# Building a Legal Citation Network: The Influence of the Court of Cassation on the Lower Judiciary

Matthias Van Der Haegen*

## Introduction

Empirical research into judicial behaviour and the legal system is bountiful in the United States, but few and far between in Europe.[1] The reasons behind this knowledge gap are plentiful,[2] but slowly more and more research is successfully being undertaken on this side of the Atlantic.[3] The next frontier is legal big data analysis. In this paper, I describe the methodological framework of my PhD research, in which I empirically investigate the influence of the Belgian Court of Cassation on the lower judiciary. The normative significance of this research is considerable: the aim is to uncover the nature of precedent in a civil law jurisdiction, viewed from a hierarchical perspective, and thus, whether or not case law is considered a source of law in daily practice. As an early *caveat*, I should emphasize that this article does not contain any results. This research is very much still a work in progress, as the software to analyse the assembled data is still in an initial testing phase.

I contextualize firstly my research question by briefly touching upon the Court of Cassation and the reigning theories on the authority of its case law (Section 1). Secondly, I will discuss the operationalization of the research and the state of the art with regard to legal citation studies, highlighting the difficulties specific to Belgium that significantly impact my methodology (Section 2). After briefly touching upon the already assembled data (Section 3), I will discuss the research hypotheses in Section 4 and my research methodology in Section 5. Normative implications of my research are dealt with in Section 6, with a conclusion in Section 7.

1   Garoupa & Ulen argue that no legal (methodological) innovation has had a profound impact on European legal academia. N. Garoupa & Th. Ulen, 'The market for legal innovation: law and economics in Europe and the United States', (2008) 59 *Alabama Law Review*, no. 5, p. 1568.
2   Reasons might be, *inter alia*: the fact that the academic career path of European lawyers is limited to law faculties, whereas the inability to pursue a PhD in Law in US legal academia means that PhDs in the social sciences are a common sight, alongside the fact that legal education is a postgraduate education; the fact that companies like Westlaw and LexisNexis already have powerful research tools that invite academics to go further; and the widespread acceptance of legal realism and other theories of judicial behaviour etc. Garoupa & Ulen discuss at length several explanations as to why law and economics is less widespread in Europe than in the United States. Elements of their analysis apply to computational legal research as well, in my opinion. Garoupa & Ulen, ibid., pp. 1578-1619.
3   See for example with regard to the Belgian Constitutional Court: J. De Jaegere, 'Inclusiviteit als deliberatieve bouwsteen van legitimiteit: een empirische analyse van het Belgische Grondwettelijk Hof', (2015) 70 *Tijdschrift voor Bestuurswetenschappen en Publiekrecht*, no. 4-5, pp. 194-215; P. Popelier & J. De Jaegere, 'Wetenschappelijk bewijs voor wetgeving in de rechtspraak van het Grondwettelijk Hof: een empirische analyse en een zoektocht naar verklaringen', (2016) 20 *Tijdschrift voor Wetgeving*, no. 4, pp. 285-302. See also the references in Section 2.1, infra.

## 1. The authoritativeness of the Court of Cassation's case law

The Court of Cassation is Belgium's highest judicial authority for all civil and criminal law matters. Founded in 1831, the Court oversees the uniform application of the law and steers its development by either confirming or annulling the legal decisions brought before it.[4] Modelled after the French *Cour de cassation*, its Belgian counterpart has no means of controlling its docket, although the specialized Bar at the Court of Cassation acts as an implicit filter in all non-penal matters.[5] Nonetheless, the Court annually makes around 3,000 decisions. The Court of Cassation operates in Belgium's two main languages, Dutch and French. The three chambers of the Court are each split into two language sections. Appeals are decided by the corresponding language section of the competent chamber. Every published decision is translated by the Court itself into Dutch or French respectively.

Belonging to the French legal tradition, Belgium does not have a *de iure* precedent system, as opposed to the common law tradition where *stare decisis* is deeply ingrained in the judicial system and culture. As such, the Court of Cassation must rely on its moral authority to command the compliance of the lower judiciary with its decisions, as the sole way to fulfil its task of enforcing the uniform application and interpretation of the law throughout the country.[6] It is generally acknowledged that a *de facto* precedent system exists, in which the authoritative case law of the Court applies throughout the country[7] and establishes 'judicial peace'.[8] Whether daily practice resembles this theory, is unknown. Lower courts are said to normally fall in line behind decisions of the Court of Cassation and accept its doctrine, even though they previously held opposing opinions.[9] Examples to the contrary exist however.[10]

Settled legal dogmas lend themselves well to empirical investigation. Unearthing empirical evidence, whether it affirms or disavows long-held assumptions, is in itself already worth the while as it expands knowledge of the relevant domain. More importantly, empirical research also invites new substantive research questions, thus setting new frontiers for future research. Therefore, the assumption of the existence of a *de facto* precedent system is worth putting to the test. The authoritativeness which is attributed to the decisions of the Court of Cassation is merely assumed. Do courts actually rely on the case law of the Court when making decisions? My research measures the influence of the Court on the lower judiciary.

This research question also has substantive normative implications. The Court's *raison d'être* is to counter legal uncertainty and guarantee the equal application of the law. If empirical results were to suggest that large gaps exist in the acceptance of the Court's doctrine, the means by which a uniform application of the law is to be reached would have been shown to fail to deliver.[11] Such a result would necessitate profound

---

4    I. Verougstraete, 'Cassatie: vereenvoudiging, gezag en democratie', in C. Parmentier et al. (eds.), *Hof van Cassatie van België: 175ᵉ verjaardag* (2008), p. 205; C. Scheyven, *Traité pratique des pourvois en cassation, de l'organisation et des attributions de la Cour suprême* (1885), p. 99.

5    J. Laenens, 'De advocaten bij het Hof van Cassatie', (2003) 66 *Rechtskundig Weekblad*, no. 28, p. 1118.

6    In cases of annulment and remittal, the decision of the Court of Cassation is binding on the lower judge to whom the case is remitted (Art. 1110, fourth paragraph Belgian Code of Civil Procedure (*Gerechtelijk Wetboek*, *Ger.W.*)). This is a recent reform made under the Act of 6 July 2017, which introduced a system similar to the one in force in the Netherlands where the judge to whom the case is remitted after annulment, is bound by the decision of the *Hoge Raad* (Art. 424 Dutch Code of Civil Procedure (*Wetboek van Burgerlijke Rechtsvordering*)). See M. Van Der Haegen, 'Potpourri V maakt cassatierechtspraak bindend na verwijzing', (2017) 19 *De Juristenkrant*, no. 354, p. 1.

7    M. Van Hoecke & M. Elst, 'Basic features of the legal system', in H. Bocken & W. De Bondt (eds.), *Introduction to Belgian Law* (2001), p. 35; R. Kruithof, 'Naar een "gouvernement des juges" in het verbintenissenrecht?', in R. Kruithof & W. Van Gerven, *Hulde aan Prof. dr. R. Kruithof* (1992), p. 57.

8    B. Bouckaert, *Hoe gemotiveerd is cassatie? Pleidooi voor een waarachtig precedentenhof en een hernieuwde motiveringscultuur* (1997), p. 13; Kruithof, supra note 7, p. 40. The same can be said of France, see F. Lawson, *A Common Lawyer looks at the Civil Law* (1953), p. 84.

9    H. De Page, *Traité élémentaire de droit civil belge* (1962), Vol. I, p. 307. For France, see J. Bell et al., *Principles of French Law* (2008), p. 30.

10   See the history of the '*Loondoorbetalingsarresten*', in which the Court states the same legal rule over and over again, which indicates that the lower judiciary still has not fully accepted the Court's case law with regard to this issue. See, *inter alia*, Cass. 19 February 2001, C.99.0014.N; Cass. 19 February 2001, C.99.0183.N; Cass. 19 February 2001, C.99.0228.N; Cass. 19 February 2001, C.00.0242.N; Cass. 20 February 2001, P.98.1629.N; Cass, 13 June 2001, P.01.0431.F; Cass. 6 November 2001, P.00.0444.N; Cass. 30 January 2002, P.01.1393.F; Cass. 4 March 2002, C.01.0284.N; Cass. 2 October 2002, P.02.0643.F; Cass., 9 April 2003, P.03.0049.F; Cass. 10 April 2003, C.01.0329.F; Cass. 18 September 2007, P.07.0005.N; and as recent as Cass. 8 February 2016, C.15.0170.N. See also Kruithof, supra note 7, p. 40, footnote 56; Van Hoecke & Elst, supra note 7, p. 29.

11   A recent poll by the University of Leuven and publishing house Wolters Kluwer of 1,000 lay people and 1,041 legal professionals indicated that slightly over half of the respondents agree that like cases are treated alike by Belgian judges. Broken down to the respondents' respective backgrounds, the numbers are the following: lay people 40%, lawyers 59%, public prosecutors 88% and judges 95%. A. Keereman & R. Boone, 'Vertrouwen van burgers in justitie zakt weg', (2016) 18 *De Juristenkrant*, no. 333, p. 9.

reform of the legal system, in search for measures or procedures that would be more successful in fulfilling the Court's task.

## 2. Operationalization of the research

### 2.1 Citations as influence

The operationalization of the research question is a critical step in this empirical research project: how does one measure influence? For research into the interaction between courts, it makes sense to scrutinize citations in court decisions, as they are easily visible remnants of the intellectual interaction of one court with its own previous decisions or those stemming from a different court. Citations will therefore be used as a proxy for influence.[12] A mere citation does not necessarily mean that the holding of the cited case was followed, let alone that it was instrumental in the reaching of the verdict of the citing decision. But it does tell us something, namely, that for one reason or another, the citing judge used it to justify his or her decision.[13] In the context of a civil law country such as Belgium, this is all the more relevant as the citing judge is never under any obligation to follow the Court's case law.[14] When a judge does refer to a decision of the Court, it is therefore a conscious and deliberate decision.

### 2.2 State of the art

The second half of the 20th century witnessed the rise of citation studies in the United States. Only recently has this area of legal research been explored in Europe as well.[15] Technological advances have allowed researchers to study citations on a larger scale, unburdened by the inherent limits of manual labour, as a result of the dawn of big data analysis and its application in legal research. As a consequence, research on legal citations has been manifold in recent years, studying (cross-)references within judicial decisions,[16] legislation[17] and scholarly papers,[18] in order to unveil the network structure in the specific (inter)national legal domain. Natural language processing tools and algorithms to extract, normalize and identify legal

---

12  US literature suggests that citation of a court's decisions is an indicator of its authority. R. Mott, 'Judicial Influence', (1936) 30 *The American Political Science Review*, no. 2, pp. 295-315.

13  See M. Gelter & M. Siems, 'Networks, Dialogue, or One-Way Traffic? An empirical analysis of Cross-Citations between Ten of Europe's Highest Courts', (2012) 8 *Utrecht Law Review*, no. 2, <http://doi.org/10.18352/ulr.196>, p. 89. On the nature of citations, see J. Merryman, 'The Authority of Authority: what the California Supreme Court cited in 1950', (1954) 6 *Stanford Law Review*, pp. 614-619. On the reasons why judges cite, see ibid. pp. 621-627.

14  Except when a case is remitted to a judge by the Court after annulment of the previous decision. In this instance, however, he or she is only obliged to adhere to that specific decision of the Court, and not to its larger body of case law. See note 6, supra.

15  Merryman, one of the pioneers of citation studies in the US, has already hinted at the possible significance of similar studies in civil law countries, which he thought would reveal more similarity between the citation practices of common law and civil law countries than generally assumed. J. Merryman, 'Toward a Theory of Citations: an Empirical Study of the Citation Practice of the California Supreme Court in 1950, 1960, and 1970', (1977-1978) 50 *Southern California Law Review*, p. 428, footnote 40.

16  A. Geist, 'Using Citation Analysis Techniques for Computer-Assisted Legal Research in Continental Jurisdictions' (2009), available at SSRN: <https://ssrn.com/abstract=1397674>; M. Bommarito et al., 'Law as a Seamless Web? Comparison of Various Network Representations of the United States Supreme Court Corpus (1791-2005)', (2009) *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, available at SSRN: <https://ssrn.com/abstract=1419525>; J. Fowler et al., 'Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court', (2007) 15 *Political Analysis*, no. 3, pp. 324-346; M. Van Opijnen, *Op en in het Web: hoe de toegankelijkheid van rechterlijke uitspraken kan verbeterd worden* (2014); R. Winkels et al., 'Determining Authority of Dutch Case Law' (2011), available at SSRN: <https://ssrn.com/abstract=2013190>; Th. Smith, 'The Web of Law, San Diego Legal Studies Research Paper No. 06-11' (2005), available at SSRN: <https://ssrn.com/abstract=642863>; M. Gelter & M. Siems, 'Language, Legal Origins, and Culture before the Courts: Cross-Citations between Supreme Courts in Europe' (2013), available at SSRN: <https://ssrn.com/abstract=1719183>.

17  P. Mazzega et al., 'The Network of French Legal Codes' (2009), available at <dl.acm.org/citation.cfm?id=1568271>; J. Landthaler et al., 'Unveiling references in legal texts: implicit versus explicit network structures' (2016), available at <https://wwwmatthes.in.tum.de/pages/1l4761wikalkl/Unveiling-References-in-Legal-Texts-Implicit-versus-Explicit-Network-Structures>; M. Koniaris, et al., 'Network Analysis in the Legal Domain: a complex model for European Union legal sources' (2015), available at <https://arxiv.org/abs/1501.05237>; Van Opijnen, supra note 16; B. Karstens et al., 'Reference Structures in National Constitutions', in R. Winkels & N. Lettieri (eds.), *NAiL 2014: 2nd international workshop on 'Network Analysis in Law'* (2014); N. Lettieri et al., 'Nets of Legal Information: Connecting and Displaying Heterogeneous Legal Sources', in R. Winkels & N. Lettieri (eds.), *NAiL 2014: 2nd international workshop on 'Network Analysis in Law'* (2014); M. Adedjouma et al., 'Automated Detection and Resolution of Legal Cross References: Approach and a Study of Luxembourg's Legislation' (2014), available at <http://ieeexplore.ieee.org/document/6912248/>.

18  Van Opijnen, supra note 16; T. Agnoloni & U. Pagallo, 'The case law of the Italian constitutional court, its power laws, and the web of scholarly opinions', in K. Atkinson & T. Sichelman (eds.), *Proceedings of the 15th International Conference on Artificial Intelligence and Law* (2014), pp. 151-155.

references have been successfully developed, handily dealing with difficulties such as non-uniform citation standards.[19] Focus has been firmly placed on the development of measures of legal relevance or information retrieval[20] or the discovery of a network structure in itself.[21] Normative implications of the existence of such network structures in the legal domain have received less attention. The building or unveiling of citation networks, in order to model and quantify legal knowledge, has been the end goal for most scholars.

### 2.3 Belgian citation practice

Research similar to mine has also relied on citations.[22] For reasons inherent to the Belgian judicial system, these can only be used as an example to a certain extent. The referencing culture of Belgian judges differs quite substantially from that of many of their foreign counterparts. Explicit references to case law in judgments are few and far between.[23] As Belgium does not have a precedent-based legal system, case law is not a binding source of law there despite its attributed authoritative value.[24] As a multitude of references to case law would subvert this theory, judges tend to uphold a legalist facade by not necessarily citing the case law on which they rely. The interpretation commonly, albeit wrongly, given to Article 6 Ger.W., namely that it would prohibit the explicit reference to other case law in the grounds of a judgment, might also explain the apparent lack of references.[25] Moreover, judges must tread very carefully when they do refer to other case law in their judgments. At the slightest impression that the judge considers himself or herself bound in any way by previous case law, the Court of Cassation will annul the judgment for violation of Article 6 Ger.W.[26] It is therefore often simply easier for judges not to refer to case law explicitly while at the same time relying on it.

Whatever the reason behind this peculiar behaviour, it remains a fact that explicit citations are few.[27] This does not necessarily mean that judges are not influenced by the Court of Cassation, or decisions of other judges in general, when these are relevant to the case at hand. Most likely, relevant and authoritative case law will either be cited in the lawyers' briefs, or the judge will come across it whilst undertaking legal research before reaching a decision. Either directly or indirectly, judges will be influenced by what their colleagues have written in the past.

### 2.4 Implicit citations

Given the scarcity of explicit references, basing my research on them would give a myopic view of the actual role the Court of Cassation plays in the decision-making process of judges. My methodology needs

---

19  Bommarito et al., supra note 16; Van Opijnen, supra note 16.

20  Agnoloni & Pagallo, supra note 18; Geist, supra note 16; Van Opijnen, supra note 16; Winkels et al., supra note 16; P. Zhang & L. Koppaka, 'Semantics-Based Legal Citation Network' (2007), available at <http://dl.acm.org/citation.cfm?id=1276342>; T. Schwartz et al., 'A Legal Citation Recommendation Engine Using Topic Modeling and Semantic Similarity' (2015), available at <https://www.semanticscholar.org/paper/A-Legal-Citation-Recommendation-Engine-Using-Topic-Schwartz-Berger/c1279eab77ff63fd6b62da8d85033e8c0a295e45>.

21  Mazzega et al., supra note 17; Landthaler et al., supra note 17; Smith, supra note 16; Karstens et al., supra note 17; A. Lyte et al., 'Network Metrics of the U.S. Code' (2015), available at <https://www.mitre.org/publications/technical-papers/network-measures-of-the-united-states-code>.

22  Agnoloni & Pagallo, supra note 18; Fowler et al, supra note 16; Van Opijnen, supra note 16; Winkels et al, supra note 16; Y. Panagis & U. Sadl, 'Making EU (case)law: Evidence from a paragraph to paragraph network on the cases concerning the citizenship of the European Union' (2015), available at <http://www.karlbranting.net/law-and-big-data-workshop/LawBD-2015_submission_3.pdf>.

23  The same is true for France, see M. Lasser, *Judicial deliberations: a comparative analysis of judicial transparency and legitimacy* (2004), p. 36.

24  It is argued that the higher a court is positioned in the judicial hierarchy, the more authoritative its decisions are: Van Hoecke & Elst, supra note 7, p. 29.

25  See H.U. Jessurun d'Oliveira, *De Meerwaarde van Rechterlijke Uitspraken: arrêts de règlement en precedenten* (1973), pp. 13-14. Art. 6 Ger.W., however, only means that a simple reference to or citation of another judgment is not a sufficient rationale for a decision, as this would consider the referred to decision to be a binding source of law, which it is not.

26  There is no violation of Art. 6 if the judge sets out the reasons why he or she follows the earlier case law, or indicates that he or she follows it not because he or she considers to be bound by it, but because he or she agrees with its content. See S. Van Overbeke, 'Verwijzing naar rechtspraak in de rechterlijke beslissing', (1994-1995) 58 *Rechtskundig Weekblad*, pp. 500-502; M. Adams, 'Een rechtstheoretische glosse bij artikel 6 van het Gerechtelijk Wetboek', (1994-1995) 58 *Rechtskundig Weekblad*, pp. 1087-1089, and their references. The same is true in France. See S. Bragdø-Ellenes, 'An introduction to French legal culture', in S. Koch et al. (eds.), *Comparing legal cultures* (2017), pp. 147-148; E. Steiner, *French Legal Method* (2002), pp. 80-81.

27  Especially in contrast to e.g. common law judgments. Common law lawyers reading Belgian judgments would most likely have the impression that Belgian judges decide completely independently of each other, without any concern for the structure, coherence, or stability of the broader legal system, basing their judgment solely on their personal interpretation of the applicable law. This impression is to a large extent mistaken. See M. Shapiro, *Courts: a comparative and political analysis* (1981), pp. 135-136; Lasser, supra note 23, pp. 38-39.

to be adapted to take into account the intricacies of Belgian citation practice. As such, implicit citations come into play. An implicit citation for the purpose of my research is a phrase the source of which is not explicitly mentioned. Any sentence might or might not be an implicit citation. Uncertainty with regard to the correctness of the attribution to that source is inherent. It is moreover impossible to determine whether the author knowingly referred to that source implicitly, or whether the use of the same or similar expression is due to an unconscious influence or pure chance.[28] Given the lack of explicit citations, my research must be based on implicit citations. The few explicit citations should be identified as implicit citations as well, and as such offer an opportunity to evaluate the performance of the method used to identify the implicit citations, as the explicit citation offers original testimony of the link between the citing and cited decision.

Implicit citations are found through measuring semantic or syntactic similarity, which expresses the similarity between two phrases as a value between zero and one. The higher that number, the more plausible the existence of an implicit reference. One method that has received traction within the field of natural language processing, is based on the co-occurrence of (lemmatized) nouns in sentences:[29] when between two phrases a predefined threshold of similarity is met, that similarity is treated as an implicit reference. More advanced methods, based on vectorization, have been used as well.[30]

## 3. Data

As with any empirical research, data either make or break the project. My research is not an exception, and thus requires data of a sufficient magnitude and quality in order to draw inferences. For this, I initially contacted the major legal publishing companies active in Belgium. Of these, unfortunately, only Wolters Kluwer Belgium was willing to cooperate by opening up its entire dataset of judicial decisions for my research. For the Court of Cassation, the decisions I obtained from Kluwer number 17,462, spanning the period between 2000 and 2015. On average, the Court publishes yearly around one-third of its decisions.[31] The assembled dataset contains the large majority of the published decisions of the last fifteen years. 9,352 of these decisions are unique, which highlights the fact that, for the majority of cases, I have both the original decision and its translation into either Dutch or French. Given the fact that I rely on implicit citations, this is a considerable advantage: implicit citations in judgments of Dutch-speaking courts to French-speaking Court of Cassation decisions, for example, do not require any more effort to identify than the identification of citations within the same language group, as I can rely on the translations.[32]

With regard to judgments of the lower judiciary, I assembled data from various sources. Kluwer provided 17,027 decisions. Almost half of those stem from the (labour) courts of appeal. Data from Juridat, the external database of the judiciary, is included in my research as well. This dataset currently contains over 200,000 decisions, although the exact numbers for distinct courts are not clear at this point. At the end of 2013, Juridat contained slightly fewer than 150,000 decisions, with higher ranking courts dominantly present.[33] Decisions from the labour courts and labour courts of appeal figure strongly in this dataset.

---

28  A court may decide a case in line with a previous Court of Cassation decision, but this does not necessarily mean that this court did so because of that previous decision. Holding otherwise would amount to a *post hoc ergo propter hoc* fallacy. See S. Wasby, 'The Supreme Court's Impact: some problems of conceptualization and measurement', (1970) 5 *Law & Society Review*, p. 46.

29  Landthaler et al., supra note 17; Bommarito et al., supra note 16. Other techniques have also been used. See for example the use of Jaccard Distance to measure similarity in: W. Alschner & D. Skougarevskiy, 'Mapping the Universe of International Investment Agreements', (2016) 19 *Journal of International Economic Law*, pp. 561-588. A lemma is the base form of a word.

30  Zhang & Koppaka, supra note 20; Schwartz et al., supra note 20.

31  P. Taelman, 'Potpourri en de burgerlijke rechtspleging', in W. Vanbiervliet (ed.), *De Potpourri-wetten: een eerste evaluatie van de waaiers aan quick wins, reparaties en grondslagen voor organisatorische vernieuwingen* (2017), p. 4. A request to obtain the non-published decisions of the Court of Cassation was, however, denied. This is not an impediment to my research, as non-published decisions can have no influence on the judiciary, with the exception of the court to which the case is referred after annulment.

32  The translation of decisions takes place under the auspices of the Court itself.

33  See Commissie voor de Modernisering van de Rechterlijke Orde, 'Verslag gewijd aan de bekendmaking van rechterlijke beslissingen: de veer, de Pelikan en de cloud' (2014), available at <https://justitie.belgium.be/nl/rechterlijke_orde/toezicht_en_advies/commissie_modernisering_rechterlijke_orde/informatie/publicaties>, pp. 11-13.

All decisions currently included in my dataset are published decisions. The total number of judicial decisions annually rendered by the Belgian judiciary stands at about one million.[34] Only a very small proportion of decisions are published, and, moreover, publication practice in Belgium does not correspond to any scientific criterion.[35] For this reason, I am in the process of obtaining unpublished decisions in bulk. Given the sorry state of the Belgian judicial data management,[36] this is not an easy task. In recent years efforts have been made and improvement has been effected, but the storage of data is done with merely judicial and administrative needs in mind, and is not adapted to its possible use by scholars or the public at large. If this attempt proves successful, however, I should obtain well over 100,000 unpublished recent decisions of the lower judiciary.

## 4. Research hypotheses

The influence of the Court of Cassation on the lower judiciary is a vague concept and needs clarification. More precisely, I intend to investigate whether, and the extent to which, judges refer to the case law of the Court of Cassation. The hypothesis is that they do, which is corroborated by literature.[37] Numerous factors impact the extent to which judges refer to the Court's decisions. Possible differences that might exist between judges in their citation behaviour are of particular interest. The existence of such differences will be researched with regard to hierarchy, geography and the influence of the local appellate court as a combination of hierarchy and geography.[38] These three areas of possible differences form distinct subhypotheses. They are mainly based on the rich US literature on impact.[39]

### 4.1 Hierarchy

The first subhypothesis centres on the position of a given court in the judicial hierarchy: the higher a court is positioned in the hierarchy, the more likely it is to cite Court of Cassation case law in its decisions. This could be explained by the fact that, compared with decisions of the courts of first instance, decisions of an appellate court are more likely to be brought before the Court of Cassation.[40] As such, judges might more readily anticipate such an appeal and go to greater lengths to make their judgment 'cassation-proof' by relying and drawing heavily on the Court's established doctrines. Differences might exist within the same hierarchical levels as well.[41] American literature suggests that specialized courts cite the US Supreme Court less frequently than courts with a more general competence.[42]

---

34  H. Van Bossuyt, 'Toegang tot de rechtspraak en toegankelijke rechtspraak', in A. Bossuyt et al. (eds.), *Liber spei et amicitiae: liber amicorum Ivan Verougstraete* (2011), p. 109.

35  See the report of the Commission for the Modernization of the Judiciary: Commissie voor de Modernisering van de Rechterlijke Orde, supra note 33, p. 4.

36  See for an introduction Taelman, supra note 31, pp. 2-7.

37  See Section 1, supra.

38  Differences will more likely than not also exist with regard to subject matter: in some areas of law, the Court's case law is extremely important due to a lack of, outdated or succinct legislation. As it is difficult to distinguish automatically between cases based on subject matter, and many cases revolve around several areas of law, this difference does not feature in my research as it currently stands. The docket number of each Cassation decision indicates to which (broad) area of the law the case belongs. This can serve as an initial indication, but little more than that.

39  *Inter alia* M. Petrick, 'The Supreme Court and Authority Acceptance', (1968) 21 *Western Political Quarterly*, no. 1, pp. 5-19; Th. Barth, 'Perception and Acceptance of Supreme Court Decisions at the State and Local Level', (1968) 17 *Journal of Public Law*, no. 2, pp. 308-350; Wasby, supra note 28, pp. 41-60; L. Baum, 'Lower-Court response to Supreme Court Decisions: reconsidering a negative picture', (1978) 3 *The Justice System Journal*, pp. 208-219; G. Gruhl, 'The Supreme Court's Impact on the Law of Libel: Compliance by Lower Federal Courts', (1980) 33 *Western Political Quarterly*, no. 4, pp. 502-519; D.R. Songer et al., 'The Hierarchy of Justice: Testing a Principal-Agent Model of Supreme Court-Circuit Court interactions', (1994) 38 *American Journal of Political Science*, no. 3, pp. 673-696; S. Benesh & M. Reddick, 'Overruled: an Event History Analysis of Lower Court Reaction to Supreme Court Alteration of Precedent', (2002) 64 *The Journal of Politics*, no. 2, pp. 534-550.

40  This is of course logical. Judgments of the courts of first instance cannot be brought before the Court of Cassation, unless they are rendered in an appeal procedure against decisions of the justice of the peace or police tribunal, who in general deal with low value civil cases and petty crimes, respectively.

41  E.g. the specialized labour court as opposed to the general court of first instance.

42  L. Baum, 'Specialization and Authority Acceptance: the Supreme Court and Lower Federal Courts', (1994) 47 *Political Research Quarterly*, no. 3, pp. 693-703.

**4.2 Geography**

The second subhypothesis is that geographical differences exist with regard to the influence of the Court of Cassation. This would be because of relevant traits shared only within a distinct geographical area. This subhypothesis can be further divided into two subdivisons: a centre-periphery difference and the existence of jurisprudential pluralism within the country.

### 4.2.1 Centre-periphery

The first subdivision concerns a centre-periphery effect, which generally speaking would entail that the farther a court is geographically located away from Brussels, the seat of the Court of Cassation, the lower the likelihood is that it refers to the Court's case law. The country's legal elite is centred in and around Brussels, which means that judges and lawyers located in the capital are more likely to frequent the same social circles and networks, both professional and leisure, as justices of Court of Cassation and lawyers of the associated Bar.[43] Being regularly confronted with the Court's main actors might make the lower judiciary more susceptible to the Court's case law. Distance from the Belgian capital itself would be too rough a method for evaluating the centre-periphery effect. The periphery itself is not monolithic but rather ranges from metropolitan to more rural areas, which must be taken into account as well. The centre-periphery effect will thus be measured through a combination of distance and the level of urban development in a given area, as expressed by the number of inhabitants within a given court's jurisdiction.[44]

### 4.2.2 Jurisprudential pluralism

The second subdivision relates to jurisprudential pluralism[45] within Belgium. The hypothesis is that significant discrepancies exist between distinct geographical regions of Belgium with regard to the influence of the Court. Unifying case law and guiding legal development is the Court's express mission. This necessarily presupposes that case law and legal development is, or at least can be, divergent within the country. The existence of such a divergence will be researched first of all in relation to the different linguistic areas, namely Dutch-speaking and French-speaking Belgium. Flanders and Wallonia are said to have different legal cultures.[46] This difference in legal culture[47] might be reflected in the way judges from both regions treat the Court's case law. Related to this is the question whether judges are more inclined to cite those decisions from the Court that stem originally from the section of their native language.[48]

Secondly, regional differences might exist as well on a smaller scale within each linguistic area. Circumstantial evidence for the existence of local differences in case law can be found in the creation of legal journals for specific geographical areas.[49] The reasons behind their founding might be manifold, but the idea that local case law is more relevant to local practitioners than 'non-local' jurisprudence, is central.[50]

---

43  See for the existence of a Washington-centred legal elite with regard to the US Supreme Court: K.T. McGuire, 'Lawyers and the U.S. Supreme Court: the Washington Community and Legal Elites', (1993) 37 *American Journal of Political Science*, pp. 365-390. In Norway, a similar centre-periphery effect is said to exist with regard to the legal elite. See G. Grendstad et al., *Policy Making in an Independent Judiciary: the Norwegian Supreme Court* (2015), pp. 86-87 and 149.

44  Based on the statistics published by the federal government, see <http://statbel.fgov.be> (last visited 27 September 2017).

45  I opted to use the term jurisprudential pluralism as opposed to judicial pluralism, since the differences pertain solely to the substance of the case law and not to the organization or structure of the different courts.

46  See D. Heirbaut, 'The Belgian legal tradition: does it exist?', in M. Kruithof & W. De Bondt (eds.), *Introduction to Belgian Law* (2017), pp. 18-19. See also P. Durand, *Le droit des obligations dans les jurisprudences française et belge* (1933), pp. 239-240; E. Van Dievoet, *Het Burgerlijk Recht in België en Nederland van 1800 tot 1940: de rechtsbronnen* (1943), pp. 350-351, where differences in interpretation between courts in Flanders and Wallonia are discussed. Durand especially attributes these to the different psychological nature and temperament of the Flemish and Walloons.

47  For a discussion of the notion of legal culture, see: J. Øyrehagen Sunde, 'Managing the Unmanageable: an essay concerning legal culture as an analytical tool', in S. Koch et al. (eds.), *Comparing legal cultures* (2017), p. 14.

48  Every chamber of the Court is divided into a Dutch- and French-speaking section. As indicated earlier, all published decisions are translated by the Court into its other working language.

49  For example: *Tijdschrift voor Gentse Rechtspraak, Tijdschrift voor West-Vlaamse Rechtspraak, Rechtspraak van de Haven van Antwerpen, Limburgs Rechtsleven, Jurisprudence de Liège, Le pli juridique: bulletin de la Commission scientifique commune aux barreaux de Charleroi, Mons et Tournai, Revue de jurisprudence du Hainaut, Recueil de la jurisprudence du Tribunal de Nivelles, Rechtspraak van sociaal recht: Brussel – Leuven – Nijvel, Jurisprudence commerciale des Flandres, Revue de jurisprudence de l'arrondissement de Charleroi,* and *Turnhouts Rechtsleven.* Not all of the mentioned journals are currently still in existence.

50  See S. Vandenbogaerde, *Vectoren van het recht: geschiedenis van de Belgische juridische tijdschriften* (forthcoming 2017), pp. 324-325.

This highlights the fact that jurisprudential differences can exist between distinct regions: if the law was applied uniformly across all jurisdictions, differentiating between courts on this basis of geography would be irrelevant for legal practice. This jurisprudential pluralism might be reflected in the extent to which courts rely on the case law of the Court of Cassation and the different nature of the cited decisions.

### 4.3 Role of the appellate courts

The third subhypothesis combines both previous subhypotheses, and states that the respective courts of appeal play a pivotal role in the acceptance of Court of Cassation case law, by bridging the gap between the Court and the lower courts located within their jurisdiction. In this hypothesis, Court of Cassation case law would only find firm ground after it has been relied upon by the appellate courts as vital intermediary actors.[51] Two reasons might explain the central role of the courts of appeal: firstly, given the vast number of decisions the Court makes each year, the courts of appeal reduce its case law to more manageable proportions by drawing the lower courts' attention to fewer Cassation decisions. Secondly, by applying Cassation doctrine in a factual context, the courts of appeal make it more digestible for lower courts.[52]

The regional jurisprudential differences would possibly follow the boundaries of the jurisdictions of the distinct courts of appeal. Courts of appeal set out a judicial policy within their jurisdiction, especially in areas within the factual realm where control by the Court of Cassation is impossible.[53] Lower courts are said to attach greater weight to the case law of the court of appeal to whose jurisdiction they belong,[54] and the courts of appeal are in turn influenced by the case law they review, namely that of the courts within their respective jurisdictions. Different courts of appeal might thus have diverging opinions on like issues, and these differences would in turn be reflected in the case law of the lower courts within their jurisdictions. If the differences between the courts of appeal in their treatment of Court of Cassation case law are also present within the courts of their respective territorial jurisdictions, additional evidence is found for the hypothesis that the courts of appeal are the brokers of Court of Cassation case law.

### 5. Methodology

The methodology to conduct this research can be split into two distinct parts: pre-network and post-network methodology.

### 5.1 Pre-network methodology

The pre-network methodology consists of those methods needed to move from simple text documents to a citation network between cases. This process is divided into three steps.

#### 5.1.1 Text parser

The judgments, both from the Court of Cassation as well as those stemming from other courts, are made public in either portable document format (PDF) or Word format. These are transformed to ordinary text files. Relevant metadata, such as date of the decision, nature and location of the court, docket number, etc., are then extracted from the plain text files in which they are embedded by a text parser. Considerable difficulties, such as non-uniform docket number and date formats, have been surmounted in a time-consuming trial and error process. Decisions of the Court of Cassation adhere to a uniform structure, which helps in the correct identification and extraction of the metadata and relevant parts of the decisions. The

---

51  See L. Baum, 'Responses of Federal District Judges to Court of Appeals Policies: an Exploration', (1980) 33 *Western Political Quarterly*, no. 2, p. 219.
52  Lower courts would be aware of the position of their court of appeal, because of the publication of those decisions or because their judgments have been revised on appeal. It is likely that some time lapses before the decisions of the lower courts and that of their court of appeal reach a state of convergence, however stable or fragile such a state may be.
53  For France, see P. Théry, 'La "jurisprudence" des cours d'appel et l'élaboration de la norme', in N. Molfessis (ed.), *La Cour de cassation et l'élaboration du droit* (2004), p. 129. See also Steiner, supra note 26, p. 86.
54  Van Hoecke & Elst, supra note 7, p. 29. The same can be said for France: Steiner, supra note 26, p. 86.

same cannot be said unfortunately for decisions of the lower courts, which do not adhere to any specific standard. All metadata necessary to test for the hypotheses are collected in this manner. Finally, the file is converted to an XML file (Extensible Markup Language), with appropriate tags for all metadata.

### 5.1.2 Caché database management system

I collaborate with Intersystems, a US-based software developer and MIT spin-off specialized in database management systems. Caché is their object-oriented database management system. Each decision is stored within this database in XML-format, together with its metadata. Data can be accessed as both object-based, and through SQL (Structured Query Language) statements and queries. The reason why Caché was chosen over other database management systems, next to the purely pragmatic reason that Intersystems was willing to provide licences and technical assistance, is the fact that Caché is supplemented by many embedded technologies, the most relevant of which is iKnow.

### 5.1.3 iKnow

iKnow is a tool for text analysis, using a bottom-up approach to explore texts. iKnow structures text into 'concepts' and 'relations', making use of pre-defined libraries.[55] A 'concept' is a word group (e.g. a noun with its adjective), whereas a 'relation' is those words, usually verbs but also adverbs, that link concepts together to form a coherent sentence. Having structured the text in this manner, documents can be analysed and compared based on the concepts and relations they contain. Using iKnow, every decision of a lower court will be analysed, and compared with all decisions of the Court of Cassation, on a sentence-by-sentence basis. When similarity meets a certain pre-defined threshold, this similarity is treated as an implicit citation. The results of this analysis are then also saved within the database. In many respects, this method is analogous to plagiarism detection software.

This is a rigid approach, but rigidness has two advantages for my research: firstly, legal language is relatively invariable, as the use of synonyms or the inclusion of other words might convey not just a mere difference in nuance but rather a change in meaning. Secondly, as the citations that I try to identify are implicit, a higher degree of rigidness in the identification process increases the chances of correctly labelling implicit references as such.[56] The main drawback of this approach is that only (quasi) copy-paste implicit citations can be found, ignoring those citations that are not an exact match. For that reason, I will complement this method by another, less rigid approach drawn from computational linguistics to identify implicit citations. This method is likely to be based on the co-occurrence of lemmatized nouns, and thus comparable to the method used by Landthaler et al.[57] and Bommarito et al.[58] The danger of this method is of course that too wide a net might be cast, falsely identifying as implicit citations phrases where the author did not intend to cite Court of Cassation case law or was not unconsciously influenced by it. A combination of both approaches would thus produce the most reliable results.

It is likely that many of the identified implicit citations will in fact not constitute citations from a lower court to the Court of Cassation, but rather be the result of both the Court of Cassation and the lower court citing the same source, for example legislation. To counter this non-negligible possibility, all found implicit citations that are the expression of a legislative rule will be excluded from the results.[59] A similar risk exists with legal scholarship. Avoiding this danger is much more difficult, if not impossible, as no single repository of Belgian (or international) legal scholarship exists that contains all relevant doctrine, which could be used in the same way as with legislation to eliminate those implicit citations whose common origin lies within legal scholarship.

---

55 These libraries are built by linguists and contain virtually the complete vocabulary and grammar of a language. Libraries are available for both Dutch and French.
56 And thus reduces the chance of false positives (type I errors).
57 Landthaler et al., supra note 17.
58 Bommarito et al., supra note 16.
59 Whether this will be done only for those expressions that are an exact copy of the legislative rule, or also for those that are similar, still needs to be decided. What appears to be a minimal change to the wording of the rule, could very well entail a significant change in meaning that follows from the case law of the Court.

## 5.2 Post-network methodology

The post-network methodology analyses the established citation network and tests the hypotheses. Two distinct quantitative methods are applied, frequently used within social sciences but relatively new to (European) legal academia: multivariate regression analysis and social network analysis.

### 5.2.1 Multivariate regression analysis

Multivariate regression analysis will be used to test for the effect of different variables on the extent to which a court implicitly cites Court of Cassation case law (dependent variable), whilst controlling for other variables. The first subhypothesis will be tested by using both the citing court's nature[60] and the distinct hierarchy levels[61] as independent variables. The second subhypothesis requires a less straightforward approach: operationalization will be done using a nominal variable (district of the citing court), an ordinal variable (number of inhabitants within the jurisdiction of the citing court, categorized) and a numerical variable (distance from Brussels in kilometres) as independent variables. Belgium has twelve judicial districts, with most of them spanning the territory of a province. Each district has between two and four seats of courts, which used to form independent districts of their own until the reform of the judicial structure in 2014, which brought the number of districts down from 27 to 12. Most of the data predates this reform. Using the post-2014 districts as a variable, allows testing for broader regional differences in the acceptance of the Court's case law, for example between different provinces.[62] Using also, however, the pre-2014 structure with 27 different districts as a variable, allows for more nuance as, within each district, areas with a distinct character exist.[63]

### 5.2.2 Social network analysis

Social network analysis is a method used to analyse relationships (ties) between research subjects (nodes), and patterns within these relationship networks.[64] For my research, courts would be the nodes with citations of the same Court of Cassation phrase[65] as ties. The second and main hypotheses will be verified using cluster analysis methods. The existence of clusters would indicate that different subsets of courts tend to rely on different case law of the Court of Cassation. This might *inter alia* be because they belong to the same geographical area (propinquity effect) or that these courts share similar cultural and regional traits (assortativity). I will use centrality measures to verify the main hypothesis. If the courts of appeal have the greatest betweenness centrality, this indicates that they are vital intermediary players in the diffusion process of Court of Cassation case law, which would provide additional evidence for the first hypothesis. With regard to software, I will rely on UCINET for the network analysis and Gephi for visualization purposes.

## 6. Normative implications

The insights gathered from my research shed a new light on the intricacies of precedent and the behaviour of judges in a judicial hierarchy and European continental context. In current doctrine, the role of the courts of appeal in the law-making process of the Court of Cassation is limited to the *ex ante* conveyance of solutions to legal problems, after which the Court hands down a judgment settling the issue, at least for the time being.[66] If supporting evidence is found for the hypothesis that courts of appeal are the brokers of

---

60   E.g. court of first instance, commercial court, labour court.
61   E.g. court of first instance as opposed to courts of appeal.
62   E.g. the Antwerp district, which can be said to be quite urbanized and densely populated, as opposed to the district of Luxembourg, which is predominantly rural and sparsely populated.
63   For example, even though both the court of first instance in Bruges and the court of first instance in Veurne have, since 2014, belonged to the same district of West-Flanders, the latter city is situated in a more rural and sparsely populated area as opposed to the more urbanized and populated area around Bruges.
64   S. Wasserman & K. Faust, *Social network analysis: methods and applications* (2006), p. 3.
65   Citations will be of phrases that form part of the Cassation lexicon, and not so much of singular decisions, as the Court frequently repeats its previously stated doctrine, which would make it difficult to identify the exact decision in which the phrase originated.
66   See for a discussion of the role of the lower judiciary with regard to the law-making of the (French) Court of Cassation: F. Zenati, 'La saisine pour avis de la Cour de cassation', (1992) *Recueil Dalloz*, no. 32, p. 247.

the Court of Cassation case law, the *ex post* role of these courts would come to light, indicating that these courts play a much more important role in shaping the law than previously anticipated.[67] This would prompt a discussion on the authority of supreme court decisions and the nature of precedent in the European continental legal system.

My research also challenges the functioning of the Court of Cassation by critically evaluating whether the long-held theory on the authoritativeness of the Court's case law withstands empirical testing. If the Court has only little influence on the majority of the courts, questions should be raised about the extent to which it is successful in fulfilling its *raison d'être*. A reform of the nature of the Court of Cassation would be unavoidable, possibly steering away from the cassation mechanism towards the direction of more precedent-based supreme courts, such as exist in the Scandinavian countries[68] and the United Kingdom.[69] The question of compliance is also intrinsically linked with legitimacy.[70] If compliance is limited, this could be attributed *inter alia* to a lack of legitimacy. Measures to heighten accountability, such as a radical change in reasoning style, would then need to be put in place. Moreover, the results of this research will also tackle the issue of the position of case law in the hierarchy of legal sources: if the decisions of the Court of Cassation are amply cited, without much variance in wording, there would be little practical difference between those non-binding decisions, and legislation.[71] As such, a higher degree of transparency in the decision-making process of the Court would be required, and the issue of the democratic legitimacy of such 'legislation from the bench'[72] would be brought to the fore.

## 7. Conclusion

Computational empirical research within legal academia is scarce in Europe, where traditional research methodology still reigns supreme.[73] The Empirical Legal Studies Movement, originating in the United States,[74] is however bound to take over Europe as well. The use of big-data analytical tools in both legal practice and research is an inevitable evolution. Both mainstream legal academia and the judiciary itself should therefore make preparations, if they do not want to miss the boat and be confronted with an unsurmountable backlog.

For researchers, the empirical study of legal phenomena can be a treacherous path to take. The legal empiricist has to bridge the gap between law, an area in which methodology is traditionally considered to be of little importance, and methodologies of the social sciences, which are often not easily transplanted into the legal domain. This might discourage those willing to venture beyond the traditional borders of legal academia. Quantitative research is moreover sometimes met with scepticism in traditional legal circles, where qualitative methods have been the hallmark of research. This scepticism can be overcome by bringing empirical research closer to lawyers, tackling topics directly relevant to and intrinsically entwined with daily legal practice, combined with a clear indication of the limits of empirical research and a straightforward communication of its results.[75] Legal academia needs to welcome and promote collaboration between its members and computer scientists, computational linguists, and big data analysts if it wishes to remain relevant in a rapidly changing research environment and not become out of touch with recent developments.

---

67   Impact studies of the jurisprudence of the US Supreme Court had a similar effect. Wasby, supra note 28, p. 42.
68   J. Øyrehagen Sunde, 'From courts of appeal to courts of precedent – Access to the highest courts in the Nordic countries', in C.H. van Rhee & Y. Fu (eds.), *Supreme Courts in Transition in China and the West* (2017), p. 53.
69   Z. Bankowski, et al., 'Precedent in the United Kingdom', in D. MacCormick & R. Summers (eds.), *Interpreting Precedents: a comparative study* (1997), p. 315.
70   Petrick, supra note 39, pp. 7-8.
71   See Merryman's discussion of primary and secondary sources. Merryman, supra note 13, pp. 620-621.
72   See for an introduction: G. Peabody, 'Legislating from the bench: a definition and a defense', (2007) 11 *Lewis and Clark Law Review*, pp. 185-232.
73   As Billiet argues, much of the European law and economics research remains hidden from international eyes as it is only published in the native language of the researchers and caters for a local audience. C. Billiet, 'Formats for law and economics in legal scholarship: views and wishes from Europe', (2011) 12 *University of Illinois Law Review*, no. 5, pp. 1501-1508. The same argument probably applies to computational empirical research.
74   For an overview of the history and aims of the Empirical Legal Studies Movement, see Th. Eisenberg, 'The origins, nature, and promise of empirical legal studies and a response to concerns', (2011) 12 *University of Illinois Law Review*, no. 5, pp. 1713-1738.
75   See Billiet, supra note 73, pp. 1509-1511.

The judiciary needs to adapt as well. Next to the obvious advantages for legal practitioners, mining the currently underexplored knowledge embedded within the amassed output of the judiciary would considerably benefit its members and facilitate their work.[76] Legal certainty and predictability would as a consequence be enhanced. Big data analysis should therefore be welcomed, but such a paradigm shift has, at least in Belgium, not yet materialized. Measures that can easily be implemented within the judiciary and that would greatly facilitate computational analysis of case law, include *inter alia* the standardization of the structure of legal decisions, the uniform inclusion of a minimum of metadata in the decision itself, and the publication of decisions in user-friendly formats such as XML or HTML.[77] These simple measures would help researchers a great deal in unlocking the rich information embedded in case law.

My research constitutes the first large-scale application of computational analysis of case law in Belgium. As such, it can function as a pilot study. Problems that I encounter highlight areas in which amendment is necessary in order to facilitate follow-up research. Methods that I have used to overcome these hurdles and conduct my research, can serve as a starting point from which to improve. Those who are willing to explore the uncharted lands of empirical legal studies may find many hidden treasures that invite empirical testing lurking in the caves and coves of traditional legal academia. Anchors away! ▌

---

76  Taelman, supra note 31, p. 7.
77  H. Van Bossuyt & B. De Groot, 'Law via the internet: free access to law and judicial decisions; solutions and challenges from a Belgian viewpoint', (2010) 3 *International Journal of Liability and Scientific Enquiry*, no. 4, p. 275; Van Opijnen, supra note 16, pp. 470-471.