport'). Only two topics do not seem to bear a direct relationship to the single market or the CAP. One relates to the Common Foreign and Security Policy ('security, cfsp, birth, person'); the other to environmental protection ('safety, assessment, waste, environmental'). Overall, Figure 4 shows considerable overlap with Figure 3, which is hardly surprising given the imbalance in the size of the three sub-corpora and the hefty share accounted for by legislative documents. If anything, the topics in Figure 4 appear even more technical and market-centred. The high prevalence of farming topics can be explained by the regulatory regime embraced by the CAP, which leans heavily on market intervention. Despite recent reforms, the CAP is not exactly known for its light-touch approach to regulation.³³ Frequent interventions concerning specific commodities translate into abundant, short-term regulations.

³³ Alan Greer, 'Post-Exceptional Politics in Agriculture: An Examination of the 2013 CAP Reform' (2017) 24 *Journal of European Public Policy* 1585.

Figure 4. Topics and topic proportion, legislative sub-corpus (1966-2017)

Table 2 shows the legislative documents with the highest topic proportion for a sample of legislative topics. For example, the document with highest proportion of topic ‘type, vehicle, vehicles, approval’ is Directive 2009/67/EC of the European Parliament and of the Council of 13 July 2009 on the installation of lighting and light-signalling devices on two or three-wheel motor vehicles. Table 2 shows how topic modelling can be used to search for legal documents pertaining to a specific topic. A topic model gives us an indication of topic proportion not just for a single document but for all documents in the corpus – which is a great advantage over methods based on keyword search. Inspecting documents with high topic proportion can also help interpret topics that may at first appear ambiguous. For example, the topic ‘appropriations, expenditure, staff, revenue’ may not look immediately obvious, but it becomes very clear when

we look at the act featuring the largest proportion of the topic, namely the Statement of the revenue and expenditure of the Translation Centre for the Bodies of the European Union for the financial year 2007.

Table 2. Acts with the largest topic proportion for a sample of legislative topics

Topic	Act	Year	Topic proportion
type, vehicle, vehicles, approval	Directive 2009/67/EC of the European Parliament and of the Council of 13 July 2009 on the installation of lighting and light-signalling devices on two or three-wheel motor vehicles	2009	99.96%
security, cfsp, birth, persons	Council Implementing Decision 2012/167/CFSP of 23 March 2012 implementing Decision 2011/486/CFSP concerning restrictive measures directed against certain individuals, groups, undertakings and entities in view of the situation in Afghanistan	2012	99.95%
fishing, vessels, vessel, fisheries	Commission Implementing Regulation (EU) 2017/1594 of 25 September 2017 adding to the 2017 fishing quotas certain quantities withheld in the year 2016 pursuant to Article 4(2) of Council Regulation (EC) No 847/96	2017	99.89%
appropriations, expenditure, staff, revenue	Statement of revenue and expenditure of the Translation Centre for the Bodies of the European Union for the financial year 2007	2007	99.81%
weight, exceeding, milk, matter	Commission Regulation (EEC) No 3567/83 of 16 December 1983 fixing the export refunds on milk and milk products	1983	99.80%
rice, grain, broken, processed	Commission Regulation (EEC) No 417/80 of 21 February 1980 fixing the export refunds on rice and broken rice	1980	98.83%

Moving from legislative to judicial discourse, Figure 5 illustrates our topic model for the CJEU texts. In comparison to legislators, EU judges devote less attention to CAP issues, although these are still well represented (e.g. topic ‘sugar, price, wine, tobacco’ and topic ‘milk, agricultural, quota, levy’). The sizable number and monolithic character of staff cases, which for a time were handled by a specialised Civil Service Tribunal, make this topic (‘staff, officials, post, committee’) the most prevalent in the judicial corpus. Probably because they rarely, if ever, result in landmark rulings with far-reaching doctrinal and policy implications, staff cases seldom get mentioned in academic publications. Yet, they form a constant source of workload for the EU judiciary.³⁴ Many topics deal with single market themes (eg customs, tariff nomenclature, VAT, product packaging, taxation of cross-border operations, competition...). Yet, we see some social policy (‘employment, workers, pension, benefits’) and citizenship themes (‘residence, family, nationals, child’), too.

Overall, our unsupervised document classification reveals the judicial sub-corpus to be more diverse than its legislative counterpart. The *homo economicus* looms large in CJEU adjudication, but not as conspicuously so as in legislation. This topic model topic gives a sense of what the typical case on the CJEU docket looks like. Topics such as ‘customs, export, imports, dumping’, ‘tariff, heading, customs, classification’, and ‘vat, tax, taxable, supply’ lend some support to Michal Bobek’s claim that these dull subjects represent the normal life of EU law.³⁵ Yet although topics like ‘employment, workers, pension, benefits’, ‘residence, family, nationals, child’ and ‘criminal, charter, penalty, penalties’ suggest that this claim must be nuanced.

Similar to Table 2, Table 3 shows the cases with the highest topic proportion for a handful of topics. The names of the parties involved suggest that the listed topics are, indeed, central to these rulings. The ruling most characteristic of the topic ‘criminal, charter, penalty, penalties’, *PPU Minister for Justice and Equality v Francis Lanigan* (Case C-237/15), involves the Irish Ministry of Justice. Not unexpectedly, given the words defining the topic, the Grand Chamber ruling involves the interpretation of both the Framework Decision on the European Arrest Warrant and the Charter of Fundamental Rights and Freedoms of the European Union. Likewise, it should not be surprising *Maria Teixeira v London Borough of Lambeth and Secretary of State for the Home Department* (Case C-480/08) involves the British Home

³⁴ The Civil Service Tribunal was created in 2005 and abolished in 2016 with its agenda reabsorbed by the General Court.

³⁵ Bobek (n 15).

Department given that it is the ruling most characteristic of the topic ‘residence, family, nationals, child’.

Figure 5: Topics and topic proportion in judicial sub-corpus (1963-2018)

Table 3. Cases with the largest topic proportion for a sample of judicial topics

Topic	Case	Year	Topic proportion
tariff, heading, customs, classification	Case 54/79 Firma Hako-Schuh Dietrich Bahner v Hauptzollamt of Frankfurt am Main-Ost	1979	98.88%
ohim, opposition, registration, euipo	Case T-144/10 Space Beach Club, SA v Office for Harmonisation in the Internal Market (Trade Marks and Designs) (OHIM)	2010	98.71%
restrictive, entities, cfsp, funds	Case T-43/12 Mohamad Hamcho and Hamcho International v Council of the European Union	2012	98.57%
tax, taxation, income, resident	Case C-35/11 Test Claimants in the FII Group Litigation v Commissioners of Inland Revenue and The Commissioners for Her Majesty's Revenue & Customs	2011	98.55%
health, product, medicinal, authorisation	Case C-31/03 Pharmacia Italia SpA, formerly Pharmacia & Upjohn SpA	2003	98.51%
customs, export, imports, dumping	Case C-34/14 A v Hauptzollamt Nürnberg	2014	98.15%
criminal, charter, penalty, penalties	Case C-237/15 PPU Minister for Justice and Equality v Francis Lanigan	2015	97.50%
residence, family, nationals, child	Case C-480/08 Maria Teixeira v London Borough of Lambeth and Secretary of State for the Home Department	2008	97.44%
vat, tax, taxable, supply	Case C-264/17 Harry Mensing v Finanzamt Hamm	2017	97.31%
air, directives, transposition, transpose	Case C-359/17 Marcel Lutz and Others v TUIfly GmbH	2017	92.35%
data, disclosure, personal, investigation	Case T-300/10 Internationaler Hilfsfonds v Commission	2010	90.04%

Turning now to the CMLR corpus, Figure 6 depicts our ‘academic’ topic model. In contrast to the legislative and judicial topic models, the CMLR topic model indicates considerably lower attention for CAP issues but considerably more attention for constitutional matters. Except for

one topic ('tariff, levy, sugar, quotas'), we do not find topics that clearly relate to agricultural commodities or CAP regulations and disputes. By contrast, several topics pertain to constitutional issues. One such topic is 'ecthr, strasbourg, constitutions, vires'. As shown in Table 4, the CMLR contribution most characteristic of this topic is Matthias Kumm's 'Who is the Final Arbiter of Constitutionality in Europe?: Three Conceptions of the Relationship Between the German Federal Constitutional Court and the European Court of Justice'. Table 4 confirms that the topic 'conseil, etat, cour, cassation', too, can be interpreted as constitutional. The document containing the highest proportion of the topic is a 1974 article titled 'How Supreme is Community Law in the National Courts?' by Gerhard Bebr.

Figure 6. Topics and topic proportion in the CMLR sub-corpus (1963-2018)

Table 4. Articles with largest topic proportion for a sample of CMLR topics

Topic	Title of CMLR Article	Year	Topic proportion
preemption, erta, exclusivity, trips	‘Implied exclusive powers in the ECJ’s Post-Lisbon jurisprudence: the continued development of the ERTA doctrine’	2018	99.84%
pension, worker, employer, pensions	‘Social security cases in the Court of Justice of the European Communities’	1982	99.82%
procurement, award, tender, utilities	‘An assessment of the new legislative package on public procurement’	2004	99.81%
gatt, wto, acp, panel	‘EC-ACP economic partnership agreements and WTO compatibility: An experiment in north-south inter-regional agreements’	2006	99.80%
tariff, levy, sugar, quotas	‘The Agricultural Policy of the Community’	1963	99.76%
women, sex, disability, sexual	‘Sex equality in the single market: New directions for the European Court of Justice’	1995	99.70%
eea, efta, withdrawal, switzerland	‘Brexit means Br(EEA)xit: The UK withdrawal from the EU and its implications for the EEA’	2018	99.69%
cfsp, kadi, military, eeas	‘The Meaning of Petersberg: Some Considerations on the Legal Scope of ESDP Operations’	2007	99.60%
processing, privacy, internet, digital	‘Online service providers and liability: A plea for a balanced approach’	2011	99.54%
ecthr, strasbourg, constitutions, vires	‘Who is the Final Arbiter of Constitutionality in Europe?: Three Conceptions of the Relationship Between the German Federal Constitutional Court and the European Court of Justice’	1999	99.21%
conseil, etat, cour, cassation	‘How Supreme is Community Law in the National Courts?’	1974	99.76%

Topics arguably constitutional in character further include ‘women, sex, disability, sexual’ and ‘preemption, erta, exclusivity, trips’. The topic ‘eea, efta, withdrawal, switzerland’, which covers articles over Brexit (see Table 4), may also be seen as possessing a constitutional dimension, at least in the broad sense of the term.

Comparing the CMLR topic models with the legislative and judicial models appears to confirm the intuition that EU law scholars tend to focus more on what occurs at the CJEU stage than at the legislative stage. This appears to buttress the aforementioned observation of Van Gestel and Micklitz about “case law journalism”.³⁶ Then, when comparing judges and academics, the latter exhibit a greater predilection for ‘sexier’ constitutional topics, which echoes the critique of Michal Bobek. To be sure, there are academics who have written about the CAP³⁷ or public procurement (as exemplified by topic “procurement, award, tender, utilities”). Yet these topics are less prevalent in academic than in judicial discourse.

Interestingly, the article most representative of the sole academic topic presenting a real connection with the CAP – ‘tariff, levy, sugar, quotas’ – is penned by a European Commission official (see Table 4). ‘The Agricultural Policy of the Community’ was published in 1963 by Giancarlo Olmi, who worked for the European Commission’s legal service. While ascertaining the affiliation of CMLR contributors is beyond the scope of the present Essay, this finding is consistent with what we said above about the impact of overlap in personnel on the focus of EU law scholarship. As Harm Schepel and Rein Wesseling have demonstrated, albeit on the basis of a smaller corpus, many CMLR contributors have links to EU institutions.³⁸

c. Temporal Variations

Figure 7 is another illustration of the points we just made about the focus of EU legal scholarship. The plot depicts the relative proportion of the phrase ‘human/fundamental rights’ across the three sub-corpora over time. Although legislators, judges and academics have all been giving greater attention to fundamental rights, the emphasis on rights is clearly and

³⁶ van Gestel and Micklitz (n 16) 293.

³⁷ The work of Francis Snyder, however, deserves mention. See, for example, Francis G Snyder, *Law of the Common Agricultural Policy* (Sweet & Maxwell 1985); Francis G Snyder, ‘The Special Legal Status of Agriculture: Assumptions and Contradictions in Economic Law’, *New Directions in European Community Law* (1990); Francis Snyder, ‘The Common Agricultural Policy in the Single European Market’, *Academy of European Law, Collected Courses of the Academy of European Law*, vol 2 (1992); Francis Snyder, ‘The Use of Legal Acts in EC Agricultural Policy’, *Sources and Categories of European Union Law: A Comparative and Reform Perspective* (1996); Francis Snyder, ‘CAP’ [2012] *The Oxford Handbook of the European Union*; Francis G Snyder and A Mahiou, *La Sécurité Alimentaire/Food Security and Food Safety* (Leiden: Martinus Nijhoff for the Hague Academy of International Law 2006). See also Burrell (n 17).

³⁸ Schepel and Wesseling (n 25).

consistently stronger in academic writings. Consistent with our concept of co-evolution, the three discourses have, to a noticeable extent, developed along similar upward trajectories but with diverging growth rates.

Interestingly, the plot shows a clear increase in attention to human/fundamental rights coinciding with the adoption of the Lisbon Treaty in both the judicial and CLMR sub-corpora – plausibly as a consequence of the entry into force of the Charter of Fundamental Rights.

Figure 7. Relative frequency of the phrase human/fundamental rights in the three sub-corpora

As the issue of human rights illustrates, topic attention may undergo substantive shifts over time. For reasons of space, we do not report time trends for all topics and reserve them for the Appendix. Instead, we focus on a selection of legislative, judicial and CMLR topics that illustrate shifts which we judged to be substantively important. These are depicted in Figure 8, 9 and 10, where shaded bands around the curves indicate how certain the model is about the true topic proportion. As a cursory look at the three figures makes plain, shaded bands are larger for CMLR topics because the topic model is constructed from a smaller number of documents.

Fewer documents mean less information to adjust the topic model to the corpus and, therefore, more uncertainty over topic proportion. Similarly, confidence bands for the CJEU and legislative topics are larger in the early years, reflecting lower judicial and legislative output in those years.

More substantively, some of the downward and upward trends in Figure 8 and 9 are illustrative of the evolution of single market economic regulation. Legislation has become less about agricultural commodities (as exemplified by the declining share of ‘wheat, flour, malt, maize’) and more about free movement of capital (‘credit, risk, assets, capital’), the liberalisation of transport markets (‘transport, services, service, airport’) and the harmonization of safety standards (‘type, vehicle, vehicles, approval’). Likewise, Figure 9 shows that attention to tariffs and farming issues has decreased while the proportion of cases dealing with taxation, public procurement and energy markets has gone up. These dynamics appear to reflect distinct phases of the construction of the single market, with the early phase focusing more on the removal of tariff barriers and the establishment of a customs union and later phases focusing on deeper regulatory obstacles to the free movement of services, products and capital and the liberalisation of utility and transport markets.³⁹ The upward slope for the topic ‘credit, risk, assets, capital’ in Figure 8, in particular, seems to reflect legislative efforts, at the end of the 1990s, to make the free movement of capital an economic and legal reality.⁴⁰ The temporal trends for these topics along with those reported in the Appendix suggest that, as far as legislators and judges are concerned, economic integration and the *homo economicus* continue to occupy centre stage, although the emphasis has shifted from ‘old’ (tariffs and farming) to ‘new’ (transport, energy, taxation, technical standards) issues.

The trends for topic ‘security, cfsp, birth, persons’ in Figure 8, topic ‘restrictive, entities, cfsp, funds’ in Figure 9 and topic ‘cfsp, kadi, military, eeas’ in Figure 10 all point to an increase in attention to security and foreign policy issues around 2000 (before that mark topic proportion is zero or close to zero). The curves in the judicial and CMLR topic models even display similar shapes. Less surprising is the increase in attention to immigration (topic ‘asylum, migration, third country, immigration’ in Figure 10) and criminal matters (‘crime, arrest, warrant, idem’) in CMLR articles. Interest in these topics is relatively new. Interests in the criminal procedure

³⁹ Neil Kinnock, ‘The Liberalisation of the European Aviation Industry’ (1996) 8 *European business journal* 8; Tooraj Jamasb and Michael Pollitt, ‘Electricity Market Reform in the European Union: Review of Progress toward Liberalization & Integration’ (2005) 26 *The Energy Journal*.

⁴⁰ Leo Flynn, ‘Coming of Age: The Free Movement of Capital Case Law 1993-2002’ (2002) 39 *Common Market Law Review* 773.

can be linked to the legislative act introducing the European Arrest Warrant and the relatively large number of CJEU cases it has spawned.⁴¹ The CJEU's case law on citizenship and immigration is also of recent origin.⁴² Discriminations ('women, sex, disability, sexual'), by contrast, has been an established topic in EU legal scholarship for somewhat longer, which, again, reflects the temporal evolution of the CJEU's case law.⁴³

⁴¹ The first case decided under the European Arrest Warrant is C-303/05, *Advocaten voor de Wereld*, issued in 2007. Until the end of 2018, there were 33 cases decided by the Court on the topic. For the overview of the cases, see Eurojust, '2018 Overview on the Case Law of the Court of Justice of the European Union (CJEU) on the European Arrest Warrant (EAW)'; Anne Pieter van der Mei, 'The European Arrest Warrant System: Recent Developments in the Case Law of the Court of Justice' (2017) 24 *Maastricht Journal of European and Comparative Law* 882; Luisa Marin, 'Effective and Legitimate?: Learning from the Lessons of 10 Years of Practice with the European Arrest Warrant' (2014) 5 *New Journal of European Criminal Law* 327.

⁴² The earliest case decided on the issue of EU citizenship dates from 2008 and concerns the Regulation 343/2003 on establishing the criteria and mechanisms for determining the Member State responsible for examining an asylum application. See Case C-19/08 *Migrationsverket v Edgar Petrosian* [2009]. Prof Dr Gerard-René de Groot and Ngo Chun Luk, 'Twenty Years of CJEU Jurisprudence on Citizenship' (2014) 15 *German Law Journal* 821; Camelia Toader and Andrei I Florea, 'Free Movement of Workers and the European Citizenship Studies and Articles' (2012) 2012 *Revista Romana de Drept European* 67; Willem Maas, 'The Origins, Evolution, and Political Objectives of EU Citizenship' (2014) 15 *German Law Journal* 797; Susanne K Schmidt, 'Judicial Europeanisation: The Case of Zambrano in Ireland' (2014) 37 *West European Politics* 769; Francesca Ippolito, 'Migration and Asylum Cases before the Court of Justice of the European Union: Putting the Eu Charter of Fundamental Rights to Test?' (2015) 17 *European Journal of Migration and Law* 1.

⁴³ Europäische Union, Europarat and Europäischer Gerichtshof für Menschenrechte (eds), *Handbook on European Non-Discrimination Law* (2018 edition, Publications Office of the European Union 2018); Evelyn Ellis, *European Community Sex Equality Law* (Clarendon Press; Oxford University Press 1991); Sonia Mazey, *Gender Mainstreaming in the EU: Principles and Practice* (Kogan Page 2001); Evelyn Ellis and Philippa Watson, *EU Anti-Discrimination Law* (OUP Oxford 2012); Craig and de Búrca (n 14); Claire Kilpatrick, 'Gender Equality: A Fundamental Dialogue' in Silvana Sciarra (ed), *Labour law in the courts: national judges and the European Court of Justice, edited by Silvana Sciarra*. (Hart Pub 2001); Claire Kilpatrick, 'The Court of Justice and Labour Law in 2010: A New EU Discrimination Law Architecture' (2011) 40 *Industrial Law Journal* 280.

Figure 8. Topic proportion over time for selected legislative topics

Shaded bands indicate 95 per cent confidence interval

Figure 9. Topic proportion over time for selected CJEU topics

Shaded bands indicate 95 per cent confidence interval

Figure 10. Topic proportion over time for selected CMLR topics

Shaded bands indicate 95per cent confidence interval

5. Conclusion

In this paper, we explored a novel, giant corpus of EU law texts using computer-based text-mining techniques. The aim of our analysis was to measure and compare issue attention in legislative, judicial and academic texts at a relatively high level of generality. Overall, our text analytics cast light on areas of both convergence and divergence, in line with the notion of co-evolution, which we suggested as a way to conceptualise the relations among the three domains. We found that, on the whole, there is greater overlap between the judicial and the legislative voice and between the judicial and the academic than between the academic and the legislative. Academics respond more to CJEU rulings than to legislative enactments, which implies that there might be a grain of truth in the “case law journalism” critique levelled at EU law

scholarship. From our analysis, legal scholars also look more inclined to shun the technical, regulatory stuff that forms the bread and butter of legislation and, to a large degree, of CJEU adjudication. The daily life of EU law, as some have put it, is indeed, more often than not about dull, unexciting matters such as VAT, tariff nomenclatures and agricultural commodities.

As EU law scholars' attention tends to be skewed towards constitutional and fundamental rights topics, a reader leafing through contributions in a law journal like CMLR might get the impression that EU law has moved away from businessmen's Europe more than it really has. According to our textual analysis, EU positive law continues to mostly concern itself with the regulation of economic activity, although the focus of market regulation has shifted somewhat, away from farming commodities and custom duties and towards capital flows, taxation and safety standards. While attention to fundamental rights in legislation and CJEU rulings seems to have increased in recent years, it is far from being central to the agenda of judges and legislators. On the other hand, the discrepancy in focus between legal scholarship and legislation and case law is something that one may put to the academics' credit. Indeed, it can be argued that it is the responsibility of legal scholars to take fundamental rights and constitutional issues seriously when judges and legislators are perceived to fail their duty.

In any case, the analysis presented in this Essay was conceived as primarily exploratory and descriptive and we did not purport to adjudicate these questions. While the methodological approach we adopted is indubitably unconventional, at least for a legal audience, recent contributions have presented corpus linguistics and text-mining techniques as providing a new methodological paradigm for European legal research.⁴⁴ Although we certainly see room for further refinements of our findings, we hope to have demonstrated what such methods can achieve in terms of uncovering patterns and dynamics in large collections of EU legal texts.

⁴⁴ Urska Sadl and Henrik Palmer Olsen, 'Can Quantitative Methods Complement Doctrinal Legal Studies? Using Citation Network and Corpus Linguistic Analysis to Understand International Courts' (2017) 30 *Leiden Journal of International Law* 327; Arthur Dyevre, Wessel Wijtvet and Nicolas Lampach, 'The Future of European Legal Scholarship: Empirical Jurisprudence' (2019) 26 *Maastricht Journal of European and Comparative Law* 348.

Appendix

Figure A1. Topic proportion over time for legislative topics.

Shaded bands indicate 95per cent confidence interval

Figure A2. Topic proportion over time for CJEU topics

Shaded bands indicate 95 per cent confidence interval

Figure A3. Topic proportion over time for CMLR topics

Shaded bands indicate 95 per cent confidence interval