

SNIS Final Report: A Quantitative Textual Approach of the European Consensus Method of Interpretation in the European Court of Human Rights

Prof. Jonathan B. Slapin

Abstract of Executive Summary:

The European Consensus method is a tool for judicial decision-making employed by the European Court of Human Rights when adjudicating on sensitive moral and social issues. Our project has employed human hand-coding combined with computational techniques to uncover the nature of European Consensus in judgments rendered by the Court. We find that European Consensus is a tool that has become increasingly common over time, is used more when the Court makes decisions in cases involving certain member states – most notably the United Kingdom, but also France – and when making decisions with respect to certain articles of the European Convention of Human Rights – most notably articles 8 (private and family life), 2 (life) 10 (freedom of expression) and 14 (non-discrimination). These findings help shape our understanding of how human rights law has evolved in Europe through the ECHR system.

Executive Summary:

How are complex and controversial moral and political questions surrounding human rights (HR) law adjudicated in Europe? The text of the European Convention on Human Rights (ECHR or the Convention), the principal bill of rights in Europe, is designed to be general and abstract. It requires interpretation to acquire concrete normative content and to provide answers to sensitive questions and to produce tangible effects in everyday life. Interpretation of the ECHR is the task, first, of national authorities and especially of the courts of the 47 member states of the Council of Europe (CoE) and, ultimately, of the European Court of Human Rights (ECtHR or the Court). When fulfilling its mandate as the authoritative interpreter of the ECHR, the ECtHR has a number of methods of interpretation at its disposal. This project has explored one of the most important of these methods, the European consensus (EuC) method.

EuC is a “tool” of legal interpretation created and often used by the Court when making decisions on morally, politically, or socially controversial, sensitive, or ambiguous human rights (HR) issues, such as abortion, same-sex partnership, or religious dress. In a nutshell, EuC is used to trace the evolution of laws regarding sensitive topics across states and other bodies, such as international organisations, enabling adjudicators to interpret human rights law in a dynamic way that reflects present-day conditions. Narrowly conceived, EuC consists of the comparative analysis of the laws and practices of CoE member states on the regulation of the HR question at issue, with a view to identifying whether a new shared understanding has emerged in Europe (or world-wide). The existence of EuC can lead to the establishment of common standards across Europe that is, in turn, a form of integration. Conversely, the absence of EuC creates more space for discretion (called margin of appreciation) by national authorities and, ultimately, sovereignty in the sense that national authorities are “free” to adopt the policy they prefer instead of having to comply with common, Pan-European standards.

But how precisely does consensus come into existence? Does consensus depend on whether particular states or a certain number of states recognise a right? Is the use of consensus in

ECtHR judgments associated with specific terminology and language? Under what conditions do the ECtHR Judges make use of this method of interpretation? Is the Court more likely to resort to this method when certain states are involved in a case? Taking a pluri-disciplinary approach, this project has borrowed quantitative tools from the social and political sciences to answer these important questions of the emergence of HR legal rules in Europe.

The project has empirically evaluated the use of EuC as a method of human rights interpretation in the ECtHR, focusing primarily on Grand Chamber judgments – the Court’s most important decisions. We find that Grand Chamber judgments make explicit reference to the language of European Consensus in approximately 24% of cases and use language that either explicitly or implicitly relates to Consensus in over 40% of cases. These numbers underestimate the prevalence of Consensus language in recent years as we find that the use of EuC has become increasingly common over time. While it was uncommon to find Consensus language in judgments written in the 1990’s it is increasingly found in more recent judgments. We find that Consensus is used more when the Court makes decisions in cases involving certain member states – most notably the United Kingdom, but also France – and when making decisions with respect to certain articles of the European Convention of Human Rights – most notably articles 8 (private and family life), 2 (life) 10 (freedom of expression) and 14 (non-discrimination).

These findings are important not only for understanding how the Court reaches decisions, but also for understanding how the Court seeks to increase the legitimacy of its decisions. Previous legal scholarship has argued that the use of EuC increases the legitimacy of Court decisions because the Court is able to argue that, rather than setting the human rights standards at issue itself, it is merely identifying and reiterating common practice of European states. It is not developing new standards out of thin air. We can now say the circumstances under which the Court uses this tool to seek legitimacy and how it does so.

Final Scientific Report:

Judges use methods of legal interpretation, or legal “tools”, to construct arguments when deciding cases. This is not only true of judges in national legal systems, but also of international judges. Judges on the European Court of Human Rights (ECtHR) -- the international court that adjudicates cases regarding human rights outlined in the European Convention of Human Rights (ECHR) -- have developed a particular method of legal interpretation, the method of European consensus (EuC). EuC is a “tool” used by ECtHR judges when making potentially controversial decisions about the emergence of new human rights, e.g., the right to have an abortion, or the right to same-sex marriage. It is a method of comparison to other countries’ laws and norms, but also to other sources of law, such as international law standards, the case law of other (international) courts, or the practice of international organizations. In short, EuC allows the Court to draw on external sources and on human rights standards developed at the national level or within various international organisations as a means of argumentation. Judges can, in effect, argue that it is the consensus stemming from and reflected in a variety of sources that a concrete human rights standard has (or has not) emerged. Such arguments, as asserted by scholars of European human rights law and the ECtHR, can confer legitimacy on ECtHR decisions and increase the likelihood that actors (e.g., judges and lawyers) in national legal systems pay them heed (Dzehtsiarou 2015).

In this project, we employ quantitative text-analytic methods, both human coding and computational methods, to gain insight into the process of legal decision-making and nature of legal argumentation. We go beyond existing literature in quantitative social science that focuses on case outcomes and their impact on international human rights law (see e.g. Voeten 2007, 2008, 2021). Until now, quantitative text-as-data methods as applied to the decisions of the ECtHR have been primarily used to predict possible outcomes of cases (Aletras et al. 2016; Medvedeva, Vols, and Wieling 2020), but they have not been used to understand the methods of legal interpretation that the Court uses to reach its decisions (Peat 2021).

Using quantitative text-analytic methods to identify and map the use of the EuC method over time leads us to a better understanding of how human rights law has changed and developed across Europe. It also provides the basis for answering more theoretically driven questions, such as when and why the ECtHR uses EuC when making judgments. Is EuC more prevalent in certain types of cases, amongst certain judges, or when certain ECHR parties are involved? We study when and where EuC is used to gain an understanding not only of how EctHR Judges make reasoned arguments and reach decisions but also to understand how new international human rights emerge across Europe. We find that the use of EuC as a tool for legal decision-making has increased significantly since the adoption of Protocol 11 ECHR. It is used more frequently when certain states are respondents in cases, and with respect to certain ECHR articles and legal issues.

In this report, we briefly introduce the EuC method and discuss how EctHR and judicial decision-making have been studied to date. We then describe our text-as-data approach to identifying the presence of EuC reasoning in judicial decision-making. Finally, we present some of our findings from the project and discuss avenues for future research.

What is European Consensus?

EuC is a “tool” of legal interpretation invented by and often used by the EctHR when making decisions on morally, politically, or socially controversial, sensitive, or ambiguous human rights issues, such as abortion, same-sex marriage, or religious clothing. In a nutshell, EuC is a comparative approach in which judges compare practices in other states and organizations to trace the evolution of (societal) norms regarding human rights. Using EuC enables adjudicators to interpret human rights law in a dynamic way that reflects present-day conditions and attitudes. This, in turn, opens opportunities for standard-setting via interpretation.

Narrowly conceived, EuC consists of the comparative analysis of the laws and practices of the Council of Europe (CoE) member states on the regulation and interpretation of the human rights question at issue, with a view to identifying whether a new shared understanding regarding human rights has emerged in Europe (or world-wide). As an example, we can take the case of same-sex marriage. While in the past there may have been no agreement across countries around the question of whether same-sex couples should have the right to marry, when deciding a case on whether such a right exists, EctHR judges can examine rules, laws and practices of member states of the CoE today and determine whether updated standards around the acceptance of same-sex marriage have emerged, justifying a decision to create such a right within the context of the EctHR. Thus far, EuC analysis has led the EctHR to conclude that no right to marriage shall exist for same-sex couples in Europe (e.g., *Hämäläinen v. Finland* [GC], 16-7-2014, 37359/09), but also that such couples shall be afforded some form of legal recognition of their union, such as civil law partnership (*Oliari and Others v. Italy*, 21-7-2015, 18766/11 and 36030/11).

Within the EuC framework, the EctHR may also engage with comparative analysis of sources of law outside of the practice of the CoE member states. For instance, it can consider the practice of international organizations, such as the CoE itself or the European Union (EU), and consider human rights standards developed by other, non-European institutions (e.g., the United Nations (UN)) or even by non-European states. If the EctHR identifies the existence of consensus, however defined, on a specific human rights issue, it may then move on to recognize pan-European standards that are binding on all states under its jurisdiction. In the absence of consensus, states and their national authorities enjoy a wider margin of appreciation¹, that is discretion to make their own regulatory choices. To that end, the EctHR’s level of judicial scrutiny, concerning in particular the test of proportionality/necessity, is lower.

Although we can offer a basic definition of EuC, as we have here, much about the method and its use by the Court remains unknown (Kapotas and Tzevelekos 2019, 9–10). EctHR judges may, themselves, lack a common understanding as to what EuC is, and perhaps also a common vocabulary. The Court will often use different terminology to refer to the same or similar concepts, making it difficult to ascertain whether the Court is truly referring to EuC. Alternatively, our ignorance about the nature of EuC may be the result of a strategic choice on the part of the Court, for whom the lack of a detailed definition of EuC may increase its flexibility when employing the method (see Kapotas and Tzevelekos 2019, 10). While we can

¹ Specifically, the term “margin of appreciation” refers to the “room for manoeuvre the Strasbourg organs are prepared to accord national authorities in fulfilling their obligations under the European Convention on Human Rights” (Greer 2000).

be confident that the EctHR uses EuC in certain judgments, and legal scholars have identified some language associated with the use of EuC, there remains a significant lack of clarity about the nature of this important legal method. This lack of clarity stands in contrast to research on other, related legal methods of interpretation, such as “margin of appreciation,” which have somewhat clearer definitions and are more obvious when employed by the Court.

In our analysis, we treat the EuC as a latent variable that we must measure, much like other latent concepts that social scientists often measure with measurement models (e.g., democracy, rule of law, and ideology) and we empirically uncover the nature of EuC using the texts of the EctHR’s judgments.

The Law and Politics of EctHR Decision-Making

Lawyers are not alone in having studied judicial decision-making. Political scientists, too, regularly study courts and judicial decision-making. While political scientists often view decision-making processes to be just as important as the outcomes of those processes – e.g., when assessing the state of democracy (Coppedge 2002; Munck and Verkuilen 2002) – when analyzing courts, they have been less likely to focus on how courts make decisions and the nature of legal argumentation that they use. Instead, their focus has been on the outcome of the judicial decision-making process – e.g., the votes of judges, themselves (see e.g. Epstein, Landes, and Posner 2012; Martin and Quinn 2002; Segal and Spaeth 1993; Voeten 2008), and the policy impact of the judicial rulings (Hafner-Burton and Tsutsui 2005; Mathew D. McCubbins, Noll, and Weingast 1987; Matthew D. McCubbins, Noll, and Weingast 1989).

Unlike in the contexts where political scientists have frequently studied rules and procedures, such as legislatures where the rules of procedures and decision-making tools are often written down, in the judicial context, the nature of the legal methods of interpretation used by judges to make decisions are not always clear and their use is more likely to be informal. Courts do not always provide detailed reasoning, and when they do, they do not always explicitly label or name the interpretative method that they have employed. This leaves room for legal scholars to debate the nature of the legal decision-making process and the nature of the legal methods employed by judges. However, it is difficult to study these legal decision-making tools systemically (but see Howard and Segal 2002, Helfer and Voeten 2021; Stone Sweet, Sandholtz, and Andenas 2021).

The empirical literature on the ECtHR, and international courts more generally, comes mostly (but not entirely) out of the political science tradition and is outcome-oriented. Generally speaking, it fits into one of the following categories: the study of judicial decisions and the prediction of outcomes (Aletras et al. 2016; Medvedeva, Vols, and Wieling 2020; Voeten 2008, 2021), the explanation of variation in compliance of judgements (Grewal and Voeten 2015; Hillebrecht 2009, 2012, 2014; Panke 2020; Stiansen 2019, 2021), and the study of citizens’ support for courts and the effect that support has on court rulings and compliance (Cichowski 2006; Dinas and Gonzalez-Ocantos 2021; Madsen 2020; Stiansen and Voeten 2018).² But none of this literature has explored the nature of legal argumentation in EctHR decision-making or its relationship to any of the outcomes of interest to these studies. While we do not explicitly look at the relationship between EuC and outcomes with respect to compliance or legitimacy in our work either, we argue that we take a first step – the conceptualization of EuC – towards understanding how legal reasoning and the use of this particular method of legal interpretation can impact outcomes of interest to both quantitative

² But see Lupu and Voeten (2012) and Helfer and Voeten (2014) for exceptions.

social scientists and legal scholars interested in the EctHR. Thus, the focus of our project is primarily to understand the nature of this important tool of legal interpretation.

Research Design

To study the nature of EuC, we turn to the analysis of the legal texts that the Court produces on the merits of a case (i.e., excluding admissibility), namely the judgments. Our analysis consists of six steps. It is focusing on Grand Chamber judgments that run through the end of 2019. As of the end of December 2019, there were 465 final Grand Chamber judgments. Cases decided by the Grand Chamber are the most important and consequential of cases adjudicated by the EctHR and they are heard by 17 judges, including the Court's President and Vice-Presidents. Cases reach the Grand Chamber either after judges in a Chamber proceeding (consisting of 7 judges) have issued a ruling and one of the parties involved requests a referral to the Grand Chamber or if the Chamber, because of the gravity/importance of a case, relinquishes jurisdiction to the Grand Chamber. Thus, by focusing on Grand Chamber judgments, we cover the most important decisions made by the EctHR, although we will ultimately expand our analysis to cover Chamber judgments, as well.

Our process is as follows:

1. *Identify the parts of the EctHR judgments where the EctHR lays out its comparative data and legal reasoning and which could potentially contain EuC reasoning.*
2. *Develop a Human Coding scheme on the basis of our current understanding of EuC*
3. *Hand code a random sample of Grand Chamber cases*
4. *Train a classifier to uncover other possible instances of EuC in cases that were not hand-coded.*
 - a. *Examine (random samples of) instances where the classifier identifies consensus to determine whether EuC is truly present or whether we have an instance of a false positive.*
 - b. *Rerun classifier to develop a definitive list of EuC language in the judgments of the Grand Chamber*
5. *Examine correlates of the Court's use of EuC language in both hand-coded and computer-coded sample: e.g., time, the government involved, the Articles of the ECHR in question (i.e., the human rights at issue), the Judges hearing the case, etc.*

We explain each of these steps in turn.

Identify the relevant parts of the EctHR judgments: Similar to previous work looking at EctHR case documents, we do not analyze the entire case. Specifically, we restrict the case documents to The Law sections and The Relevant Law and Materials sub-section (which sometimes appears in The Procedure or The Facts sections) as these are the parts of the cases that contain the legal argumentation on behalf of the Court. Additionally, we are only interested in analyzing parts of the cases where the Court is "speaking". Within a case, there are arguments from both sides, from third-party interveners, from individual judges (separate opinions) as well as the analysis from the Court. EuC reasoning can only logically be present in the section of the judgment where the court is "speaking," (i.e., text where the Court is

developing its own reasoning rather than reciting the arguments of the parties to the dispute or of interventions by third parties). To identify when the Court is speaking, we analyzed the structure of the cases as well as built a machine learning classifier based on the language the Court used in the Grand Chamber.

Develop conceptualization of EuC and a plan for Human Coding: We first require significant input from humans, in particular highly trained human rights lawyers, to conceptualize what is EuC language. We relied upon our team of academic lawyers who specialize in European human rights law to identify certain language and patterns indicative of the EuC method. Nevertheless, a significant challenge facing our research is that even these highly trained human rights lawyers who study decision-making in the EctHR do not always agree on what constitutes the use of EuC by the Court in cases where the Court is not explicit in this respect. This problem is not unique to the study of EuC but crops up whenever studying and attempting to measure imprecisely defined latent concepts (e.g., democracy, hate speech, populist rhetoric to name a few). While some instances of EuC language are easy to identify, even for non-trained coders, there are instances where even highly trained lawyers do not agree.

We tackle this problem in two ways. First, we decide to focus on the occurrences of *EuC language* rather than the use of EuC, itself. The difference is subtle; while in most instances, when the Court uses language associated with EuC, it is, in fact, using EuC reasoning. However, there may be instances in which language is used but not necessarily to the tool of EuC, itself. The lawyers in the project are more readily able to agree on what constitutes EuC language than whether the Court uses EuC, itself, in specific instances. Second, we develop a tagging scheme to highlight this language at the paragraph level within the judgments.

To identify EuC language, we focus on identifying two factors: one, references to external sources within the cases; and two, the use of EuC language. First, it is important to identify references to external sources as sources that are extraneous or external to EctHR text a necessary, but not sufficient, precondition for the use of EuC. The EuC method draws on domestic human rights standards (set by domestic courts or by national law makers for instance), the practice of international institutions, or even the practice of states other than the ECHR parties for various comparative analysis purposes, including to explore whether these sources and instances of practice of states or organizations (such as the EU) justify setting a common European standard that should be the practice of all ECHR parties as well. To do this, not only did we identify external sources, but we identified types of external sources. These sources ranged from domestic practice of the ECHR parties, domestic practice of third states, sources stemming from EU legal order, sources stemming from the CoE system outside of EctHR text and case law, other sources including hard or soft international law or case law of third international courts, and others (such as scientific evidence or a bibliography). Together, these forms of external sources make up the universe of “ingredients” or of the “constitutive elements” for consensus to be identified – yet they are a necessary but not a sufficient condition, because references to such sources can be done for different (albeit possibly contiguous) interpretive purposes besides EuC, such as evolutive interpretation that does not involve consensus or systemic integration, that is, alignment of the ECHR interpretation with the rules of and the trends prevailing within international law.

After identifying references to external sources, which are a necessary precondition for EuC, the lawyers focused on identifying EuC language by tagging paragraphs within judgments as follows: Tag 1 is used for the explicit use of the term “consensus” by the Court itself, which

would (almost) always indicate that the Judges are, indeed, employing consensus. We tag the word alone and when used on its own or in conjunction with other terms (e.g., “European consensus”, “Scientific consensus”, “emerging consensus”, and “international consensus”). The second category – Tag 2 – indicates the use of quasi-explicit language and is a little more nuanced than Tag 1, but still clearly indicative to lawyers that the Court is employing EuC language. Examples of Tag 2 language include phrases such as “the vast majority of [states/countries/member states/contracting parties]”, “no uniform approach”, “trend”, “a significant number of [states/countries/member states/contracting parties]”. These phrases indicate that the Court is engaging in comparative analysis of the practice of the ECHR state parties or of other states or more generally types of sources (e.g., international law standards) to determine whether a new consensus around the existence of a human right is emerging. The final tag – Tag 3 – is given to paragraphs when Tag 1 or Tag 2 language is absent, but the Court refers to extraneous sources and the context suggests to the lawyer-coders that the Court could be using EuC logic. Tag 3 paragraphs were then examined by the whole team of lawyers and further broken down into those paragraphs where the lawyers were more confident that EuC is being employed and those where it is not. This tagging scheme was developed by the team of lawyers after a careful reading of a random sample of Grand Chamber cases.

Hand code a random sample of Grand Chamber cases: Having developed the coding scheme, the team of lawyer-coders applied it to a random sample of 236 of the Grand Chamber judgments. To ensure the highest ease of functionality across our team, we did all the hand coding in Microsoft Word with the use of highlighting and commenting functions. Then, we imported these documents in R for text processing and quantitative analyses. We found that this process worked well across different skill levels, and it allowed the structure of the documents to be preserved (such as the different section titles, font size, indentation, etc.).

Develop and apply a dictionary of terms that are definitely or highly likely to be indicative of the Court applying EuC: After hand-coding 236 EctHR Grand Chamber judgments, we developed a dictionary analysis based on the language used to identify paragraphs as either Tag 1 or Tag 2. The goal of this dictionary analysis is threefold. First, in building the dictionary, we can determine whether there are additional instances of Tag 1 or Tag 2 language that the human coders did not code, either because they missed the language or because, despite the presence of the language, the coders felt that, in the particular instance, it was not indicative of consensus. Second, we can determine whether there is additional language that regularly occurs in paragraphs tagged as Tag 1 or 2 that we had not thought of, and which should be included in the dictionary. Finally, after refining the dictionary, we apply it to the entire corpus of Grand Chamber and, eventually, Chamber decisions to identify all instances of the use of explicit and quasi-explicit consensus language, and not just instances in those randomly chosen hand-coded cases. We report the terms included in our dictionary in the Appendix.

Train and run a classifier to uncover possible instances of EuC not identified by the dictionary. The dictionary can readily identify instances of Tag 1 and Tag 2 language, but by definition, it is not possible to include Tag 3 language in a dictionary. This language is more implicit and context dependent. By building a classifier, we can determine whether, based upon systematic use of language, we can uncover instances of paragraphs that lawyers would agree could be tagged using Tag 3. We start by training the classifier on the sample of hand-

coded texts.³ We then refine it and run it out-of-sample on the whole Grand Chamber corpus. We then sample the paragraphs that the classifier has indicated to contain consensus. These paragraphs were then given to the lawyers to determine whether they should be coded as containing any type of consensus language. Each paragraph was given to multiple lawyers to code. Coding was blind so that lawyers did not know how other lawyers had coded the same paragraph. This information is then used to evaluate the functioning of the dictionary analysis as well as to further refine the classifier, which is finally applied to the whole corpus. The combined results of the dictionary and classifier analyses give us a comprehensive picture of the use of European consensus within the EctHR Grand Chamber.

Examine correlates of the Court's use of EuC language. In a final step, we can use the results of analyses to identify correlates of EuC. Scholars of the EctHR have noted that the Court's use of EuC has increased over time. Others have hypothesized that the Court may be more likely to use the tool when cases involve certain member states. For example, the Court may use EuC to demonstrate to certain more skeptical or recalcitrant governments (e.g., the United Kingdom or Russia) that many or most other states hold a different view about the human right in question. Likewise, certain Judges or Judges from certain states may be more likely to engage in EuC reasoning. And EuC might be used more with regard certain human rights than others.

Findings

As described above, our first task was to develop our coding scheme and apply it to a sample of Grand Chamber judgments. We ultimately use the tagged text from this coding process to train our classifier. But before we train the classifier, we analyze the results in our hand-coded sample and use our tagging and dictionaries to perform some basic descriptive analysis of the presence of EuC language. Because the hand-coded were all read by academic human rights lawyers, we are most confident about our coding in these cases.

Here we uncover several key results. First, as suspected, the Court's use of EuC has increased steadily over time. Figure 1 shows this steady increase.

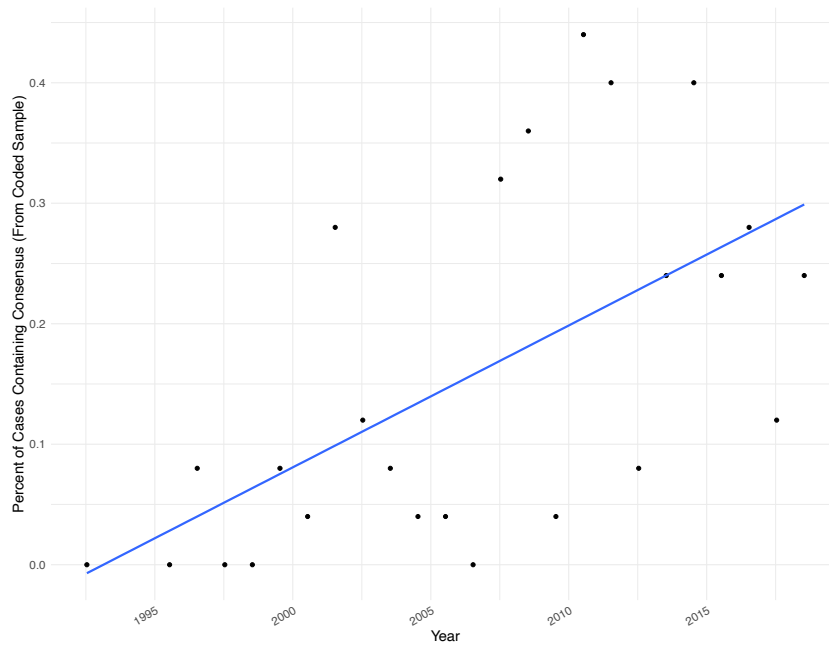


Figure 1: Percentage of Cases Containing Consensus Language By Year

The figure shows the percentage of hand-coded cases containing consensus language (y-axis) in each year (x-axis). Each dot represents a year for which we have hand-coded cases and the blue line provides the linear trend. The percentage of cases in sample in which consensus language is clearly increasing over time since the mid 90s.

Our data also show that some states are more likely to be subject to judgments using EuC than others. Figure 2 presents a bar plot of the number cases in our sample using consensus language by respondent state. It also shows the total number of cases in our sample in which that state was a respondent state.

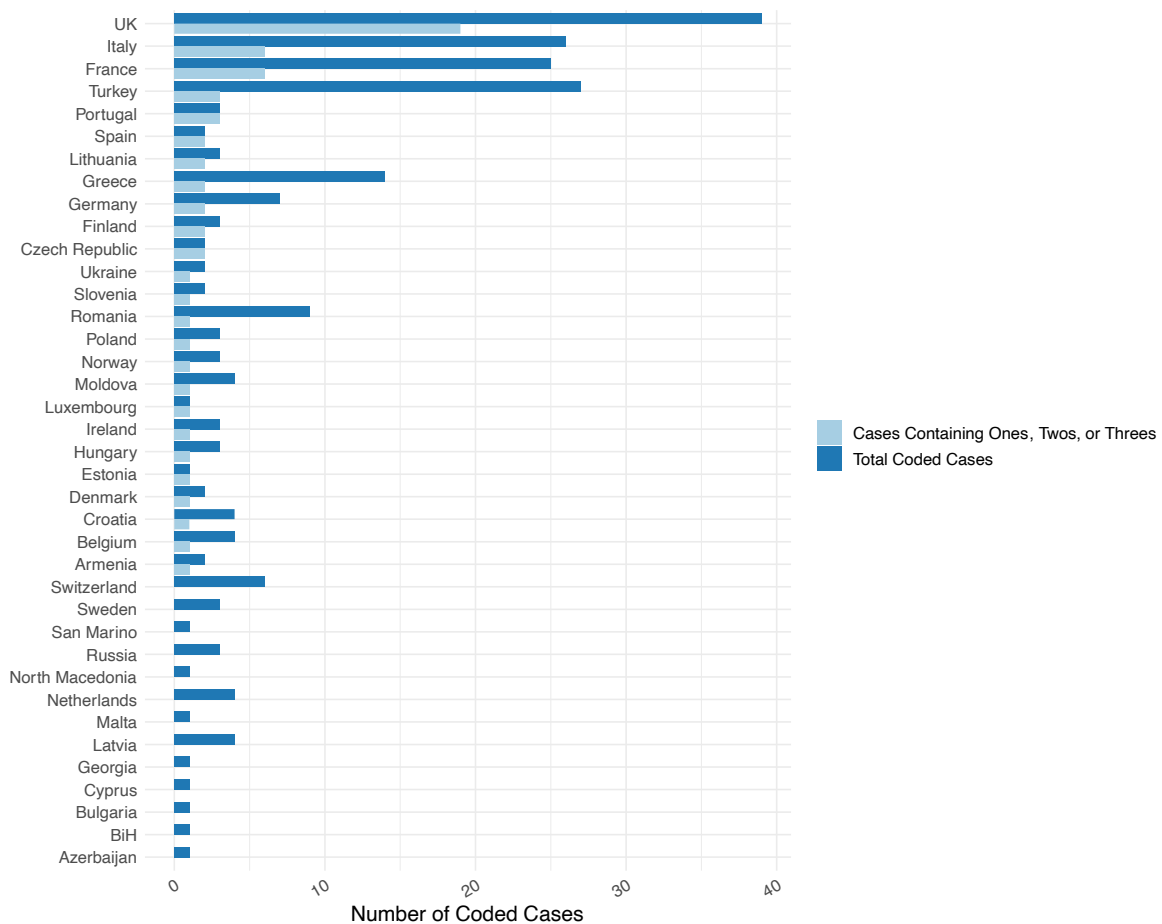


Figure 2: Frequency of Total Cases and Cases Using Consensus by Respondent State in Hand-Coded Sample

Figure 2 shows that the United Kingdom is both the state with the most cases in which it was the respondent and also the respondent country with the most number of judgments using consensus. The United Kingdom is followed by Italy, France and Turkey. Many but certainly not all of the countries at the top of the list are stable western European democracies with highly developed systems of rule of law. That these countries should be involved in high number of cases using consensus is not surprising. Because of their highly developed legal systems and the fact that they are large states, they are simply involved in a large number of cases as the dark blue bars show. Additionally, many of cases that they are involved in are likely to involve highly contentious and unsettled issues that have a tendency to divide and spark controversy among western states and within western legal systems. These are precisely the type of case where the EctHR has used consensus.

Next, we examine the types of cases in which the Court uses consensus language. Specifically, we look at the articles of the ECHR that consensus is used in conjunction with. The bar graph in Figure 3 shows the number of times consensus language occurs with respect to a particular article of the treaty (light blue bar) and the total number of cases in our sample involving that article (dark blue bar).

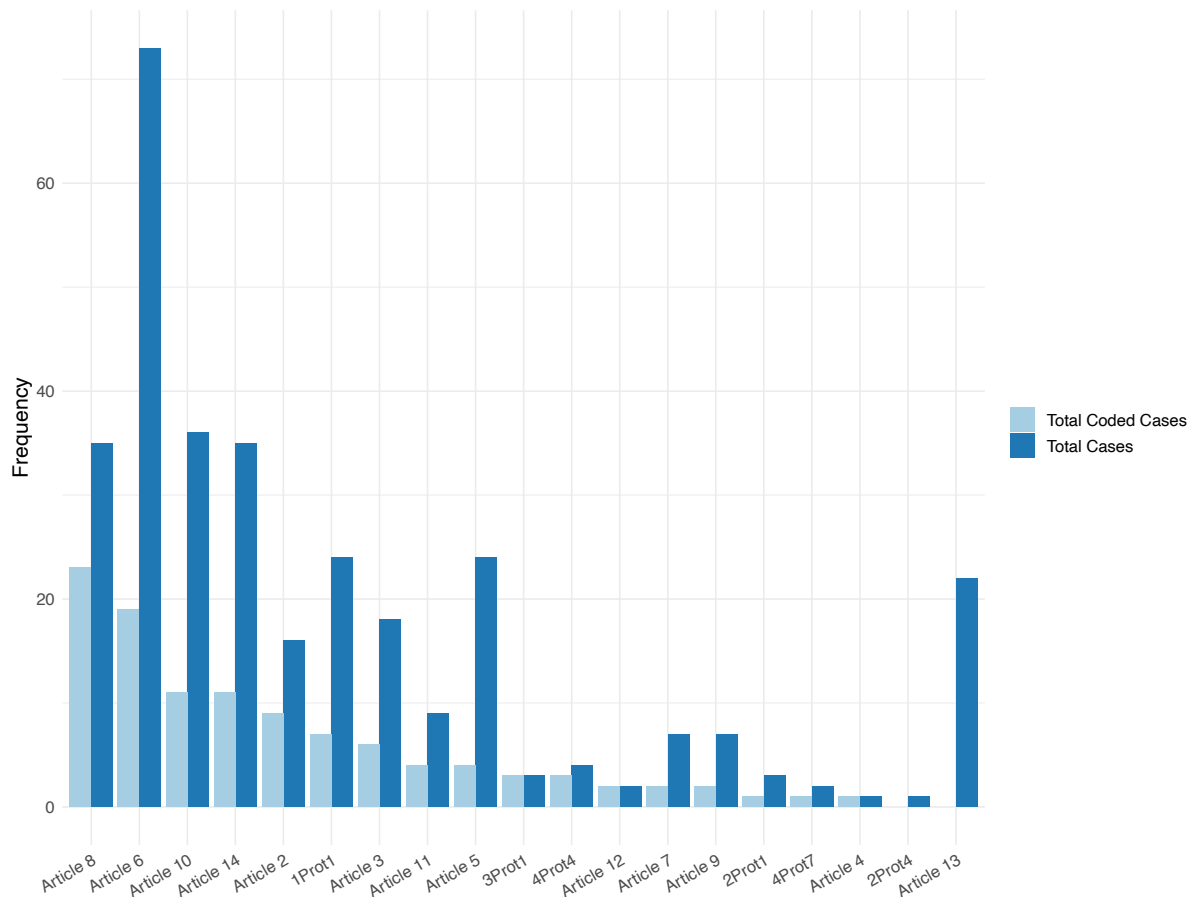


Figure 3: Frequency of Consensus by ECHR Article in Hand-Coded Sample.

In our sample, consensus is most often associated with Article 8, followed by Articles 6, 10, 14 and 2. Concentrating on the more common articles (those with more than 10 judgments in our sample), Art. 8 and 2 experience the highest percentage of judgments using consensus. Art. 8 involves the right to respect for private and family life, Art. 6 sets out the right to fair trial, Art. 10 is about freedom of expression, Art. 14 establishes the prohibition of discrimination, and Art. 2 covers the right to life.

The final analysis that we conduct in our hand-coded sample examines which states contribute to the building of consensus. In addition to tagging consensus language, our team of lawyers also highlighted mentions of state names (and other actors) that were used by the Court in a comparative analysis of law. A comparative analysis of state practice is a fundamental building block of consensus and necessary to establish whether a consensus exists, either among signatory states of the ECHR or more broadly. The names of all states, except the respondent state in the case, were highlighted. Figure 4 shows the number of times that a state name is highlighted in our sample of cases.

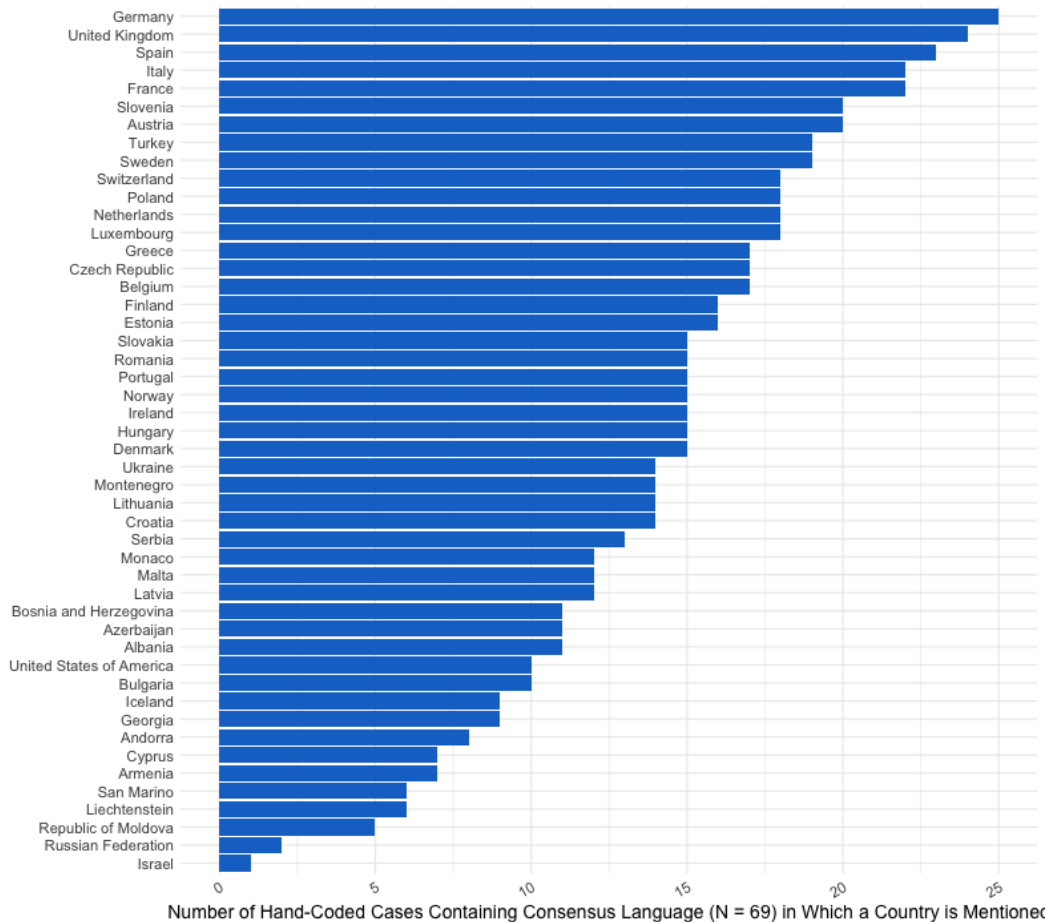


Figure 4: Frequency of Mentions of Country Names Highlighted with Hand-Coded Sample

Here we see that the countries that are mentioned most often in judgments when the Court conducts comparative analysis are Germany, UK, Spain, Italy and France. These are the states that the Court most frequently looks to when seeking to establish whether a consensus exists by means of explicit references to the practice at the domestic level of named CoE member states. It is interesting that the countries that are most often respondent states in cases using consensus language are also many of the states mentioned most often in the comparative analysis of state practice seeking to establish it. Of non-signatory states, only the United States and Israel are mentioned. Other large western democracies (e.g., Canada, Australia, Japan, New Zealand) are not explicitly mentioned.

Classifier Analysis We start training our machine learning classifier using the 237 hand-coded cases. These cases are sufficiently representative to understand general trends and the overall performance of the classifier. We run multiple different classifiers with various forms of text preprocessing and transformation. We run all analysis at the level of the paragraph. Due to the strong imbalance of consensus labelled paragraphs to non-consensus labelled paragraphs, we up-sample consensus language and down-sample non-consensus language in our training set. Ultimately, this leads us to a training set of a ratio of around 1:2 consensus to non-consensus language paragraphs, which we randomize into 10 different training and test sets to run the different machine learning classifiers. This allows us to get a baseline performance of the different models while ensuring that specific paragraphs are not driving the results.

The classifier includes all tags in the training and test sets (Tag 1, Tag 2, and Tag 3) where it tries to predict whether the paragraph contains consensus or not (EuC language is operationalized as a binary variable). The best performing model is the support-vector machine model (SVM) with unigrams and the least amount of text preprocessing. Our average F-1 score for the iterations of classifiers, which is an overall performance measure, is 0.58. This is rather low, however, our overall specificity scores are high (a mean of close to 1 at 0.98). This is encouraging as it means that if the classifier identifies a paragraph as consensus, it is almost always contains consensus language as identified by the lawyers. This high specificity score is coupled with a low precision score, meaning that there are also a high number of false positives (non-consensus language identified as consensus by the classifier). It is these false positives that drive down our F-1 scores. The persistent presence of a high number of false positives means that it remains difficult to use our classifier results to answer further substantive questions.

Discussion and Conclusion

EuC is an increasingly important tool of legal interpretation that is employed by EctHR judges when making decisions regarding contentious human rights issues across Europe. Judges take time and effort to craft legal arguments in hopes that their legal reasoning, not only the outcome of a particular case, will have an impact on human rights law in the contracting parties across Europe. Solid legal reasoning, and the use EuC in particular, may confer legitimacy on an EctHR ruling and make it more likely that actors in national legal systems take the judgment seriously when making decisions at home.

Despite its importance, EuC has proven a difficult legal method to study, to conceptualize, and certainly to quantify. Even our team of highly qualified academic human rights lawyers have disagreements over what does and does not constitute consensus language when this is not (semi-)explicit in the text of the judgment. EuC language cannot be identified without a lot of qualitative input, both because it is such a rare event within the text of judgments, and also because it is very context dependent. We created a coding scheme that measures three forms of EuC language: explicit consensus, semi-explicit consensus, and implicit, meaning the text contains reference to extraneous sources to suggest EuC logic. With this coding scheme, we manually coded half of all Grand Chamber cases. With these coded cases, and our dictionary analysis, we can confirm that the Court has increased its usage of EuC language since the 1990s, especially in the last 10 years.

We have found that EuC language is also hard for a machine learning classifier to identify. We are able to obtain high levels of specificity, meaning that the classifier tags almost all instances of consensus language. However, this comes with the caveat that the classifier also tags many false positives.

References

- Aletras, Nikolaos, Dimitrios Tsarapatsanis, Daniel Preoțiu-Pietro, and Vasileios Lampos. 2016. "Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective." *PeerJ Computer Science* 2: e93.
- Cichowski, Rachel A. 2006. "Courts, Rights, and Democratic Participation." *Comparative Political Studies* 39(1): 50–75.
- Coppedge, Michael. 2002. "Democracy and Dimensions: Comments on Munck and Verkuilen." *Comparative Political Studies* 35(1): 35–39.

- Cox, Gary W., and Matthew D. McCubbins. 2005. *Setting the Agenda: Responsible Party Government in the U.S. House of Representatives*. Cambridge ; New York: Cambridge University Press.
- . 2007. *Legislative Leviathan: Party Government in the House*. 2nd ed. Cambridge ; New York: Cambridge University Press.
- Dinas, Elias, and Ezequiel Gonzalez-Ocantos. 2021. “Defending the European Court of Human Rights: Experimental Evidence from Britain.” *European Journal of Political Research* 60(2): 397–417.
- Dzehtsiarou, Kanstantsin. 2015. *European Consensus and the Legitimacy of the European Court of Human Rights*. Cambridge: Cambridge University Press.
<http://ebooks.cambridge.org/ref/id/CBO9781139644471> (January 28, 2022).
- Epstein, Lee, William M. Landes, and Richard A. Posner. 2012. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice*. Harvard University Press.
<https://www.degruyter.com/document/doi/10.4159/harvard.9780674067325/html> (January 18, 2022).
- Finnemore, Martha, and Kathryn Sikkink. 1998. “International Norm Dynamics and Political Change.” *International Organization* 52(4): 887–917.
- Gilardi, Fabrizio. 2013. “Transnational Diffusion: Norms, Ideas and Policies.” In: Carlsnaes, Walter; Risse, Thomas; Simmons, Beth A. (eds.). *Handbook of international relations*. Thousand Oaks: SAGE, Second Edition, 453-477.
- Greer, Steven. 2000. *The Margin of Appreciation: Interpretation and Discretion under the European Convention on Human Rights*. Strasbourg: Council of Europe Publishing.
- Grewal, Sharanbir, and Erik Voeten. 2015. “Are New Democracies Better Human Rights Compliers?” *International Organization* 69(2): 497–518.
- Hafner-Burton, Emilie M. 2021. “Elite Decision-Making and International Law: Promises and Perils of the Behavioral Revolution.” *AJIL Unbound* 115: 242–47.
- Hafner-Burton, Emilie M., D. Alex Hughes, and David G. Victor. 2013. “The Cognitive Revolution and the Political Psychology of Elite Decision Making.” *Perspectives on Politics* 11(2): 368–86.
- Hafner-Burton, Emilie M., and Kiyoteru Tsutsui. 2005. “Human Rights in a Globalizing World: The Paradox of Empty Promises.” *American Journal of Sociology* 110(5): 1373–1411.
- Helfer, Laurence R., and Erik Voeten. 2014. “International Courts as Agents of Legal Change: Evidence from LGBT Rights in Europe.” *International Organization* 68(1): 77–110.
- Hillebrecht, Courtney. 2009. “Rethinking Compliance: The Challenges and Prospects of Measuring Compliance with International Human Rights Tribunals.” *Journal of Human Rights Practice* 1(3): 362–79.
- . 2012. “Implementing International Human Rights Law at Home: Domestic Politics and the European Court of Human Rights.” *Human Rights Review* 13(3): 279–301.
- . 2014. “The Power of Human Rights Tribunals: Compliance with the European Court of Human Rights and Domestic Policy Change.” *European Journal of International Relations* 20(4): 1100–1123.
- Howard, Robert M., and Jeffrey A. Segal. 2002. “An Original Look at Originalism.” *Law & Society Review* 36(1): 113.
- Kapotas, Panos, and Vassilis P. Tzevelekos, eds. 2019. *Building Consensus on European Consensus: Judicial Interpretation of Human Rights in Europe and Beyond*. 1st ed. Cambridge University Press.
<https://www.cambridge.org/core/product/identifier/9781108564779/type/book> (January 24, 2022).

- Khosla, Madhav. 2011. "Inclusive Constitutional Comparison: Reflections on India's Sodom Decision." *American Journal of Comparative Law* 59(4): 909–34.
- Laver, Michael, Kenneth Benoit, and John Garry. 2003. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97(2): 311–31.
- Lupu, Yonatan, and Erik 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): 413–39.
- Madsen, Mikael Rask. 2020. "Two-Level Politics and the Backlash against International Courts: Evidence from the Politicisation of the European Court of Human Rights." *The British Journal of Politics and International Relations* 22(4): 728–38.
- Martin, Andrew D., and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999." *Political Analysis* 10(2): 134–53.
- McCubbins, Mathew D., Roger G. Noll, and Barry R. Weingast. 1987. "Administrative Procedures as Instruments of Political Control." *Journal of Law, Economics, & Organization* 3(2): 243–77.
- McCubbins, Matthew D., Roger G. Noll, and Barry R. Weingast. 1989. "Structure and Process, Politics and Policy: Administrative Arrangements and the Political Control of Agencies." *Virginia Law Review* 75(2): 431.
- Medvedeva, Masha, Michel Vols, and Martijn Wieling. 2020. "Using Machine Learning to Predict Decisions of the European Court of Human Rights." *Artificial Intelligence and Law* 28(2): 237–66.
- Moyer, Laura P., and Susan B. Haire. 2015. "Trailblazers and Those That Followed: Personal Experiences, Gender, and Judicial Empathy: Trailblazers and Those That Followed." *Law & Society Review* 49(3): 665–89.
- Munck, Gerardo L., and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35(1): 5–34.
- Palmer, Barbara. 2001. "Women in the American Judiciary: Their Influence and Impact." *Women & Politics* 23(3): 91–101.
- Panke, Diana. 2020. "The European Court of Human Rights under Scrutiny: Explaining Variation in Non-Compliance Judgments." *Comparative European Politics* 18(2): 151–70.
- Peat, Daniel C. 2021. "The Tyranny of Choice and the Interpretation of Standards: Why the European Court of Human Rights Uses Consensus." *NYU Journal of International Law and Politics* 53(2): 381–432.
- Reh, Christine, Adrienne Héritier, Edoardo Bressanelli, and Christel Koop. 2013. "The Informal Politics of Legislation: Explaining Secluded Decision Making in the European Union." *Comparative Political Studies* 46(9): 1112–42.
- Riker, William H. 1986. *The Art of Political Manipulation*. New Haven: Yale University Press.
- Segal, Jeffrey Allan, and Harold J. Spaeth. 1993. *The Supreme Court and the Attitudinal Model*. Cambridge ; New York, NY: Cambridge University Press.
- Shepsle, Kenneth A. 1979. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models." *American Journal of Political Science* 23(1): 27.
- Stiansen, Øyvind. 2019. "Delayed but Not Derailed: Legislative Compliance with European Court of Human Rights Judgments." *The International Journal of Human Rights* 23(8): 1221–47.
- . 2021. "Directing Compliance? Remedial Approach and Compliance with European Court of Human Rights Judgments." *British Journal of Political Science* 51(2): 899–907.

Stiansen, Øyvind, and Erik Voeten. 2018. "Backlash and Judicial Restraint: Evidence From the European Court of Human Rights." *SSRN Electronic Journal*.

<https://www.ssrn.com/abstract=3166110> (September 26, 2021).

Tomz, Michael. 2008. "Reputation and the Effect of International Law on Preferences and Beliefs." Working Paper, Stanford, CA: Stanford University. Available at:

<https://tomz.people.stanford.edu/sites/g/files/sbiybj4711/f/tomz-intllaw-2008-02-11a.pdf>.

Accessed 2 June 2022.

Voeten, Erik. 2007. "The Politics of International Judicial Appointments: Evidence from the European Court of Human Rights." *International Organization* 61(04).

http://www.journals.cambridge.org/abstract_S0020818307070233 (September 26, 2021).

———. 2008. "The Impartiality of International Judges: Evidence from the European Court of Human Rights." *American Political Science Review* 102(4): 417–33.

———. 2021. "Gender and Judging: Evidence from the European Court of Human Rights." *Journal of European Public Policy* 28(9): 1453–73.