SEARCHING FOR A REFERENCE: USING AUTOMATED TEXT ANALYSIS TO STUDY JUDICIAL COMPLIANCE*

by

KATARÍNA ŠIPULOVÁ**, HUBERT SMEKAL***, JOZEF JANOVSKÝ****

The concept of judicial compliance has attracted plenty of attention in the last two decades. Yet, despite the growing scholarly interest, important research questions remain largely unresolved. This is partly due to the persistent use of unsystematic research, built on the cherry picking of cases. The content of only a few well-known judgments has been thoroughly examined, and the rest remains largely ignored by the legal scholarship. The aim of this article is to introduce a sketch of a new three-level approach for improving research on judicial compliance in a multi-level arena. We show how the use of automated text analysis in combination with more traditional legal methods might shed more light on the concept of judicial compliance and judicial dialogues. We explain the procedure of the automated collection of data and their coding and also point out the risks of using automated text analysis when studying judicial compliance. The approach is demonstrated on a single case study of the use of European Court of Human Rights rulings by Czech apex courts. This study assesses how often and in what way the domestic courts engage with the European Court of Human Rights case law.

* The authors would like to express their gratitude to the reviewers and MUJLT editor Marek Pivoda, who immensely helped to improve the quality of the manuscript. The research leading to this article has received funding from the Czech Science Foundation under Grant Agreement no. 16-09415S, Panel P408 ("Beyond Compliance – Domestic Implementation of International Human Rights Case Law").

** katarina.sipulova@law.muni.cz, Senior Researcher, Judicial Studies Institute, Law Faculty, Masaryk University, Czech Republic.

*** hsmekal@fss.muni.cz, Senior Researcher, Faculty of Social Studies, and Judicial Studies Institute, Law Faculty, Masaryk University, Czech Republic.

**** janovsky@mail.muni.cz, Law Faculty, Masaryk University, Czech Republic.
KEY WORDS
Automated Text Analysis, Compliance, European Court Of Human Rights, National Courts, References

1. INTRODUCTION
The last two decades have seen a proliferation in various international judicial bodies.¹ The phenomenon does not solely include a numerical increase of judicial bodies, but also an increase in their influence and engagement in both domestic and international politics.² This applies especially to the bodies that hold the benefits of compulsory jurisdiction, high levels of independence from national governments, and a big caseload thanks to individual petitions.³ Yet, these courts usually lack the power to oversee and foster the execution of their own decisions. With law as the dominant regulatory tool in modern states, the judiciary lays the foundations of its power on the possession of legal expertise⁴ and on the reputation of learned interpreters of law.

The European Court of Human Rights (“ECtHR”) is no exception. The literature generally acknowledges the importance of the ECtHR, calling it the most successful international adjudication and enforcement regime for the protection of human rights.⁵ The ECtHR’s judgments influence the functioning of all branches of power, with possible significant intrusion into national balance of powers.⁶ The Committee of Ministers supervises the implementation of adverse judgments at the national level.⁷ Yet,
the ECtHR still depends largely on the cooperation of domestic courts, which are vital for the execution of its judgments.

Several noteworthy questions were raised by existing research, e.g., Do domestic courts function as transmission belts for the EctHR? What is the form of judicial compliance? Do domestic courts engage in judicial dialogue with Strasbourg? If so, in what form? Why, or in what instances, do courts refer to the ECtHR? How did the reference (compliance) patterns evolve over time? However straightforward these questions may appear, we still lack clear answers.

The research on interactions between national and international courts has become voluminous, yet considerable gaps remain. While it is widely acknowledged that international courts’ case law is reflected by domestic courts, we do not know exactly how or to what extent. Legal papers have typically built on a rather low number of the most important cases and overlooked the big picture of the ordinary, but the most frequent cases. We therefore do not know much about how often the domestic courts use international case law and what its typical use is in daily practice. Recent enormous developments in technology have significantly improved accessibility to the data, the process of data collection, coding and analysis. In our project, we utilize these developments to enrich traditional legal research methods and examine the core research question of whether and how domestic courts comply with ECtHR’s case law. Accordingly, this paper aims to introduce a framework for the systemic research on the use of international case law by domestic courts.

Our three-level framework utilizes both manual and automated methods of data collection and coding, as well as quantitative and qualitative methods of analysis. Thus, it does not rely only on a classical, detailed legal analysis of the most important cases nor does it employ only traditional hand coding, but it builds on more cases and also uses automated text analysis. We find this rich mix of methods of data collection, coding and analysis as the most promising way of acquiring knowledge about

---

7 Although the Committee of Ministers lacks ‘hard’ execution powers, it can refer a question to the ECtHR on whether a Party has failed to fulfil its obligation to abide by the final judgment of the Court (see Art. 46 para. 4 of the European Convention on Human Rights, as amended by Protocol No. 14).

the phenomenon under study. However, based on our experience with carrying out the research project on judicial references between the national and international levels, we warn that the automated methods need careful validation because many false positives and negatives can occur.

This paper describes the advantages of the three-level framework and shows its potential to improve current scholarship in the field of judicial compliance with international human rights case law. The paper focuses on the methodological aspects of the research on the use of international case law by domestic courts, mainly on the problem of how to approach judicial compliance both quantitatively and qualitatively. The paper covers issues of sampling, data collection and text recognition, and coding, while the data analysis is only briefly introduced. Snapshots of the preliminary results of the use of ECtHR case law by Czech apex courts are presented in order to illustrate the potential of our approach; a comprehensive presentation of the results would require a book-long enterprise.

The paper first addresses the use of reference-based research in the study of domestic judicial compliance with international case law (Section 2). Then, we present the fundamentals of our three-level framework, which advances the research on how often, and how, national courts participate in the compliance exercise (Section 3). Section 4 points to the challenges and potential problems that may occur when conducting research based on our three-level framework, while Section 5 presents in detail the methodology of the project, specifically data collection and coding that incorporates automated methods. Application of the framework is demonstrated on the case of the Czech Supreme Administrative Court’s references to the ECtHR case law (Section 6). Section 7 concludes.

2. JUDICIAL COMPLIANCE AND REFERENCE-BASED LEGAL RESEARCH IN THE INTERNATIONAL SETTING
This section briefly introduces the field. First, we address the concept of judicial compliance with international human rights law and explain why the research of references is essential for its understanding. Then we

---


10 Such a book will be the final outcome of our whole research project.
demonstrate how this field of research might benefit from the use of mix methods, especially from the inclusion of automated text analysis.

We have already noted above that the increasing importance of the courts, including international courts, is widely acknowledged. Both national and international courts can benefit from their co-existence. Domestic courts may use citations of the international case law to support their own reasoning, while reminding the executive and legislative branches that they themselves agreed to and ratified the international treaty establishing the international court. Vice versa, international courts largely depend on the cooperation of domestic courts, especially those standing at the top of the judicial hierarchy. Apex courts unify domestic jurisprudence and may help with the domestication of international case law by frequently referencing it. Most of the existing research focuses on the relationship between domestic courts and the Court of Justice of the European Union (“CJEU”) or the EctHR.11

Our research focuses on the use of ECtHR case law by the Czech apex courts – the Constitutional Court, the Supreme Court and the Supreme Administrative Court. We assume that when courts use the existing case law, they then cite it. References to international case law indicate that domestic judges feel the urge to engage with international judgments. A reference does not automatically mean judicial compliance (in the sense of conformity, see infra), e.g. in cases when domestic judges expressly oppose an international judgment. However, even such a reference provides an important hint that domestic judges take international case law seriously, because they consider it important to expressly acknowledge their opposition to the direction international case law is taking.

When domestic courts approvingly refer to the international case law, then they significantly contribute to compliance with the case law. Usually, compliance is understood as a state of conformity of practice or policy with legal norms.12 It has been widely argued that national courts belong among the most important compliance partners of international courts.13 Frequent references to international case law indicate its domestication by national

---

judges and subsequently higher compliance. Thus, references by national courts help in building the legitimacy of international courts and their case law.\(^{14}\)

Yet the systematic tracing of references has only slowly started making its way to the study of interactions between national and international levels. The automated reference recognition has developed remarkably in legal informatics, but it typically remains confined within the boundaries of one legal system, national or international,\(^{15}\) and unlike our project does not try to connect the two levels. Moreover, the use of the automated reference recognition has not significantly infiltrated the general legal research.\(^{16}\) Legal research still relies on traditional methods of analysis and views new trends with suspicion.\(^{17}\)

We position our paper in the discussion on the use of references to international case law by national courts, rather than in the discussion on automated reference recognition. However, for the latter debate, or for the field of legal informatics, our three-level framework might serve as a manifestation of the use of automated methods in the larger research project. Moreover, it points to the practical difficulties that arise when carrying out such research (see Sections 3, 4 and 5).


Systematic research using references in a large number of cases would not have been possible without the use of modern technologies. Courts worldwide produce immense amounts of decisions. It became impossible a long time ago for scholars to fully process such a sheer volume of data (read judgments, analyse them and comment, systematize, etc.). The use of computers can, however, significantly help, at least in some stages of research. Huge advancements in information technology over the last few decades have made a previously unimaginable quantity of data available and ready to be analysed, and processed for practical use.\(^\text{18}\)

Researching judicial compliance and references to ECtHR case law requires access to domestic courts’ databases, as well as a firm understanding of the wider context in which these references were used. This places high demands on both the understanding of language and the legal background, including the more hidden, shadowy life of decisions embedded in the social and cultural particularities of every society. The use of mixed methods should bring comprehensive understanding to the problem at hand, uncovering both the big picture and also more fine-grained processes. On one hand, solely quantitative research of references does not tell us much about how important the role of the international case law is in domestic judgments, if it is followed, and what long-term consequences it brings. Only legal experts with deep expertise in the functioning of the apex courts can alert the research team when the data show something unexpected (but hidden to a layperson).\(^\text{19}\)

On the other hand, detailed qualitative legal studies of the most important judgments can paint an overly optimistic picture of the influence of international case law on domestic case law and overlook the (possibly low) extent of the overall use of the international case law.

Joining forces with social science methods brings huge promise for legal scholarship. Automated, computer-driven text analysis promises to reduce the costs of reading enormous collections of decisions so that we are able to explore the thus-far unreachable and unknown territory of courts’


\(^{19}\) For example, when a judge who does not have a reputation for being a human rights champion (which is a piece of information known only to the expert legal community) records a very high number of references to ECtHR case law. Only then does one check in more detail and finds that the high numbers are due to copy-pasting a short passage of text including a reference to an ECtHR judgment in decisions rejecting petitions as manifestly unfounded.
decision-making. Applying these methods to case law seems to be especially useful in jurisdictions where apex courts have very weak filtering mechanisms and thus issue many thousands of decisions each year, which makes a complex academic examination close to impossible.

The most immediate inspiration for our work was Wind’s article, which used automated techniques to count the frequency of references to ECtHR case law by Scandinavian courts. Such an approach however does not say anything about how ECtHR case law is used (What was the purpose of the reference?) or what its significance is (Would the case be decided differently without the reference?). Nor does it distinguish if a domestic court used the reference in its own reasoning, or if it appeared only in the summary of the proceedings before lower courts or of the arguments of the parties. The next section thus presents an overview of our three-level approach, which also answers these important questions, and shows the benefits of including automated techniques.

3. THREE-LEVEL APPROACH TO THE STUDY OF JUDICIAL COMPLIANCE
As noted above, the cooperation of domestic courts is indispensable for international judicial bodies. International human rights norms come to life with their domestic application. In order to understand the domestic judicial compliance with international case law, we first have to identify the set of domestic decisions which use it. Then we can focus on the extent of the use of the international case law – i.e. the frequency of references to international cases in domestic decisions and its development over time. This basic descriptive statistical exercise is important for a rough mapping of the terrain, but does not tell us much about how the international case law is used by domestic judges. Specifically, we are interested in the significance of international case law (i.e. Is it important for deciding domestic cases?), to what extent domestic judges follow their international peers, how extensively they consider international case law and how carefully they refer to it. We employ a dynamic perspective, which means that we are interested in developments in all these categories over time.

For our research, we developed a more nuanced understanding of compliance and attempted to implement it with a mix of quantitative and

---

qualitative methods such as descriptive statistics (frequencies and crosstabs) and traditional legal analysis (rich description of cases in context). We argue that judicial compliance might be understood in both a broader and a narrower sense. The narrow understanding of judicial compliance reflects only the implementation of adverse ECtHR judgments against a particular country by domestic courts. In other words, narrow judicial compliance focuses only on how domestic courts implement judgments that found that the country in question violated the Convention.

We are convinced that a broad understanding of judicial compliance is better for the systemic research of international case law’s impact on the domestic judicial practice than the narrow one because it oversteps the limitation of compliance focusing solely on adverse judgments against the home country. A broad understanding enables us to include all mutual interactions in the case law of both the domestic and international levels and thus more comprehensively examine domestic judicial compliance with the international case law.

First, the macro-level encompasses all references that domestic apex courts have ever made to ECtHR case law. The macro-level asks how frequently courts refer over time to ECtHR case law. We provide the answer by measuring the annual development in frequency of domestic apex courts’ references to ECtHR case law. In other words, we record the number of references (as well as the number of decisions with references) and compare it against the population (i.e. the total number) of all decisions of the respective apex courts in order to have a rough idea how often ECtHR case law is used and how the frequency develops over time.

The automated text analysis allows us to instruct computer programs to detect the use of certain words and phrases in texts. It replaces human hand coding, but still needs some human involvement. While computers can learn clustering, whether supervised by human beings or not, the validation of the results requires precise and often time consuming human involvement. As Grimmer and Steward point out, all automated methods, due to the complexity of the language, and particularly so of legal reasoning, are based on inexact language models. The following sections

---


22 Ibid.
therefore introduce our model and validation process in detail (see particularly Section 5).23

After the quantitative macro-level analysis of all domestic apex court judgments referring to ECtHR case law, there is only a stratified sample of few hundred judgments that work with a reference ECtHR case law in their reasoning. This sample is more closely examined in order to understand what the typical mode of use of the ECtHR’s case law is and what significance it has. Do the domestic courts follow the ECtHR’s case law? To what extent do they engage with it (automatically accept, or critically discuss)? Is the ECtHR’s case law important for the outcome of the case, or would the case be decided in the same way even in the absence of the reference to ECtHR case law? This meso-level analysis relies on close reading by humans, as coding requires expert understanding of judicial interpretation (see Section 5.3).

Finally, the qualitative micro-level analysis of a few carefully hand-picked cases enables us to focus, based on the knowledge gained from the macro- and meso-level analyses, on both typical and atypical features of referencing ECtHR case law, and to evaluate the more far-reaching consequences of ECtHR case law in domestic judicial practice (see Section 5.4).

<table>
<thead>
<tr>
<th>Level</th>
<th>Main question</th>
<th>Sampling</th>
<th>Coding</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macro</td>
<td>How much do the apex courts refer to ECtHR case law? What are the most cited ECtHR cases? Which judges cite ECtHR case law?</td>
<td>All cases including references to ECtHR case law in reasoning</td>
<td>Automated with manual validation</td>
<td>Descriptive statistics (frequencies and crosstabs)</td>
</tr>
<tr>
<td>Meso</td>
<td>How do the apex courts use ECtHR case law? What is its significance for deciding domestic cases? Do domestic judges follow ECtHR case law? How detailed is the engagement of domestic judges with ECtHR case law? How carefully do domestic judges cite ECtHR case law?</td>
<td>Stratified sampling based on the macro-level sample</td>
<td>Hand-coding based on the detailed codebook</td>
<td>Descriptive statistics (frequencies and crosstabs)</td>
</tr>
<tr>
<td>Micro</td>
<td>What are the consequences of the use of the ECtHR case law for domestic legal practice? Significant importance of ECtHR case law for selected fields (e.g. discrimination) Adoption of techniques used by the ECtHR by domestic courts (e.g. proportionality test)</td>
<td>Purposive sampling</td>
<td></td>
<td>Contextually rich legal case studies</td>
</tr>
</tbody>
</table>

Table 1: Three-level Approach to Judicial Compliance24

24 Source: authors.
4. RESEARCH CHALLENGES AND POTENTIAL INACCURACIES

There are several caveats which need to be addressed in relation to research of this complexity. We identified four possible problematic factors that challenge the research aim and the results. These are: 1) reduction of the research scope to the apex courts, 2) reference to ECtHR case law as a relevant indicator of judicial compliance, 3) the problem of silent (indirect) references, and 4) identification of cases in which courts do not refer to ECtHR case law.

First, in our case study, we decided to analyse the case law of apex courts, disregarding the courts of lower instances. Lower courts play a similarly important role in bringing ECtHR case law into practice. Nevertheless, the top courts are typically seen as key actors in judicial compliance. This is especially true in the Czech context, where the Constitutional Court oversees the protection of human rights and other two apex courts are responsible for unifying the case law of domestic courts.

We therefore believe that the apex courts function as transmission belts for international human rights bodies. They transmit the respective information and good practices in two directions: towards lower domestic courts, and towards other state actors and bodies.

Furthermore, in Czechia, similarly to many other countries, obtaining databases of lower courts’ decisions is virtually impossible. Apart from the data being heavily protected, lower courts have rarely developed their own online databases. Most typically, lower courts store only written files, which makes any sort of research highly difficult. That being said, it would undoubtedly be interesting for future research to probe at least a sample study into the lower courts’ case law and their comprehension of ECtHR case law, particularly so due to the different personal characteristics of judges sitting on the courts (education, profile, academic background), or the material factors influencing the performance of courts (presence of analytical units, assistants to judges, etc.).

Second, some readers might wonder whether references to ECtHR case law are indeed a relevant indicator for assessing the role of apex courts in judicial compliance with ECtHR case law. Merely counting references

---

to ECtHR case law hardly amounts to measuring the impact and importance the courts assign that case law. Being conscious of the fact that not all references are of equal importance, we tackled this issue by incorporating a qualitative content analysis on the meso-level, looking deeper into how exactly the courts use references, in what instances, what their position in the reasoning is and their influence on the result of the case.

Third, it often happens that once the domestic court delivers a very detailed, well-reasoned judgment referring to ECtHR case law and pioneering a new line of jurisprudence, future similar cases refer only to this domestic pioneering judgment and omit the baseline ECtHR case law that originally inspired the domestic court. We are convinced, however, that the mere decision of a judge to cite or disregard a reference to ECtHR case law has a certain symbolic meaning and value and repercussion for the ECtHR’s legitimacy as perceived by domestic judges. The future research might, however, attempt to build on our results and address the “silent references” through a network analysis. It will be helpful to collect and analyse indirect references to these very important domestic pioneer cases that transmitted the ECtHR’s case law into the domestic jurisprudence for the very first time.

Lastly, some references remain unattributed. Courts might often comply with the ECtHR’s opinions without explicitly referring to its case law, either consciously, when trying to avoid controversies potentially caused by adhering to an opinion of an international body in politically and socially salient cases, 26 or unconsciously, mostly when quoting and referring to a plethora of consistent domestic law. One might argue that without including the cases in which the apex courts should have referred but did not (either intentionally or unintentionally, and either when complying or non-complying with ECtHR case law), our analysis paints only an incomplete picture. Nevertheless, we built our research on the pre-understanding that only direct, explicit references are a valid indicator of the position of domestic courts towards ECtHR case law. The reason for using a reference in a case is to strengthen the persuasive authority of the court’s reasoning. In order to add another layer of legitimacy for its

---

claims and findings, judges refer to judgments and decisions of those domestic and international bodies which they consider legitimate and authoritative. References therefore hold a strategic place in judicial reasoning and judges enhance the prestige of other institutions by incorporating references to their case law.

5. DATA COLLECTION AND CODING

Most of the existing studies treat all references as relevant arguments and reasoning sources. Such a pre-understanding fails to capture the real practice of domestic courts. Great many references appear in other parts of judgments than reasoning (especially in the introductory part of a decision summing up the facts of the case and the arguments raised by the parties to proceedings, or in separate opinions). These references do not have the capacity to influence the core dispute underlying the case. When examining judicial compliance, it is vital to filter out those references which do not appear in substantive reasoning. Still, distinguishing the placement of the reference does not tell us much about its impact on the result of the case. To tackle these significant issues, we developed a funnel-like filtering mechanism (Figure 1), which 1) cleaned our dataset of references which did not appear in a court’s reasoning, and 2) further categorised references due to their impact on the result of a case (the substance).

![Figure 1: Visualisation of Filtering/Reduction of the Unit of Analysis](image)


29 Source: authors.
In order to capture these different levels and understandings of judicial compliance, we divided our research into 3 levels: the macro-level, focusing on general patterns present in domestic case law, the meso-level, which digs deeper and inquires into the impact of references on the substance of domestic disputes, and finally, the micro-level, which offers an in-depth qualitative analysis of individual cases. The following section serves not only as an overview of the problems related to automated coding, but also offers a guidebook for future research.

5.1 IN SEARCH OF THE DATA
One of the reasons why there is no comprehensive research on the use of references to international HR bodies’ case law by domestic courts dwells, undoubtedly, in the accessibility of the data. In many jurisdictions, access to the case law of lower courts is virtually non-existent; some courts do not have online databases, and if they do, they usually do not cover older decisions. Although these courts do have their own archives, obtaining and processing actual case files is highly costly. In most European countries, the situation is a little less gruesome when it comes to apex courts (as is also the case in Czechia). Yet some difficulties still remain. As previously mentioned, our research project builds on the assumption that apex courts function as transmission belts, promoting the domestic application and compliance with ECtHR case law. A proxy for measuring this compliance is a reference to the ECtHR’s case law, i.e. a citation of respective decisions delivered by the ECtHR (for more on the composition of the reference see below). First, we therefore had to collect entire datasets of cases delivered by the top three Czech apex courts: the Constitutional Court (“CC”), the Supreme Court (“SC”) and the Supreme Administrative Court (“SAC”). All three databases are publicly accessible; nevertheless, not all of these databases allow a user-friendly download of metadata. We also realized that some courts do not publish some of their decisions online.30 We had to implement data scraping for some metadata in the cases of the CC and the SAC (subject area, judge rapporteur, etc.). The CC’s dataset lacked case file numbers in their identification, so we had to proceed with automatically obtaining the file numbers from the text of the decisions. Lastly, all obtained documents have

30 Sometimes by omission, sometimes intentionally, e.g. purely technical decisions, some decisions on recognition of judgments, etc.
been converted to UTF8 format, which, compared to .doc, does not have a structure and allows for smooth processing in the programme R. Special and lengthy attention was devoted to optimizing special characters, mostly typical for the Czech language. For the CC, our dataset encompasses 60,403 decisions\(^{31}\) delivered between 1.1.1993 and 31.12.2015. For the SC we collected 84,374 decisions delivered between 1.1.1993 and 31.12.2015. The case law of the SAC covers a shorter period, from 1.1.2003 to 31.12.2015, as the court was established only in 2002. This dataset covers altogether 39,477 decisions.

After obtaining the domestic datasets, we devoted similar attention to the ECtHR’s case law. Similarly, we data scraped the HUDOC database\(^{32}\) for all decisions and judgments delivered by the Strasbourg Court and its predecessor (European Commission for Human Rights), together with metadata identifying the title of the case (name of the party to the proceedings), date of issuance, ECHR body that issued the decision, subject area, result of the case (violation – non-violation), respective Convention rights, and adverse country. We scraped these data using the Excel macro VBA. This phase resulted in the creation of a list of ECtHR cases which could have potentially appeared as an object of a reference in domestic apex courts cases. We then proceeded with tracing these references on the macro-level.

5.2 MACRO-LEVEL ANALYSIS: GOTTA CATCH ‘EM ALL

As previously mentioned, the macro-level of our analysis aims to uncover very broad patterns in which domestic apex courts refer to ECtHR case law.

One of the crucial issues was therefore to establish the population for our research. As it is virtually impossible to identify the cases of non-application (i.e. cases where a domestic court should have or could have referred to the ECtHR case law\(^{33}\), but did not do so), we eliminated all cases where the courts did not refer to the ECtHR case law and focused only on cases with a reference to the ECtHR case law.

The core puzzle of the macro-level part of the research was how to tease out the information on the presence of a reference in the text

\(^{31}\) Here, we use the broad term decision, meaning both decisions and judgments.

\(^{32}\) www.hudoc.echr.coe.int.

\(^{33}\) We use the general term “ECtHR case law” even if sometimes the courts refer only to the ECtHR (or Strasbourg court) without any further specification. It is obvious that when referring, the case law of the ECtHR is being meant, not the ECtHR as such.
of the judgment automatically. Similarly to the ECtHR, the Czech apex courts also face a crisis in the overflowing number of petitions. Since 1993, the top three courts combined decided over 180 thousand cases. As it would be impossible to manually code such an amount of case law, we decided to use an automated text analysis consisting of a three-step process. First, we defined “a reference to ECtHR case law” and identified its constituents (see below). Second, we created an R algorithm able to recognise a reference according to these constituents. Third, we amended the algorithm in order to semi-automatically recognise which part of the domestic decision the reference was raised in. Each one of these steps has been manually validated.

Figure 2: Process of Automated Recognition of References

As illustrated in Figure 2, our first step was to identify constituents of a reference to ECtHR’s case law, i.e. indicators that a particular set of words represents a reference. For this purpose, we constructed an R algorithm based on several gazetteers.

Having data scraped the HUDOC database, we created a list of all ECtHR case file numbers and consequently searched for the presence of these file numbers in domestic decisions. If we found the respective file number, the algorithm extracted a small paragraph of text surrounding the reference, which allowed us to validate the results. This first stage left us with many false positives. Typically, shorter ECtHR case file numbers coincided with file numbers of decisions issued by domestic courts, particularly so in the case of the CC. Eventually, we eliminated all false positives with further amendments of the algorithm after several rounds of manual validation. During this first phase, we also validated eventual false negatives (sample of 200 cases), checking whether cases where the algorithm did not report any reference indeed did not encompass one.

Nevertheless, using an ECtHR file number is only one way in which domestic courts refer to its case law and, both from our first validation and

---

34 2002 for the Supreme Administrative Court.
35 Source: authors.
from practical experience, we knew that Czech apex courts are not particularly consistent in using the file numbers. We therefore amended the constituents of a reference to include either a file number or a name of the party to the proceedings (Table 2). A word of caution should be raised here, however. Some ECtHR judgments are better known by popular titles rather than by the names of the parties to the proceedings (e.g. Skoullos family, no. 55819/00) and some names are translated differently into Czech (e.g. Handölsdalen Sami village, no. 39013/04, in Czech referred to as “Sámská vesnice Handölsdalen”). We therefore created a new list of popular titles and Czech translations and added them to the respective titles in our algorithm. Yet, given the particularly high number of cases and the length of the texts, we had to proceed further, as adding the extensive list of names and popular titles into our algorithm for all domestic cases would lead to extremely lengthy and time-consuming processing. Therefore, we opted to start with a presumption that every time a domestic court refers to a particular ECtHR case, it also mentions the ECtHR itself. In other words, we first searched for all references to the ‘ECtHR’ and its Czech variations (Figure 3) and then searched for the case titles and names in the vicinity of 1000 words surrounding the general reference to ECtHR. Those results that did not match with any name of a party/title of a case were deleted as void general references.36

<table>
<thead>
<tr>
<th>General reference to ECtHR</th>
<th>ESLP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evrop* soud*</td>
</tr>
<tr>
<td>Reference to a case</td>
<td>File number</td>
</tr>
<tr>
<td></td>
<td>Name of the party to the proceedings</td>
</tr>
</tbody>
</table>

Table 2: Constituents of a Reference37

During the process of validation, we had to amend our algorithm several times in order to capture various acronyms, typos, or incorrect terms. The most tedious part of the analysis was a validation of whether the found reference was not in fact a reference to the CJEU (often simply called the European Court) or if the file number did not match with some other

36 It is quite common practice with Czech courts to refer to international bodies in very general terms (‘as follows also from the case law of the European Court…’). Nevertheless, we did not consider such broad references as adding to the concept of judicial compliance as they did not, in fact, refer to a particular decision of the ECtHR.

37 Source: authors.
domestic authority file number. Nevertheless, after the second phase, we managed to interactively assign broad references to the ECtHR to individual respective ECtHR file numbers.

During the process of validation, we had to amend our algorithm several times in order to capture various acronyms, typos, or incorrect terms. The most tedious part of the analysis was a validation of whether the found reference was not in fact a reference to the CJEU (often simply called the European Court) or if the file number did not match with some other domestic authority file number. Nevertheless, after the second phase, we managed to interactively assign broad references to the ECtHR to individual respective ECtHR file numbers.

The macro-analysis also allowed us to prepare samples for the meso-level of analysis, which focuses on a more nuanced analysis of how domestic courts use the references in their reasoning (Figure 3).

<table>
<thead>
<tr>
<th>POPULATION</th>
<th>SAC</th>
<th>SC</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases 1993–2015</td>
<td>39 477</td>
<td>84 374</td>
<td>60 403</td>
</tr>
<tr>
<td>MACRO-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cases</td>
<td>1 913</td>
<td>1 309</td>
<td>4 184</td>
</tr>
<tr>
<td>references</td>
<td>5 894</td>
<td>7 122</td>
<td>11 977</td>
</tr>
<tr>
<td>MESO-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cases</td>
<td>1 594</td>
<td>1 080</td>
<td>3 908</td>
</tr>
<tr>
<td>references</td>
<td>4 344</td>
<td>4 161</td>
<td>10 399</td>
</tr>
</tbody>
</table>

Figure 3: Populations for our Macro- and Meso-level Analyses

5.3 MESO-LEVEL: ONLY CATCH SOME, BUT CATCH THE ONES THAT COUNT
For the purposes of the meso-level analysis, we went one step further and examined the mode of application of the references. As previously mentioned, in the last step of the macro analysis we reduced the original population of references by eliminating references in the narrative parts and the separate opinions of domestic decisions. From the remaining cases, we selected samples of a few hundred judgments for each Czech apex court. We then, with the help of a team of coders, manually coded the form, the quality of the reference in these samples, and its impact on the dispute

38 2003 for the Supreme Administrative Court.
39 Source: authors.
at the heart of the case. This means that apart from looking at the formal characteristics of references (Does the reference contain a direct quote? Does the reference refer to a particular paragraph of the ECtHR’s case? Did the court use a full file number, the name of the party, etc.? Was the reference accompanied by a literature review?), we also reviewed how and why the apex court referred to ECtHR case law. We asked whether domestic courts used references only to support a reasoning based in domestic provisions, or to substantively change the outcome of the case; whether the apex courts’ use of the reference conforms to the ECtHR’s reasoning or, on the contrary, whether domestic courts refer to the ECtHR’s judgment only to refuse its application. Accordingly, the meso-level analysis proceeded based on a detailed and elaborated codebook (Figure 4).

<table>
<thead>
<tr>
<th>How the domestic court follows IHRB ruling</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Following</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distinguishing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusing</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Influence of the IHRB ruling on domestic decision</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supportive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substantive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How the IHRB ruling is used</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invalidation of a domestic legal norm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct application</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conforming interpretation of the domestic legal norm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filling the gap in domestic law</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirmation of domestic law</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level of the detail of the reference</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference to a specific IHRB ruling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference to a specific part of the IHRB ruling</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Meso-level Reference Coding – Clipping from the Codebook

If we were to return to the reduction-funnel (Figure 1), the aim of the meso-level analysis was to get to the narrowest part of the funnel and learn which references do indeed have a substantive impact on domestic cases. The clipping from the codebook (Figure 4) therefore captures the most important categories coded manually for every decision in the meso-level sample:

A) How is the ECtHR’s decision followed in a domestic decision: with the apex court either following (and confirming) the ECtHR’s finding, distinguishing that the case at hand and the legal question raised is different from the ECtHR’s case (therefore, the ECtHR judgment cannot be applied), or directly refusing to implement the ECtHR’s finding in the domestic case at hand.

40 Source: authors.
B) Influence of the ECtHR’s decision on a domestic decision: which can either play a merely supporting role (used as a mere ornament) or can substantively change the result of the case (i.e. the domestic court would decide differently should there be no ECtHR case law).

C) Technique of the use of the reference: here, only in case that the reference is of substantive relevance, we presume that apex courts can use it either to invalidate a domestic legal norm (using the argument of the strength and primacy of international obligations), assign the ECtHR judgment primacy without invalidating any legal norm, interpret the domestic legal provisions in conformity to the referred case, use the reference to the ECtHR decision to fill in the gaps in domestic legislation, or, to confirm the content of the domestic law.

We should, however, raise a few notions regarding the sampling method implemented in our meso-level of analysis. We already noted that in order to code the references manually, we used stratified samples capturing the various importance of decisions within apex courts’ case law. The consideration underlying this decision reflected the over-representation of procedural, unpublished decisions in the case law of apex courts that have a lower probability of citing a reference that substantively influences the result of the case. Each sample therefore captures a certain percentage of published and unpublished decisions and judgments, corresponding to the composition of these categories in the respective years (Figure 5).

![Figure 5: Clipping from the Stratification of Samples](source: authors)
5.4 MICRO-LEVEL ANALYSIS: GOTTA CATCH THE UNIQUE ONES

In the micro-level analysis, we use an in-depth qualitative analysis to address some of the issues uncovered in the previous part. We concentrate on compliance in a narrow sense, looking at the development in selected adverse cases held against Czechia, particularly those in which domestic courts’ interpretations contributed to a violation of the Convention rights. Close examination of individual cases should enhance our knowledge of the system, especially in the cases which are exceptional in some sense.

6. PILOT STUDY

In order to make the methodology of our project as comprehensible as possible, we tested it out on a pilot study of the SAC. In this pilot, we were particularly interested in whether the youngest of the Czech top judicial bodies (1) uses the references to ECtHR case law, (2) how often, (3) with what impact, and, (4) how its reputation as a young, liberal, active and pro-international body translates into the use of references.

We first analysed all existing cases issued by the SAC between 1 January 2003 and 31 December 2015. In accordance with theories on judicial dialogues and existing scholarly works on the SAC, we expected the SAC to be a particularly active actor. Figure 6 captures the development of the use of references to the ECtHR case law both by the count of references (orange line) and the count of the SAC’s decisions containing a reference (orange area).
Perhaps surprisingly, the lower part of Figure 6, does not clearly support this expectation, as, although the Court refers quite significantly, especially so in the last 2 or 3 years, the orange dashed line shows that it does so in fact only in a very limited percentage of all its case law (3% on average, a little over 5% at most).

There are, however, some minor difficulties which should be mentioned before diving into the next level of the funnel-filter. The SAC is in a peculiar position as, generally, the ECHR does not cover the whole area of administrative law. Therefore, although substantively permeating e.g. asylum protection, the right to freedom of assembly, regulation of political parties, or questions of a fair trial, there are certain areas of the SAC’s decision-making where the SAC cannot rely on the Convention or on ECtHR case law. This might suggest that further reduction of the population of cases substantially related to the SAC’s jurisdiction would be helpful, nevertheless, such a reduction is not feasible. When further analysing the metadata of the ECtHR’s decisions referred to most often by the SAC, we found that when it comes to alleged violations of Convention articles, the composition of these decisions is quite diverse, and not necessarily limited to the jurisdiction ratio materiae of the SAC.

We then proceeded with the next step of the macro-level analysis and filtered out those references which appeared outside the SAC’s own reasoning. In this respect, Figure 7 reports very interesting results. While at the very beginning of the SAC’s functioning, most of the references to ECtHR case law occurred in the parts summarizing previous proceedings.

42 Source: authors.
and arguments of the parties (i.e. outside the SAC’s own reasoning), this trend changed around 2008 with references in the SAC’s own reasoning becoming dominant. Several explanations come to mind here. Either the SAC significantly changed the language of its decisions and of the overall drafting process, or, in the first years of its existence, references to ECtHR decisions were raised significantly more often by parties to proceedings than by the SAC itself. It is true that our analysis confirmed the suggestions of previous scholarship about common flaws in the formal treatment of the ECtHR’s case law by domestic courts, especially the excessive anonymization, which not only complicated the recognition of references, but also made their use in reasoning weaker and less persuasive. Either way, the results clearly suggest that work with ECtHR case law has gradually gained prominence in the SAC’s own reasoning. This conclusion would be obscured if we relied only on the initial rough count of references.

![SAC Use of References in Different Parts of Decisions](image)

**Figure 7: Do the SAC’s References to ECtHR Case Law Appear in Recitals or in Reasoning?**

Yet, even the count of references cleaned of references outside of the SAC’s reasoning does not tell us much about how courts work with the references, why and on what occasions they use them. The meso-analysis of references allows us here to dig deeper and to zero in on those areas where the SAC’s case law is significantly influenced by the ECtHR’s

---


44 Source: authors.
decisions. We therefore coded cases where the SAC uses references of substantive influence, or, on the contrary, references merely supporting the reasoning and conclusion derived from the national law. Figure 8 (introducing the results of our meso-analysis) indeed shows that on most occasions, the SAC uses the references to the ECtHR’s decisions as supportive arguments in its reasoning. Moreover, we also found that the quality of the work with references changes in clear patterns: when the SAC invokes a reference in order to substantially influence its reasoning, the treatment of the ECtHR’s case law has a higher quality, offering more precise references and longer explanations.

A further step in the meso-analysis was then to concentrate on the most important patterns of reference techniques (Figure 9). While supportive references to ECtHR case law are undoubtedly also important for the ex-post control of domestic decision-making, the core interest of judicial compliance lies with cases where the apex court might push forward the compliance with the ECtHR case law despite its discord with the domestic legislation. Figure 9, in this respect, also shows that the SAC quite often uses references to ECtHR case law in order to decide on novel questions and problems unanswered by domestic legal norms (see “Fill in the gaps”), and, on the contrary, almost never – at least openly – refuses to implement the ECtHR’s findings.

---
45 Source: authors.
The abovementioned conclusions clearly illustrate the deficiencies of the automated text analysis used as a sole method for the research on judicial compliance. Although the automated text analysis is essential for the processing of huge datasets and for getting a rough overview, we need to know more about the content of judgments using references to the ECtHR case law and their context in order to form a proper understanding of judicial compliance. An analysis based purely on an automated text analysis does not provide us with the information of whether the reference on its own adds up to a court’s compliance with the ECtHR, or whether its use is purely incidental, or, even more importantly, if a domestic court rebels against the ECHR and does not follow its case law. A rough count of references cannot provide us with this deeper understanding of compliance, and therefore a combination of various methods, as envisaged by our three-level analysis, is vital.

7. CONCLUSION

Empirical research methods have seen significant progress in the last decades, especially due to the use of computers. Vast volumes of data can be analysed in a user-friendly way with software for both quantitative and qualitative methods. Data accessibility goes hand in hand with developments in analysis methods. Were the data not accessible, even the most sophisticated methods would be useless if they could not be applied to anything. Vice versa, research with highly accessible data would
be a horrendously time consuming exercise without the assistance of information technology.

Even law, as a rather conservative research field that does not usually stand at the forefront of scientific discoveries, has slowly attracted the increasing attention of researchers using the new tools of inquiry. Legal institutions, especially courts, produce huge quantities of text and machines can help in examining their outputs. While computers can prove extremely useful in acquiring, storing, processing, and analysing data, human supervision and creative input is needed throughout the whole process. Validation of the results returned from an automated text analysis is vital to the success of the whole research endeavour. Although reports on the successes of algorithms in predicting results of legal disputes sound marvellous and computers might even outperform legal experts,47 they are not free from criticism. They arguably gloss over the question of social and human meaning in legal systems, their use might reinforce existing biases and might exacerbate inequality in access to justice, and thus undermine the legitimacy of the legal system.48

The use of machines for text analysis is only slowly permeating our subfield – compliance and international human rights case law – and we perceive automated text analysis to be a helpful tool, but not a panacea. Previous works inspired us, especially Wind’s,49 but at the same time we sought to rectify some of its drawbacks. Particularly, one should not equate a reference to ECtHR case law in the narrative part of a domestic judicial decision (where courts only sum up the proceedings and submissions of the parties) with references to ECtHR case law in a reasoning that includes a court’s own judicial consideration of ECtHR case law and its impact on the case under consideration. The use of automated text analysis helped immensely for sorting out the references in the narrative and reasoning parts of the judgments, but at the same time required extensive


validation due to quite widespread incidence of false positives and false negatives.

The macro-level analysis depended on the computerized techniques, as going through the complete, massive body of domestic apex courts case law is an impossible task. The macro-level analysis can thus provide an overall picture of the main characteristics of the use of the E CtHR’s case law by Czech apex courts and it served as a building block for the selection of a sample of cases that was then used in the meso-level analysis. Hand-coding at the meso-level, which required good understanding of the reasoning, was able to capture more finely nuanced usage of E CtHR case law. The micro-level analysis demanded deep immersion into the subtleties of individual cases, which is still better done by human researchers than by machines. Our three-level analysis combining automated and traditional methods to collect and code data with a mix of quantitative and qualitative methods of analysis has the potential to contribute to a better understanding of the use of references to ECtHR case law and might shed more light on the concept of judicial compliance.

Our project on the judicial compliance of domestic apex courts cannot and does not aspire to cover all research questions of the field. However, it does contribute to a more nuanced and more systematic picture of judicial compliance, through discussion of the domestic judicial use of ECtHR case law. We move beyond standard compliance debates and analyse more broadly how ECtHR case law affects domestic jurisprudence and how it permeates the judicial reasoning of national courts. Such a thorough analysis gives us a more accurate picture of how domestic courts actually make use of ECtHR case law and how it affects their approach and their legal reasoning.

Finally, one of the main contributions of our three-level approach is its wide applicability, which stretches far beyond the ECtHR’s and the Czech apex courts’ case law. Not only it can be used on any given country and any given international human rights regime, but can also include any domestic or international law, case law or even literature (in short – any element used by anybody). The framework is especially useful when the number of cases referring to the element of interest is large, which would otherwise make an in-depth study of all cases in the population unfeasible.
LIST OF REFERENCES


