

# crese

CENTRE DE RECHERCHE  
SUR LES STRATÉGIES ÉCONOMIQUES

## How is Digital Evidence Used in the International Criminal Court? An Exploratory Analysis

YANN LECORPS, KHAOULA NAILI, MARIE OBIDZINSKI,

YVES OYTANA, TÉA TOUTOUNJI

May 2026

**Working paper No. 2026 – 04**

**CRESE**

30, avenue de l'Observatoire  
25009 Besançon  
France  
<http://crese.univ-fcomte.fr/>

The views expressed are those of the authors  
and do not necessarily reflect those of CRESE.

**UNIVERSITÉ**  
**MARIE & LOUIS**  
**PASTEUR**

# How is Digital Evidence Used in the International Criminal Court? An Exploratory Analysis\*

Yann Lecorps<sup>†</sup>   Khaoula Naili<sup>‡</sup>   Marie Obidzinski<sup>§</sup>   Yves Oytana<sup>¶</sup>  
Téa Toutounji<sup>||</sup>

May 29, 2026

## Abstract

We examine the use of digital *versus* classical evidence in cases handled by the International Criminal Court (ICC). We study and exploit variations in the use of evidence (i) between stages of proceedings, namely pre-trial and trial, and (ii) between parties in ICC trials. Since the standard of proof varies between stages of the proceedings, we frame the situation using a simple model and discuss the results we can empirically expect. The model predicts that the prosecution's reliance on both types of evidence should increase from the pre-trial to the trial stage, although the relative magnitude of the increase remains ambiguous. Our empirical results confirm that references to both categories of evidence are more frequent at trial and show that the rise is proportionally larger for classical evidence than for digital evidence. For variations in the use of evidence between the parties involved in the trial, our approach is purely empirical. We find that the relative intensity of references to digital evidence by parties varies across cases and across trial sub-phases.

**Keywords** : International Criminal Court, digital evidence, standard of proof, online platform, textual analysis, sentiment analysis

**JEL classification code** : K4

---

\*Research carried out with the support of the Institut des Etudes et de la Recherche sur le Droit et la Justice. We thank Cécile Bourreau-Dubois, Sarah Jamal, Umberto Nizza, Olivier Schaub and two anonymous referees for valuable comments and suggestions. All remaining errors are ours.

<sup>†</sup>Université Paris-Panthéon-Assas, CRED (UR 7321), 75014 Paris, France. e-mail: yann.lecorps@assas-universite.fr

<sup>‡</sup>Université Marie et Louis Pasteur, CRESE (UR 3190), F-25000 Besançon, France. e-mail: khaoula.naili@univ-fcomte.fr

<sup>§</sup>Université Paris-Panthéon-Assas, CRED (UR 7321), 75014 Paris, France. e-mail: marie.obidzinski@assas-universite.fr

<sup>¶</sup>Université Marie et Louis Pasteur, CRESE (UR 3190), F-25000 Besançon, France. e-mail: yves.oytana@univ-fcomte.fr

<sup>||</sup>Université Paris-Panthéon-Assas, CRED (UR 7321), 75014 Paris, France. e-mail: tea.toutounji@assas-universite.fr

# 1 Introduction

In international crime investigations, the use of digital evidence can be crucial, particularly when investigators are unable to access the territory where the crimes allegedly occurred. Individuals (whether witnesses, victims, or even perpetrators) may post photos or videos online (Aronson, 2018). The situation in Syria illustrates this point particularly well. Although Syria closed its borders to United Nations investigators in March 2011, the Independent International Commission of Inquiry on the Syrian Arab Republic was nevertheless able to establish facts of human rights violations. This was possible only because the Commission obtained numerous videos posted on online platforms.<sup>1</sup> However, the probative value of such information collected from the Internet is frequently questioned in criminal proceedings (Laux, 2018). Digital evidence poses specific challenges, including the difficulty of formally identifying and authenticating the source and the risk that the information is inaccurate (Hellwig, 2021).

In this article, we adopt a positive law-and-economics perspective to analyze the use of digital *versus* classical evidence in cases handled by the International Criminal Court (hereinafter, ICC). We define digital evidence as information obtained through the Internet and digital devices (most notably video material) while classical evidence refers exclusively to witness testimony, which are a key feature of ICC trials.

We begin by examining the intensity with which classical and digital evidence are referenced. The two main procedural phases, namely pre-trial and trial, are characterized, among other differences, by distinct evidentiary requirements.<sup>2</sup> To understand how this shift in evidentiary thresholds may affect references to digital as opposed to classical evidence, we develop a theoretical model.<sup>3</sup> This model builds on the economics of judicial proof literature originating with Posner (1999), and relates both to the literature on evidence production (*e.g.*, Emons and Fluet, 2009, Froeb and Kobayashi, 2001, Kaplow, 2012) and to the literature on standards of proof (*e.g.*, Christmann and Kirstein, 2023, Demougin and Fluet, 2008, Mungan, 2011, Obidzinski and Oytana, 2019, Rizzolli and Saraceno, 2013).<sup>4</sup> Rather than asking which procedural rules are optimal, we examine how the Office of the Prosecutor (OTP) adjusts the amount and type of evidence it collects in response to changes in the standard of proof. Our theoretical model shows that as the standard of proof becomes more stringent, the quantities of both digital and classical evidence generally increase. Moreover, we find that the relative increase in each type depends on their marginal collection costs and their degree of complementarity (complementarity may arise, for example, when digital evidence is used primarily to corroborate classical evidence). In particular, if classical evidence is more costly (or harder to obtain) and if complementarity is sufficiently strong, then as the standard of proof increases (as it does when moving from the pre-trial to the trial phase), the additional amount of digital evidence gathered by the OTP will tend to rise proportionally more than the amount of classical evidence (and *vice versa*).

---

<sup>1</sup>See, for example, the “human rights channel” on YouTube: <https://www.youtube.com/humanrights>.

<sup>2</sup>The pre-trial phase typically lasts only a few days and primarily consists in the factual presentation of evidence, whereas the trial phase extends over several years and provide considerable more scope for oral expressions, especially testimonies.

<sup>3</sup>The literature on the ICC remains limited and primarily focuses on the deterrent effect of the ICC. See, for instance, Gilligan (2006), Lecorps (2021), and Sutter et al. (2006).

<sup>4</sup>In particular, Christmann and Kirstein (2023) analyze the intensity of prosecutorial investigative efforts within an inquisitorial framework. Their analysis, however, focuses on coordination issues between the judge and the prosecutor in a setting where both actors can investigate the evidence. They show that increasing the standard of proof can strengthen incentives to invest in evidence production, but also exacerbates coordination failures by promoting duplication of efforts and increasing the likelihood of errors.

We then examine references to classical and digital evidence at the two stages (pre-trial and trial) by conducting a textual analysis dictionary-based count of 756 hearing transcripts of the three ICC cases involving crimes committed no later than 2010 and that have completed the full trial stage: *Gbagbo & Blé Goudé*, *Al Hassan*, and *Yekatom & Ngaiïssona*. Our main empirical findings regarding references to each type of evidence at the pre-trial *versus* trial stage are as follows: first, the proportion of transcripts containing references to both classical and digital evidence is higher at trial than at pre-trial, and this effect persists after controlling for case and word count; second, the increase is relatively larger for classical evidence than for digital evidence. These results seem to corroborate the mechanisms identified in the theoretical model: between the two phases, the collection of both types of evidence increases (result 1); classical evidence appears to have a higher marginal probative value than digital evidence, and its (relative) marginal cost of collection is likely not very high (interpretation of result 2).

Finally, we ask whether the parties differ in their references to the two categories of evidence. To address this question, we refine our textual analysis by examining references to evidence *by party*, with particular emphasis on the OTP and the defense. We examine separately the subphases of the trial stage, distinguishing in particular between the OTP and the defense trial hearings, as the role and position of the OTP and the defense vary across these subphases. Although we do not observe significant differences in the parties' references to digital evidence across the three cases, both parties refer to digital evidence less frequently during the subphase devoted to the defense's presentation of evidence than during the OTP's corresponding subphase. Moreover, although the ratio of references to digital *versus* classical evidence tends to be higher for the OTP in the *Al Hassan* case, this difference largely disappears in the *Yekatom & Ngaiïssona* case. In the latter, the overall ratio of digital to classical evidence is higher for both parties. Finally, we conduct an exploratory sentiment analysis of the words spoken by the OTP and the defense across the subphases of each case. We find that their sentiment indexes tend to be positively correlated. However, we do not observe any clear relationship between their sentiment indexes and their ratios of references to digital *versus* classical evidence.

The remainder of the paper is organized as follows. Section 2 presents the institutional context and describes the database. Section 3 develops a theoretical model and provides empirical evidence on the references to digital *versus* classical evidence across the pre-trial and trial stages. Section 4 extends the textual analysis to examine how each party (focusing primarily on the OTP and the defense) relies on the two categories of evidence. Section 5 concludes.

## 2 Institutional context and database description

Established in 2002, the International Criminal Court (ICC) is a court of last resort whose mandate is to prosecute and deter the most serious international crimes, including crimes against humanity, genocide, war crimes, and crimes of aggression (Articles 5, 6, 7, 8, and *8bis* of the Rome Statute).<sup>5</sup> A total of 124 States have ratified the Rome Statute, the Court's founding treaty. A situation may be brought before the ICC in three principal ways: (i) at the request of a State Party to the Rome Statute; (ii) at the request of the United Nations Security Council; or (iii) through the Prosecutor's *proprio motu* powers to open an investigation (Articles 13, 14, 15, *15bis*, *15ter*). Criminal proceedings

---

<sup>5</sup>The Court's jurisdiction is complementary to that of national courts and applies only when States are unable or unwilling to prosecute. For a list of States Parties to the Rome Statute, see <https://asp.icc-cpi.int/states-parties>. The Rome Statute is available at: <https://www.icc-cpi.int/sites/default/files/RS-Eng.pdf>.

before the ICC are organized into four main stages, in addition to a possible appeal stage and the enforcement of sentences:

1. **Preliminary examination.** The OTP assesses whether there is sufficient information indicating that a crime within the ICC’s jurisdiction has been committed and whether the gravity of the situation warrants further action.
2. **Investigation.** The OTP collects evidence and identifies a suspect.<sup>6</sup>
3. **Pre-Trial.** During the confirmation of charges hearing, which follows the initial appearance, a panel of three pre-trial judges hears submissions from the prosecution, the defense, and the legal representatives of victims. The judges then determine whether there is *sufficient evidence to establish substantial grounds to believe that the person committed each of the crimes charged*. If so, the case proceeds to trial.
4. **Trial.** Before a bench of three trial judges, the OTP must prove the guilt of the accused. The Court may convict only if it is convinced *beyond reasonable doubt*.

Note that the standard of proof<sup>7</sup> applied at trial (*beyond reasonable doubt*) is more stringent than the standard used at the pre-trial stage (*substantial grounds to believe*).<sup>8</sup> In what follows, we focus on the pre-trial and trial phases.

An ICC trial is structured into three main successive phases: (i) a first phase during which the Office of the Prosecutor (OTP) presents its evidence; (ii) a second, very brief phase during which the victims’ “views and concerns” are presented; and (iii) a third phase devoted to the presentation of the defense’s evidence. During phase (i), the OTP introduces the incriminating classical and/or digital evidence it has collected. However, the defense is not passive and may cross-examine the prosecution’s evidence in accordance with the adversarial principle. A similar dynamic applies in phase (iii): the defense presents the exculpatory evidence in its possession, after which the prosecution may cross-examine that evidence in order to challenge it. Regarding phase (ii), victims’ representatives may present evidence only with the express authorization of the Trial Chamber. When such evidence is admitted, it may be cross-examined by both the defense and the prosecution.

We compiled a database of 756 transcripts from the ICC pre-trial and trial hearings of three cases. A transcript is the official written version of the audio (or video) recording of an ICC hearing and includes the statements made by the various participants. Each hearing corresponds to a single day of proceedings and may last from a few minutes to several hours.<sup>9</sup> These documents are publicly available in the ICC Legal Tools Database and on the ICC website.<sup>10</sup> The sample is restricted to the most recent cases that have completed the full trial stage by 2025 at the latest and whose crimes occurred no later than 2010.<sup>11</sup> The year 2010 has been chosen to allow for some development of the

---

<sup>6</sup>A warrant of arrest may be issued at any point after the initiation of an investigation by the pre-trial Chamber if there are *reasonable grounds* to believe that the person has committed a crime within the Court’s jurisdiction, particularly to ensure the person’s appearance at trial (Article 58 of the Rome Statute).

<sup>7</sup>The standard of proof is “the level of certainty the adjudicative authority has to reach in order to establish guilt in a criminal proceeding” (Rizzolli, 2016).

<sup>8</sup>Articles 61 and 66 of the Rome Statute of the ICC.

<sup>9</sup>Consequently, a transcript may range from a single page to more than one hundred pages, representing several hours of presentations before the judges.

<sup>10</sup>A small number of transcripts for each case are missing because they are not available on either platform.

<sup>11</sup>We therefore excluded cases that remained at the pre-trial stage because the defendant was not in the Court’s custody, or because proceedings were interrupted by the death of the accused or by a guilty plea prior to the start of the trial (e.g., *The Prosecutor v. Ahmad Al Faqi Al Mahdi*).

internet and social media. Three cases meet these criteria: *The Prosecutor v. Laurent Gbagbo and Charles Blé Goudé*, *The Prosecutor v. Al Hassan Ag Abdoul Aziz Ag Mohamed Ag Mahmoud*, and *The Prosecutor v. Alfred Yekatom and Patrice-Edouard Ngaïssona*.

The oldest case concerns crimes committed in the context of the post-election conflict in Côte d’Ivoire between December 16, 2010, and April 12, 2011. The conflict has been characterized by large-scale violence and population displacement. Laurent Gbagbo, former President, and Charles Blé Goudé were accused of participating in a common plan to retain power by force following Alassane Ouattara’s electoral victory. Laurent Gbagbo was arrested in April 2011, and Alassane Ouattara was declared president of the Republic the following month. A. Ouattara then promptly referred the situation to the ICC Prosecutor for an investigation into the massacres that had occurred during the Ivorian crisis. The trials of the two defendants were joined in March 2015, and the joint trial began in January 2016. In 2018, at the end of the prosecution’s presentation of evidence, the Court invited the defense team to choose either to present their own witnesses or to request an immediate acquittal (*no case to answer*). The defense elected to file a motion to dismiss the case. In January 2019, the Court acquitted L. Gbagbo and C. Blé Goudé of all charges, ruling that the evidence presented by the Prosecutor was insufficient.

The case of *The Prosecutor v. Al Hassan Ag Abdoul Aziz Ag Mohamed Ag Mahmoud* concerns the situation in Mali. The accused, a member of the jihadist group Ansar Eddine, was convicted on June 26, 2024, of war crimes and crimes against humanity (torture, mutilation, unlawful trials, and religious persecution) committed between April 2, 2012, and January 29, 2013, in Timbuktu, northern Mali. The conflict, which erupted in northern Mali in 2012 following the mobilization of a coalition of jihadist and separatist groups, remains ongoing. Since then, the areas affected by the fighting have progressively expanded southward.

Finally, we examined the case of *The Prosecutor v. Alfred Yekatom and Patrice-Édouard Ngaïssona* (Central African Republic). The two defendants were convicted at first instance on July 24, 2025, for war crimes and crimes against humanity committed between September 2013 and February 2014 (murder, torture, attacks against civilians, forced displacement of populations, destruction of places of worship, and persecution). Alfred Yekatom commanded a group of approximately 3,000 fighters affiliated with the anti-Balaka militia, while Patrice-Édouard Ngaïssona served as the national coordinator of the anti-Balaka. Muslim civilians were among the primary targets of these crimes. The situation in the country is still highly unstable to this day.

In these three countries (Côte d’Ivoire, Mali, and the Central African Republic), internet access remained limited in 2010. According to the International Telecommunication Union (ITU), these countries were among the last offering 3G service commercially, relying primarily on 2G networks instead. Moreover, in 2010, in Africa, only 9.6% of the population used the internet, a figure far below both the world average (30%) and the average for developing countries (21%).<sup>12</sup> The figures are even lower for the countries concerned at the time of the events, as shown in table 1. While internet penetration remained extremely low, the number of mobile-cellular subscriptions per 100 inhabitants was significantly higher, reflecting the widespread diffusion of mobile phones despite the limited availability of internet access at the time (see “Mobile subscriptions (per 100 inh.)” in table 1). Furthermore, we observe some variation between countries. Mali is the most advanced country in terms of Internet and mobile network development, while the Central African Republic is the least advanced.

---

<sup>12</sup>Source: ICT Indicators Database.

Table 1: Individuals using the Internet (% of population) and mobile-cellular subscriptions (per 100 inh.) at the time of the crimes

Country	Year(s)	Internet users (%)	Mobile subscriptions (per 100 inh.)
Côte d'Ivoire	2010	2.7	69.4
Côte d'Ivoire	2011	2.9	75.5
Mali	2012	2.8	86.2
Mali	2013	3.5	113
Central African Republic	2013	2.5	29.3
Central African Republic	2014	2.7	25

Source: Indicator “Individuals using the Internet (% of population)” and “Mobile-cellular subscriptions (per 100 inh.)” International Telecommunication Union (ITU).<sup>13</sup>

Table 2 summarizes the key information concerning the three cases, *i.e.*, the names of the accused, the country, the period during which the crimes were committed, and the number and proportion of transcripts associated with each case in our database.

Table 2: Transcripts by case

Accused	Country	Years of crimes	n	%
Gbagbo & Blé Goudé	Côte d'Ivoire	2010-2011	243	32.1
Al Hassan	Mali	2012-2013	207	27.4
Yekatom & Ngaissona	Central African Republic	2013-2014	306	40.5
<b>Total</b>			756	100

Table 3 shows that 95.2% of the hearing transcripts originate from the trial phase. This is consistent with the fact that the trial phase is significantly longer than the pre-trial phase. In general, the pre-trial phase lasts only a few days, whereas the trial phase may extend over several years. Furthermore, some portions of the hearings are held in private session and therefore are not transcribed. Private sessions occur when witnesses face serious risks of retaliation or when sensitive information is presented. In such cases, the notation “page redacted” appears, and the corresponding page is left blank. The fact that part of the hearings were not made public for reasons of witness protection suggests that the number of references to classical evidence may have been underestimated. Indeed, approximately 12% of the pages of the transcripts were blanked.<sup>14</sup>

Table 3: Transcripts by stage of proceedings

	n	%
Pre-trial	36	4.8
Trial	720	95.2
<b>Total</b>	756	100

<sup>13</sup>See <https://datahub.itu.int>.

<sup>14</sup>The proportions of pages that were redacted, by phase and by case, are reported in table 7 in Appendix A.

Regarding the references to classical and digital evidence in the hearing transcripts, we defined a list of keywords, including their lexical variants identified using the Thesaurus API.<sup>15</sup> Classical evidence refers here exclusively to witness testimonies, whereas digital evidence refers to information obtained through the Internet and digital devices.<sup>16</sup> This definition is narrower than that used in Hellwig (2021), where digital evidence is defined as “electronic evidence that is generated or converted to a numerical format.”<sup>17</sup>

All transcripts were cleaned, and we counted the occurrences of keywords related to digital and classical evidence, enabling comparisons of their frequency across stages of the proceedings (Section 3) and across parties involved in the case (Section 4). The procedure for cleaning the transcripts and counting keywords is detailed in Appendix D.<sup>18</sup>

With respect to classical evidence, the most frequent word is “witness,” which appears on average 80.8 times per hearing transcript, while the word “testimony” appears on average 6.8 times per transcript. With respect to digital evidence, we relied on a list of keywords corresponding to a broad view of references to material originating from online social networks and the Internet. The word “video” is the most frequent, with an average of 11.4 references per transcript, whereas several keywords—particularly those related to online social networking—do not appear in any hearing transcript. Unsurprisingly, the three online platforms with the highest average number of citations are Facebook, Youtube, and Whatsapp, although there is substantial heterogeneity across cases.<sup>19</sup>

Table 4 compares the proportion of transcripts containing references to each category of keywords. We find that the share of transcripts containing references to digital evidence (81%) is slightly lower than the share containing references to classical evidence (98.4%).<sup>20</sup> Moreover, no transcript contains a reference to digital evidence without also containing at least one reference to classical evidence.<sup>21</sup>

Table 4: Frequency of digital and classical evidence references

	Ref digital evidence	No ref digital evidence	Sum
Ref classical evidence	81%	17.4%	98.4%
No ref classical evidence	0%	1.6%	1.6%
Sum	81%	19%	100%

<sup>15</sup>The list of keywords related to classical (digital) evidence appears in Table 8 (Table 9) in the Appendix A. Lexical variants and synonyms were identified using the Thesaurus API from API Ninjas: <https://api-ninjas.com/api/thesaurus>.

<sup>16</sup>Fact-finding mechanisms, such as the International, Impartial and Independent Mechanism (IIIM, see <https://iiim.un.org/>), gather information from multiple sources, including the Internet, and distinguish among open-source, online, and social media content.

<sup>17</sup>Because our aim is to emphasize information obtained from online platforms and the Internet, our list of keywords does not include terms such as “satellite imagery” or “GPS information.”

<sup>18</sup>Importantly, the words spoken by a witness called by a party are not attributed to that party.

<sup>19</sup>Facebook, Youtube and Twitter are categorized as social media, and created respectively in 2004, 2005 and 2006. Whatsapp, which is an interpersonal messenger, was created in 2009. The Instagram app has been released to the public in 2012. Telegram was created in 2013, and Snapchat in 2011. Tiktok was launched in 2016. Linkedin and Viadeo were created in 2003 and 2004, but are a professional digital networks.

<sup>20</sup>See Tables 10 and 11 for the number and percentage of hearing transcripts that contain at least one reference to each category of keywords.

<sup>21</sup>One possible interpretation is that digital evidence is primarily used to corroborate classical evidence and is therefore almost always invoked alongside it.

### 3 Use of digital and classical evidence between the pre-trial and trial stages

In this section, we examine both theoretically and empirically digital evidence, in comparison with classical evidence, at the different stages of the proceedings. The first subsection theoretically analyzes how the Office of the Prosecutor (OTP)’s reliance on digital *versus* classical evidence is affected by the stringency of the standard of proof. We choose to focus specifically on the effect of the stringency of the standard of proof because it constitutes a major procedural difference (although certainly not the only one) between the pre-trial and trial phases. The second subsection presents empirical evidence on how references to each type of evidence vary across phases (pre-trial *versus* trial).<sup>22</sup>

#### 3.1 A model of evidence collection

In this subsection, we present a simple model to study how the OTP allocation of time and effort between the collection of classical *versus* digital evidence is affected by the stringency of the standard of proof.<sup>23</sup> The way we model evidence production builds on the classical literature on rules of proof (Kaplow, 2012, Demougin and Fluet, 2008). Although the simple model we propose below cannot capture the full complexity of reality, we believe it can nevertheless offer insights into how the OTP may trade off between the use of both types of evidence.

##### 3.1.1 The basic framework

There are two types of evidence: classical (in quantity  $C$ ) and digital (in quantity  $D$ ). Evidence is represented as producing a single indicator  $f(C, D)$ , which is defined as the total probative value of all admissible evidence (both classical and digital).<sup>24</sup> The total probative value can result to be above or below the evidence threshold  $\bar{x}$ .

We assume  $f_C(C, D) > 0$ ,  $f_D(C, D) > 0$ ,  $f_{CC}(C, D) < 0$ ,  $f_{DD}(C, D) < 0$ : the probative value of evidence increases with the quantities of both types of evidence, but at a decreasing rate.<sup>25</sup>

---

<sup>22</sup>All source code used in this section, as well as in Section 4, is available from the authors upon request.

<sup>23</sup>We focus on the OTP because it is the party that bears the burden of producing evidence. It would be possible to extend our model to analyze the defense’s choice between relying on digital *versus* classical evidence. However, doing so would require introducing additional complexity to account for strategic interactions between the two parties. In such a model, the trade-off identified here would likely apply, at least in part, to the defense as well. The main difference is that, given its role, the defense primarily responds to the evidence produced by the prosecution and therefore has less discretion over the type of evidence it relies upon.

<sup>24</sup>Note that we do not address the issue of evidentiary admissibility. The quantities  $C$  and  $D$  refer to evidence that has already been deemed admissible at an earlier stage of the proceedings, which we do not model here. Our focus is instead on whether the set of admissible evidence possesses sufficient probative value to satisfy the standard of proof. There is, however, a substantial law and economics literature concerned with the rules of admissibility. For example, Posner (1999) argues that “searches should be deemed illegal only if the evidentiary benefits do not equal or exceed the costs of the search to the victim.” This principle is consistent with Article 69 of the Rome Statute, which instructs the ICC to take into account “the probative value of the evidence and any prejudice that such evidence may cause to a fair trial or to a fair evaluation of the testimony of a witness.” Other rationales have been advanced in the literature to justify the exclusion of certain forms of evidence. These include for instance the cognitive costs imposed on fact-finders when interpreting complex evidence to reach a verdict (Lester et al., 2012), as well as the risk that some forms of hard evidence may mislead decision-makers when the beliefs of litigants and adjudicators are not well coordinated (Bull and Watson, 2019).

<sup>25</sup> $f(C, D)$  can be interpreted as a production function, with the inputs being  $C$  and  $D$ .

Additionally, we make the following assumptions:<sup>26</sup>

- i.  $f_C(C, D) > f_D(C, D)$  if  $C = D$ .
- ii.  $f_{CD}(C, D) > 0$ .

Assumption (i) encapsulates the idea that classical evidence has a higher marginal probative value than digital evidence. Assumption (ii) means that the two types of evidence are complements, in the sense that increasing the quantity of one type of evidence increases the marginal probative value of the other type of evidence.<sup>27</sup> We focus primarily on the scenario in which digital and classical evidence are complements, as this scenario is the most likely: digital evidence, in order to be regarded as credible, regularly needs to be supported by other types of more classical evidence. Indeed, as de Arcos Tejerizo (2023) notes, “digital evidence, regardless of its provenance, should be reasonably supported by either forensic science and/or witness statements in order to be reliable and effectively used in the courtroom.” Freeman (2018), in the context of international criminal investigations, provides several real-world examples of cases in which digital evidence is used to corroborate other forms of evidence, arguing that “the evidentiary value of [digital evidence] is still considered low, while other investigative methods such as forensic missions might lead to probative evidence, but at a prohibitively high cost.”<sup>28</sup>

Collecting evidence is costly. We denote by  $\alpha$  ( $\beta$ ) the marginal cost of gathering a unit of classical (digital) evidence. We assume that the prosecutor aims at minimizing the total cost of producing evidence, while still satisfying the standard of proof, with  $\bar{x}$  the level (or stringency) of that standard. Therefore, the problem faced by the prosecutor is:

$$\begin{cases} \min_{C, D} & \alpha C + \beta D \\ \text{subject to} & f(C, D) \geq \bar{x} \end{cases} \quad (1)$$

Solving (1) shows that the optimal levels of classical evidence ( $C^*$ ) and digital evidence ( $D^*$ ) are characterized by:<sup>29</sup>

$$\begin{cases} \frac{f_C(C^*, D^*)}{f_D(C^*, D^*)} = \frac{\alpha}{\beta} \\ f(C^*, D^*) = \bar{x} \end{cases} \quad (2)$$

The first equation of (2), together with assumption (i), imply that *if the cost of gathering each type of evidence is the same (i.e.  $\alpha = \beta$ ), then the optimal level of classical evidence is higher than the optimal level of digital evidence.*

---

<sup>26</sup>Subscripts denote partial derivatives.

<sup>27</sup>Alternatively, one could imagine that the two types of evidence are substitutable. We discuss this case at the end of this subsection.

<sup>28</sup>Another argument in support of the hypothesis of evidentiary complementarity is that one would expect complementarity to promote the admissibility of evidence (even though, as noted above, we do not model the choice of the Court regarding the admissibility of evidence). Indeed, in one of their extensions, Bull and Watson (2019) consider a setting in which two distinct pieces of material evidence may exist (*e.g.*, classical and digital evidence). They show that the court may sometimes require that the evidence be admissible only as a bundle (*e.g.*, some classical and digital evidence must be presented jointly). Moreover, Lester et al. (2012) show that the exclusion of evidence is unnecessary when “all the available evidence fits together tightly into one coherent story (formally, when all pieces of evidence are complementary with each other).”

<sup>29</sup>The details to solve this program and for the results given in this subsection are given in the Appendix B.

### 3.1.2 The effect of a more stringent standard of proof on evidence collection

In the following, we study the effect of a change in the level of the standard of proof ( $\bar{x}$ ) on  $C^*$  and  $D^*$  ( $C^*$  and  $D^*$  are functions of  $\bar{x}$ , although we do not make that explicit in order to streamline the notation).

**Result 1.** *When the standard of proof becomes more stringent, the quantities of both digital and classical evidence that the prosecutor gathers increase.*

Note that this result implies that the total cost of gathering evidence ( $\alpha C^* + \beta D^*$ ) increases with the standard.

Does an increase in the standard of proof increases relatively more the gathering of classical evidence, or conversely?

**Result 2.** *For a small increase in the standard, the quantity of classical evidence increases more than the quantity of digital evidence if:*

$$\beta f_{CC}(C^*, D^*) - \alpha f_{DD}(C^*, D^*) + (\beta - \alpha) f_{CD}(C^*, D^*) > 0 \quad (3)$$

Condition (3) may or may not be satisfied. Thus, *when the standard of proof rises, the quantity of classical evidence may increase more than the quantity of digital evidence, or conversely.* More specifically, whether that condition holds depends on three effects (corresponding to the three terms on the left-hand side of (3)).

The first effect (the first term in (3)) is negative, in the sense that the stronger is this effect, the more the prosecutor tends to favor the collect of digital rather than classical evidence. It is related to the decreasing rate with which the probative value of classical evidence increases with its quantity (*i.e.* the assumption  $f_{CC}(C, D) < 0$ ). The second effect (the second term in (3)) is similar to the first one, except that it is positive and related to the decrease in the marginal probative value of digital evidence. Taken together, these two effects mean that, *when the standard of proof increases, if the probative value of an additional unit of classical evidence decreases sharply with its quantity (when compared to the decrease in the probative value of an additional unit of digital evidence), then the prosecutor may want to intensify relatively more his effort to gather digital rather than classical evidence (and conversely).*

The third effect (the third term in (3)) relates to the degree of complementarity between the two types of evidence. To illustrate the intuition behind this effect, let us assume that the prosecutor wants to increase the probative value originating from classical evidence to reach a more stringent standard of proof,<sup>30</sup> and that the marginal cost of gathering classical evidence is *higher* than the cost for gathering digital evidence ( $\alpha > \beta$ ). In this specific case, the prosecutor has basically two choices: (i) increasing the quantity of classical evidence to *directly* increase its probative value, or (ii) increasing the quantity of digital evidence, thus increasing *indirectly* the probative value of classical

---

<sup>30</sup>This is a simplification for illustration purpose, since in our model, the probative values of each type of evidence are always linked through the “production” function  $f$ .

evidence due to the complementarity of the two types of evidence (increasing the quantity of digital evidence reinforce the probative value of each unit of classical evidence already gathered). Because the cost of a unit of classical evidence is greater than the cost of a unit of digital evidence, the prosecutor may lean toward option (ii) (that is, increasing the quantity of digital evidence), because it is cheaper to do so rather than choosing option (i) (investing directly in gathering more classical evidence). This effect is stronger when the degree of complementarity increases (*ceteris paribus*). Moreover, note that when the marginal costs are the same for each type of evidence ( $\alpha = \beta$ ), that third effect cancels out, while it works in the opposite direction when the marginal cost to produce evidence is relatively *lower* for classical evidence ( $\alpha < \beta$ ). To summarize, *when the standard of proof increases, if the marginal cost of classical evidence is higher (lower) than the marginal cost of digital evidence, the more the two types of evidence complement each other, the more (the less) the prosecutor will intensify his effort to gather digital evidence (when compared to his effort to gather classical evidence).*

What happens if we relax assumption (ii) and instead assume that the two types of evidence are substitutable? Two scenarios may arise. First, if the two types of evidence are highly substitutable, we obtain a corner solution in which only one type of evidence is collected by the OTP in equilibrium. For example, if  $\alpha = \beta$  (the marginal cost of collecting each type of evidence is identical), then the OTP collects only classical evidence, since assumption (i) implies that, for an equal quantity, classical evidence has a higher probative weight. Second, if the two types of evidence are substitutable but to a lesser extent, our results remain generally qualitatively similar to those obtained under assumption (ii).<sup>31</sup>

### 3.2 References to digital and classical evidence during the pre-trial and trial stages

In this section, we examine the frequency of references to digital and classical evidence at the pre-trial and trial stages. During the pre-trial phase, judges hear from the prosecution, the defense, and the legal representatives of the victims. Then, they shall, on the basis of the hearing, “determine whether there is sufficient evidence to establish *substantial grounds to believe* that the person committed each of the crimes charged” (Article 61 of the Rome Statute). During the trial, the prosecution must prove the defendant’s guilt. The court must be convinced *beyond reasonable doubt* to convict the accused.

Cross-table 5 provides an initial overview of the relationship between references to evidence and procedural stages<sup>32</sup>. First, we observe that references to both types of evidence are more frequent during the trial stage, with a clear under-representation ( $\approx -10$  pp) of transcripts containing references to classical and digital evidence at the pre-trial stage. Second, references to classical evidence outnumber those to digital evidence at both the pre-trial ( $\approx +14$  pp) and trial ( $\approx +18$  pp) stages, with a lower digital-to-classical ratio at trial. Third, Table 5 does not reveal any clear over- or under-representation across cases with respect to references to classical evidence, whereas there is

---

<sup>31</sup>Note that it remains theoretically possible for the degree of substitutability to be low enough to preserve an interior solution, yet high enough to reverse some of our comparative statics results. In particular, it may be the case that the amount of one type of evidence decreases when the standard of proof increases, contrary to what Result 1 suggests.

<sup>32</sup>The raw differences in evidence references for each specific (*case*  $\times$  *phase*) are available in table 12 in Appendix C.

a slight over-representation ( $\approx +2$  pp) of transcripts containing references to digital evidence in the *Yekatom & Ngaïssona* case. Finally, we note that the digital-to-classical ratio reaches its highest level during the pre-trial phase (0.84) and for the *Yekatom & Ngaïssona* case (0.84), perhaps indicating a comparatively greater reliance on digital evidence in that phase and in that case.

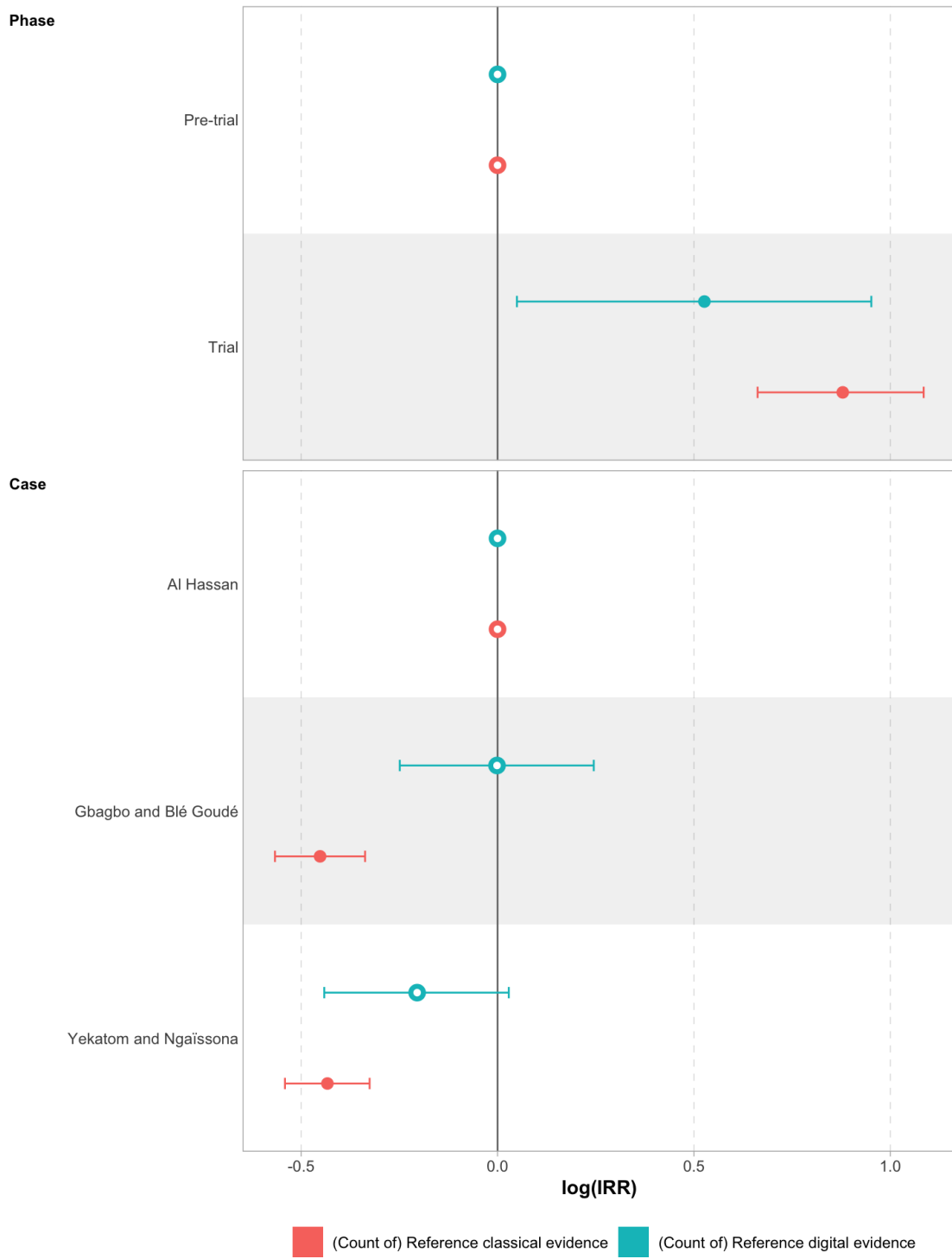
Table 5: Proportion of transcripts with references to classical and digital evidence by stage and case

	Classical evidence	Digital evidence	Ratio digital/classical
Pre-trial	86.1%	72.2%	0.84
Trial	99%	81.4%	0.82
Gbagbo & Blé Goudé	97.9%	79%	0.81
Al Hassan	98.6%	79.7%	0.81
Yekatom & Ngaïssona	98.7%	83.3%	0.84
All	98.4%	81%	0.82

Estimates from a negative binomial regression, modeling the count of references to classical (digital) evidence as a function of the stage, the case, and the word count (offset) per transcript, help to disentangle the specific effects of each variable. For greater readability, we present a coefficient plot in figure 1, while the detailed table of regression results is provided in Appendix C (Table 13). These estimates show that the “trial” effect persists even after controlling for case and word count. Note that the coefficient associated with each modality reflects the difference in log-rates relative to the reference modality. For example, the coefficient of 0.879 associated with the “Trial” modality in Model 1 indicates that the trial phase is associated with 2.4 ( $\exp(0.879)$ ) times more references to classical evidence than the pre-trial phase, holding all other variables constant. Thus, unsurprisingly, the estimates predict a higher number of references to both types of evidence at trial than at pre-trial, although the increase is relatively larger for classical evidence than for digital evidence. Furthermore, it should be noted that, as table 7 shows, the proportion of redacted pages in the trial phase is higher than in the pre-trial phase. This implies that if, as one might expect, redacted pages lead to an underestimation of references to classical evidence, then the observed effect for classical evidence should be even more reinforced. Concerning the “case” variable, we find no statistically significant differences in references to digital evidence by transcription across cases. However, the *Yekatom & Ngaïssona* and *Gbagbo & Blé Goudé* cases are associated with approximately 35% ( $\approx (1 - \exp(-0.4))$ ) fewer references to classical evidence compared to the *Al Hassan* case.<sup>33</sup>

<sup>33</sup>Regarding the *Gbagbo & Blé Goudé* case, recall that the proceedings have stopped at the end of the OTP’s presentation of evidence during trial, with both accused being acquitted. Therefore, only references to evidence in this initial OTP subphase are taken into account, as there is no “victims’ phase” nor “defense phase” in this context.

Figure 1: References (count) to classical and digital evidence



These results are consistent with our theoretical Result 1, which states that as the standard of proof becomes more stringent (as is the case when moving from pre-trial to trial in ICC proceedings), the amount of both types of evidence increases. In relation to our theoretical Result 2, they also provide (indirect) support for the assumption that classical evidence has a higher marginal probative value, and that the cost of collecting classical evidence should not be too high relative to that of collecting digital evidence.<sup>34</sup>

The empirical analysis above provides a general overview of how digital and classical evidence are referenced at the two main stages of the judicial proceedings. The next section goes a step further by distinguishing between the words used by each party (i.e., the prosecution and the defense). The aim is to shed light on the strategies employed by each party in their use of digital and classical evidence.

## 4 Use of digital and classical evidence by parties

This section examines whether members of the prosecution and defense counsel differ in their use and presentation of digital *versus* classical evidence. These two parties indeed occupy distinct roles in the trial process. During the first main stage of the trial, the prosecution—bearing the burden of proof (Kaplow, 2012, Talley, 2013)—must provide the court with sufficiently compelling evidence to establish that the accused committed the alleged crimes. More precisely, the prosecutor must persuade the judges *beyond any reasonable doubt*. In practice, members of the Office of the Prosecutor (OTP) typically work on multiple cases and rely on a wide range of professionals, including investigators and analysts. As a result, and in order to meet the burden of proof, we can expect them to be better positioned to draw on diverse sources of evidence than defense teams.<sup>35</sup> Although highly experienced lawyers, defense counsel have a distinct role: they must show that guilt has not been established beyond a reasonable doubt, notably by cross-examining prosecution witnesses and by presenting their own witnesses and evidence.<sup>36</sup> In this context, it is plausible that the prosecution and the defense adopt different strategies in their references to digital evidence. Building on these considerations, we hypothesize that prosecutors are more likely to rely on digital evidence than defense attorneys.

In this section, we focus mainly on the *Al Hassan* and *Yekatom & Ngaiïssona* cases. We also present the main findings from the *Gbagbo & Blé Goudé* case. However, note that the procedural history of this case is highly atypical, which makes meaningful comparison with the *Al Hassan* and *Yekatom & Ngaiïssona* cases more challenging. First, the proceedings against Laurent Gbagbo and Charles Blé Goudé were initially registered as two separate cases, with distinct transcripts for the preliminary stages. Second, after the OTP completed its presentation of evidence, the defense filed a motion for

---

<sup>34</sup>Note that, in the theoretical model, we compared the use of classical *versus* digital evidence by looking at the quantity of admissible evidence of each type brought forward by the OTP, whereas in the empirical analysis we rely on the number of references. Thus, caution is warranted when interpreting our results, as contextual factors (such as the way evidence is introduced in the courtroom and pattern of discourses) may differently bias the number of references to classical and digital evidence, when the number of references is used as a proxy for the quantity of admissible evidence. However, our main objective is to compare the relative use of each type of evidence, and that comparison should remain valid as long as the aforementioned biases remain consistent across stages.

<sup>35</sup>At the ICC, no rule of evidence specifies *ex ante* the admissible methods of proof or their probative value. Judges are free to assess the weight of evidence.

<sup>36</sup>Another difference between the two sides is that members of the prosecution can be viewed as “repeat players,” whereas defense counsel are often closer to “one-shot players.” Consequently, each side may be expected to “play the litigation game differently” (Galanter, 1974).

acquittal on the grounds that the evidence presented up to that point was insufficient. The Court ultimately acquitted both defendants, a decision that was later upheld on appeal. Because of the timing of this acquittal, the phases involving the victims’ representatives and the defense did not take place.

We conduct a textual analysis of the transcripts to assess how frequently words related to digital and classical evidence appear in statements made by each party. To do so, we first compiled a list identifying all speakers present at each hearing, together with their respective roles in the proceedings. This classification allows us to distinguish among members of the prosecution team, defense counsel, legal representatives of victims, the judges, and other participants (mainly witnesses, but also some administrative staff). We then applied Natural Language Processing (NLP) tools to search, for each party, the predefined sets of keywords associated with digital and classical evidence (see Tables 8 and 9).<sup>37</sup>

Because this section focuses on party-specific behavior, we distinguish, beyond the pre-trial stage, several subphases of the trial: OTP hearings, hearings involving the legal representatives of victims, and defense hearings.<sup>38</sup> Table 6 presents an overview of the relative size of these phases, measured in total word counts. In all cases, the pre-trial phase is considerably shorter than the trial phase. Within the trial phase, the largest share of the transcripts corresponds to the OTP hearings, followed by the defense hearings.

Table 6: Number of words in pre-trial and trial

Case	Phase	Subphase	Number of words
Gbagbo & Blé Goudé	Pre-trial	All	299,834
	Trial	Trial hearing (OTP)	3,279,632
		Trial hearing (victims)	0
		Trial hearing (defense)	0
		Others	419,746
Al Hassan	Pre-trial	All	88,105
	Trial	Trial hearing (OTP)	1,477,342
		Trial hearing (victims)	33,386
		Trial hearing (defense)	536,187
		Others	234,695
Yekatom & Ngaiissona	Pre-trial	All	130,436
	Trial	Trial hearing (OTP)	3,396,273
		Trial hearing (victims)	129,954
		Trial hearing (defense)	567,567
		Others	275,381

Figures 2, 3, and 4 present, for each case, the distribution of spoken words by participant and by procedural phase. The analysis includes only the interventions of the OTP, the defense, the legal representatives of victims, and the judges. Across the cases, defense counsel speak the most, followed

<sup>37</sup>See Frankenreiter and Livermore (2020) for a comprehensive review of the literature on computational methods, including NLP, applied to legal texts.

<sup>38</sup>Further methodological details are provided in Appendix D.

by the OTP, the judges, and, finally, the legal representatives of victims. During the trial phase, the defense and the OTP are the two most prominent participants. In both the *Al Hassan* and *Yekatom & Ngaïssona* cases, the relative distribution of speaking time between these two parties remains broadly consistent across the OTP hearings and the defense hearings. The legal representatives of victims intervene only marginally, except during the phase specifically dedicated to them (after the OTP hearings and before the defense hearings). Finally, the judges play a relatively limited role during the pre-trial phase but become more active once the trial begins.

Figure 2: Distribution of words by participant for each phase (*Al Hassan*)

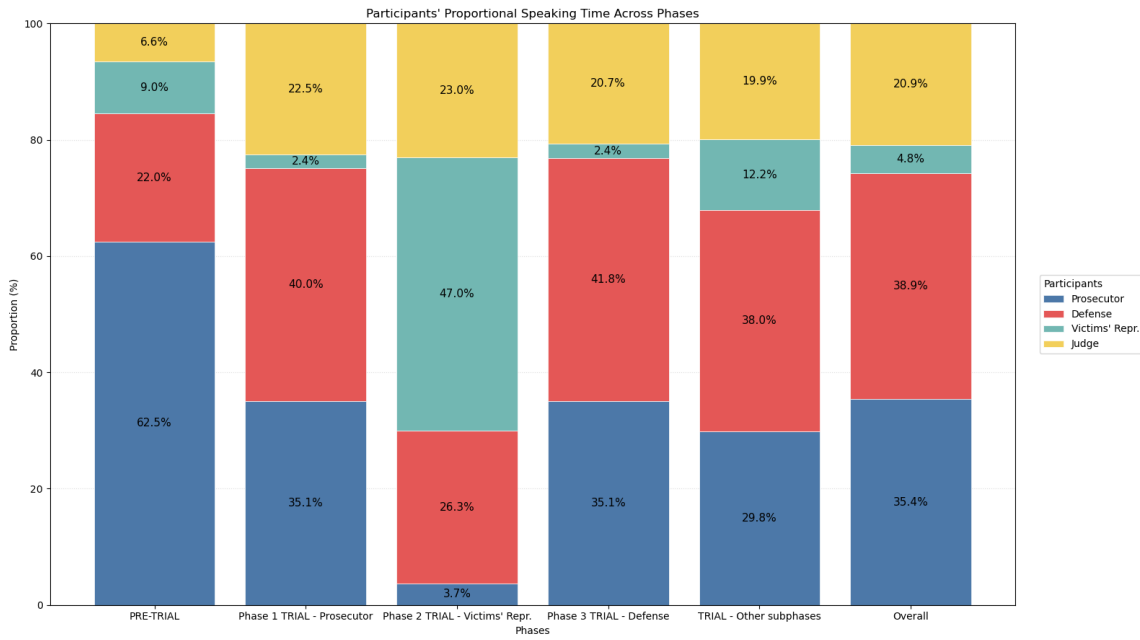


Figure 3: Distribution of words by participant for each phase (*Yekatom & Ngaiissona*)

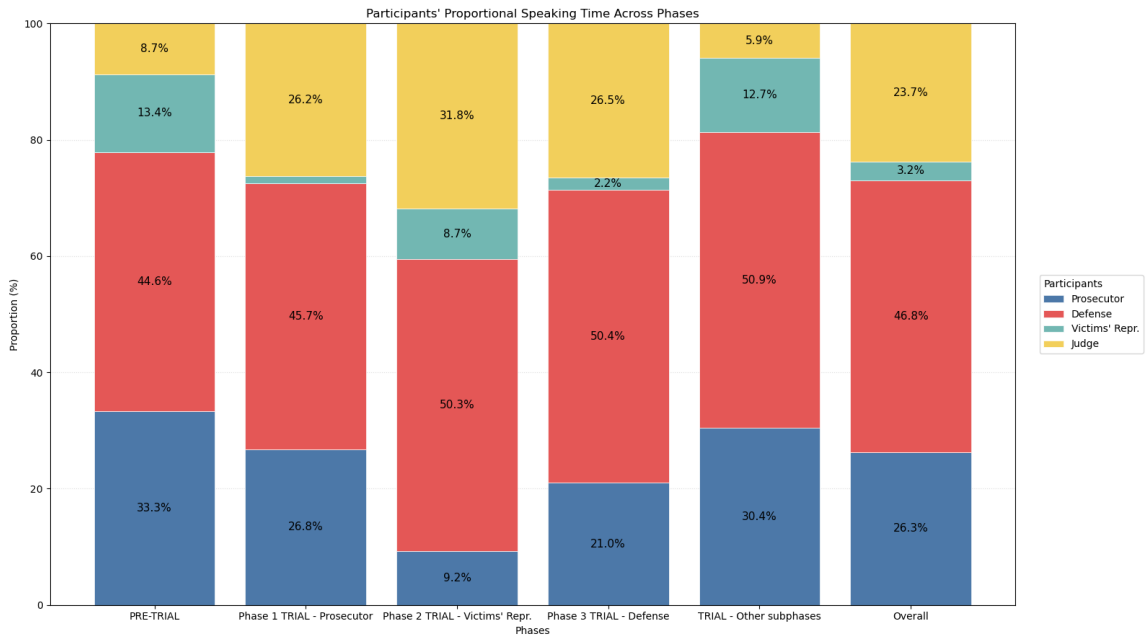
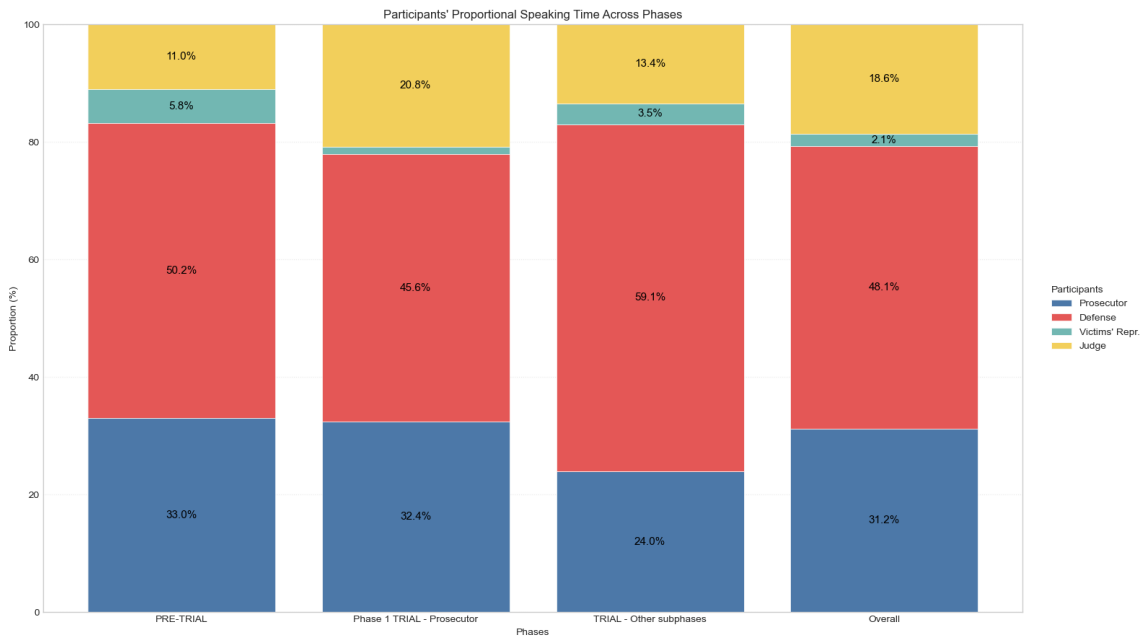
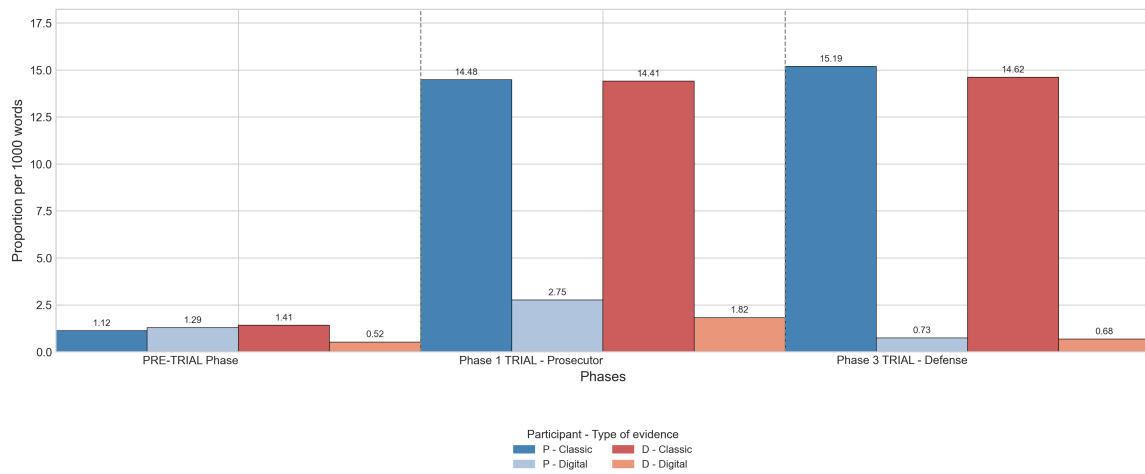


Figure 4: Distribution of words by participant for each phase (*Gbagbo & Blé Goudé*)



In what follows, we focus specifically on the strategies of the OTP and the defense through references to classical vs. digital evidence. Figures 5, 6, and 7 show the proportion of references (per 1000 words) to digital and classical evidence made by these parties during the pre-trial and trial phases (focusing in particular on the OTP and defense hearings when available). These proportions are computed by dividing the number of references to digital or classical evidence made by a given party in a given phase by the total number of words spoken by that party in the same phase, and then multiplying the result by 1000.<sup>39</sup> The figures indicate that, at both the pre-trial stage and trial subphases, the proportion of references to digital evidence made by a party is generally lower than that of references to classical evidence (although this result may reflect the composition of our keyword lists).

Figure 5: References (per 1000 words) to digital and classical evidence by participant (*Al Hassan*)



<sup>39</sup>For instance, the proportion of references to classical evidence made by the prosecution in the pre-trial phase is calculated as:

$$\frac{\text{number of references to classical evidence made by the prosecution in pre-trial}}{\text{number of words spoken by the prosecution in pre-trial}} \times 1000.$$

Figure 6: References (per 1000 words) to digital and classical evidence by participant (*Yekatom & Ngaiïssona*)

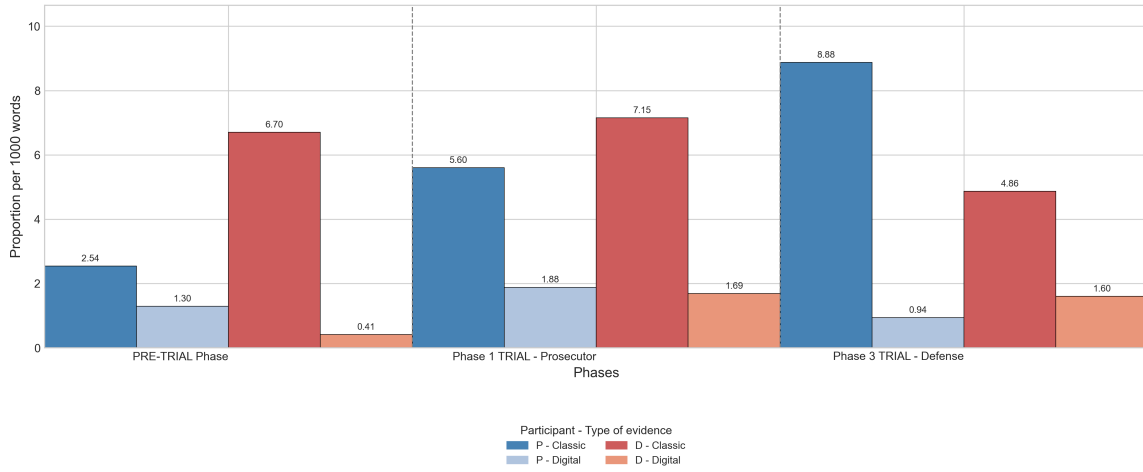
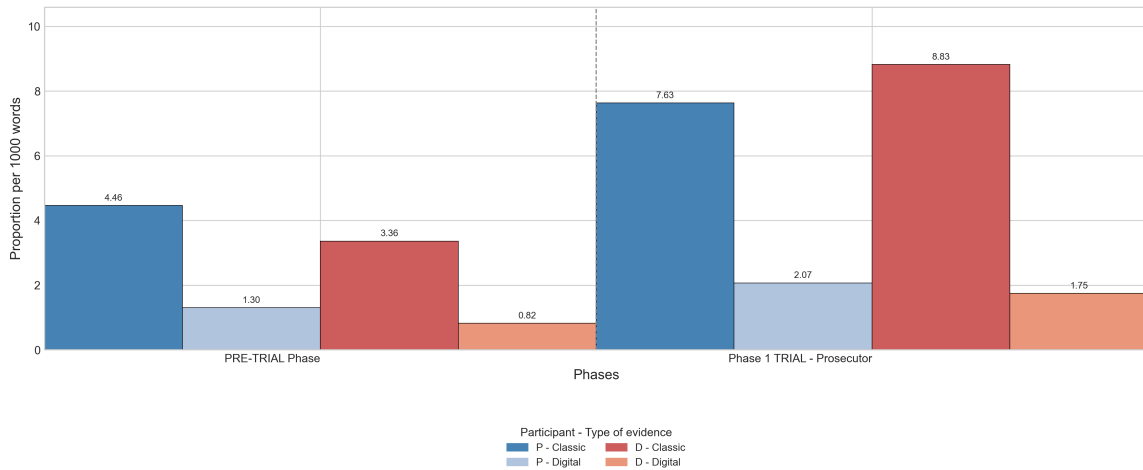


Figure 7: References (per 1000 words) to digital and classical evidence by participant (*Gbagbo & Blé Goudé*)



Next, to compare the intensity of references to digital evidence (versus classical evidence) across cases, we compute, for both the OTP and the defense, the ratio of references to digital evidence to references to classical evidence during their respective trial hearings.<sup>40</sup> In the *Al Hassan* case, these ratios suggest that the OTP employs digital evidence relatively more intensively than the defense. By

<sup>40</sup>During Phase 1 (trial hearings of the OTP) in the *Al Hassan* case, the ratio of the prosecutor's references to digital evidence relative to classical evidence is approximately 0.1901. By comparison, the same ratio for Phase 3 (trial hearings of the defense) in the *Al Hassan* case, calculated for the defense, is 0.0467. In the *Yekatom & Ngaiïssona* case, the corresponding ratios for the OTP and for the defense are approximately 0.3359 and 0.3299, respectively.

comparison, in the *Yekatom & Ngaïssona* case, the intensities of references to digital evidence relative to classical evidence become nearly identical for the OTP and the defense. Moreover, compared with the values obtained in the *Al Hassan* case, these ratios are considerably higher, suggesting that both parties rely much more heavily on digital evidence (relative to classical evidence) in the *Yekatom & Ngaïssona* case.<sup>41</sup>

In the *Al Hassan* case, the difference in the ratios of references to digital evidence relative to classical evidence suggests that the OTP’s strategy is distinct from that of the defense. This may be linked to the respective roles and levels of experience of the two parties. First, because the Prosecution bears the burden of proof and is the first party to present evidence at trial, demonstrating the defendant’s guilt *beyond a reasonable doubt* may require collecting incriminating material from a wide range of sources, including social media. As noted by Hellwig (2021), “as the Prosecutor holds the burden of proof, the Prosecution must prove all elements of the crime [...]. Crimes prosecuted under [international criminal law (ICL)] present some peculiarities with evidentiary implications, as the prosecution must provide context-related evidence, crime-based evidence and linkage evidence.” By contrast, the exculpatory evidence relied upon by the defense may differ in nature from that used by the OTP and may therefore be less readily available on digital platforms. Indeed, according to de Arcos Tejerizo (2023), “the OTP counts on a wide range of resources to collect evidence on the ground, while the defence may often have access to privileged information from the accused, which allows them to focus their search for exonerating circumstances.” Second, the OTP is often a repeat player in ICC trials, whereas defense teams participate less frequently. As a result, prosecutors may be more accustomed to working with digital evidence in the courtroom. Third, the OTP relies on a wide range of professionals, including investigators, analysts, and specialists in public information and communication.<sup>42</sup> By contrast, “when approaching potential sources of evidence, defence teams may face considerable budgetary and resource constraints” (de Arcos Tejerizo, 2023). Consequently, we may expect that exploiting digital evidence is less demanding for the OTP than for the defense.

In the *Yekatom & Ngaïssona* case, the OTP’s and defense’s strategies regarding the use of digital *versus* classical evidence appear to be much more similar. These observations may be related to the later period in which the crimes were committed in the *Yekatom & Ngaïssona* case, which made digital evidence more widely accessible to both parties. Indeed, the transcripts clearly show a higher number of references to digital evidence, relative to classical evidence, in the context of the *Yekatom & Ngaïssona* case. The nature of the digital evidence relied upon by the parties also differs across cases. For example, in the *Yekatom & Ngaïssona* case, the defense makes extensive reference to the social media platform Facebook (459 references, compared with 86 by the OTP), whereas such references are almost absent in the *Al Hassan* case (49 in total) and in the *Gbagbo & Blé Goudé* case (28 in total). Overall, in both *Al Hassan* and *Gbagbo & Blé Goudé*, references to digital evidence relate predominantly to videos.<sup>43</sup>

Finally, the results of the negative binomial regressions for the OTP and the defense, reported in

---

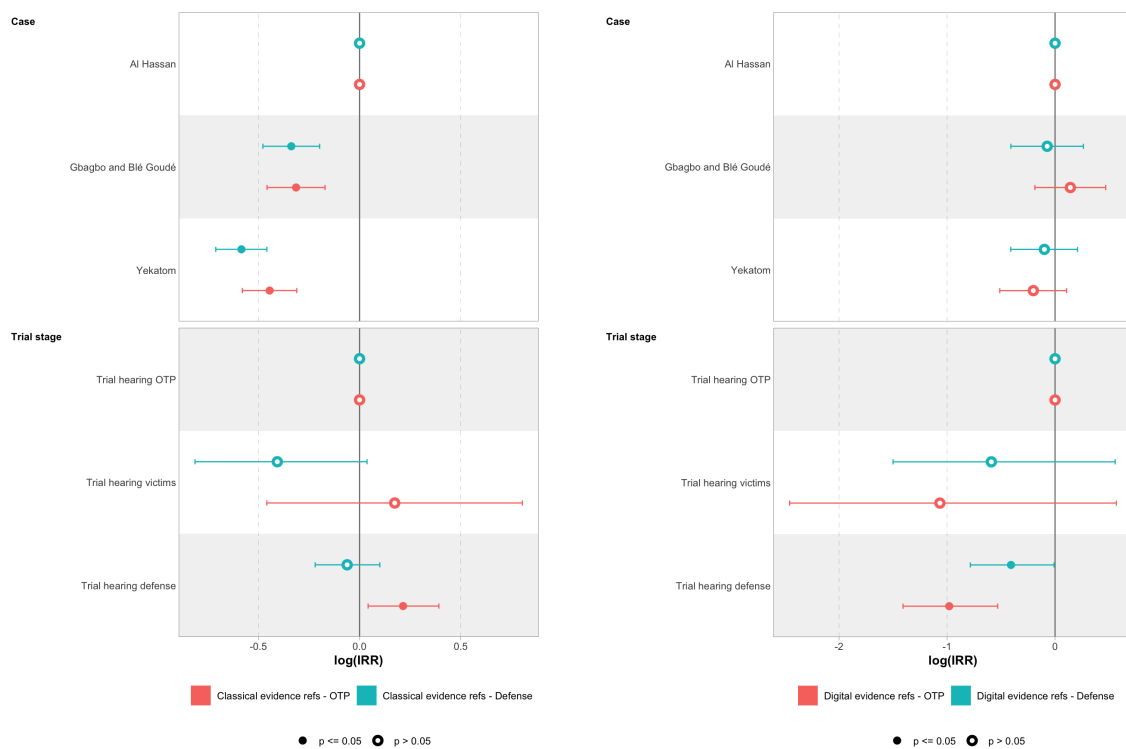
<sup>41</sup>In the *Gbagbo & Blé Goudé* case, the ratio of the Prosecutor’s references to digital evidence relative to classical evidence is approximately 0.2828, making it comparable to (but lower than) that observed in the *Yekatom & Ngaïssona* case.

<sup>42</sup>“The OTP benefits from the services of approximately 380 dedicated staff members from over 80 different nationalities, including members of the legal profession, investigators and analysts, psycho-social experts, individuals with experience in diplomacy and international relations, public information and communication, and more” (see <https://www.icc-cpi.int/about/otp>).

<sup>43</sup>In the *Yekatom & Ngaïssona* case, 74.34% of references to digital keywords correspond to the terms “video,” “video footage,” “footage,” or “open source video,” compared with 85.42% in the *Al Hassan* case and 87.36% in the *Gbagbo & Blé Goudé* case.

figure 8 (with detailed estimates in Tables 14 and 15 in Appendix E), are consistent with those obtained in figures 5, 6 and 7. In particular, both *Gbagbo & Blé Goudé* and *Yekatom & Ngaiïssona* trial are associated with statistically fewer references to classical evidence by both the prosecution and the defense compared to the *Al Hassan* case.<sup>44</sup> As a reminder, when comparing the ratios of references to digital evidence (*versus* classical evidence) across cases, we showed that this ratio is higher in the *Gbagbo & Blé Goudé* and *Yekatom & Ngaiïssona* cases than in the *Al Hassan* case. Figure 8 indicates that this result is not driven by a greater number of references to digital evidence in *Gbagbo & Blé Goudé* and *Yekatom & Ngaiïssona*, but rather by a smaller number of references to classical evidence, when compared to *Al Hassan*. With respect to the subphases of the trial,<sup>45</sup> we observe that the final subphase (defense hearings) is associated with significantly fewer references to digital evidence than the first subphase (OTP hearings) for both parties.

Figure 8: References to classical and digital evidence by the OTP and the defense



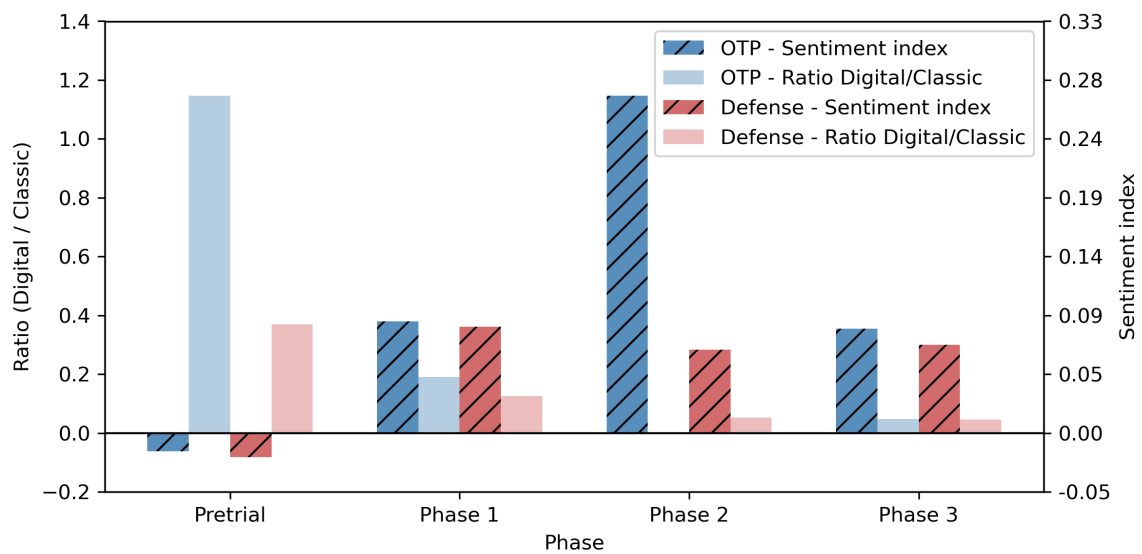
<sup>44</sup>As a reminder, unlike the other two cases, the *Gbagbo & Blé Goudé* proceedings resulted in an acquittal, as the evidence presented by the OTP was deemed insufficient. However, we do not identify any particular feature of the *Gbagbo & Blé Goudé* case that would explain this outcome when compared with the other cases. Indeed, according to figure 8 (left panel), the number of references to classical evidence by the OTP is significantly lower in *Gbagbo & Blé Goudé* than in *Al Hassan*, with an even larger gap when compared with the *Yekatom & Ngaiïssona* case. Furthermore, as shown in figure 8 (right panel), the model does not capture any significant difference in the number of references to digital evidence by the OTP in the *Gbagbo & Blé Goudé* case relative to the *Al Hassan* case.

<sup>45</sup>It should be noted that the variance captured by the coefficients associated with the “trial stage” variable comes solely from the *Al Hassan* and *Yekatom & Ngaiïssona* cases. Indeed, there is no variation for this variable in the *Gbagbo & Blé Goudé* case, as the trial phase ended at the end of the prosecutor’s sub-phase.

As a final exploratory step, we use sentiment analysis to detect the presence of emotions in exchanges between the parties and to assess whether emotional intensity differs when referring to digital *versus* classical evidence. Sentiment analysis, also known as opinion mining, is a subfield of NLP increasingly applied to legal texts to extract measures of tone (Ash et al., 2022) and subjectivity (Cao et al., 2020). We rely on TextBlob as our sentiment analysis algorithm.<sup>46</sup> This algorithm allows us to compute a polarity index for each transcript and each participant in the trial. The polarity index (hereafter sentiment index) ranges from  $-1$  to  $1$ , with higher values indicating more positive sentiment.

The sentiment indexes and the ratios of references to digital *versus* classical evidence for both the OTP and the Defense, across the pre-trial and main trial subphases, are presented separately for each case in figures 9, 10, and 11.<sup>47</sup>

Figure 9: Sentiment and ratio digital/classical for OTP and defense (*Al Hassan*)



<sup>46</sup>TextBlob is a widely used Python library that provides a simple interface for common NLP tasks (see <https://textblob.readthedocs.io/en/dev/index.html>). Because TextBlob is a general-purpose library and may lack specificity for the context we examine, we also tested LegalBERT (<https://opensource.legal/projects/LegalBERT>) as a more specialized alternative, given that LegalBERT is trained on legal texts. However, after experimenting with both tools, we decided to retain TextBlob, as LegalBERT’s specialization in formal legal language appears less suited to the conversational and interactive style characteristic of ICC hearings. This is a limitation of our approach, and ideally one would train a model specifically tailored to our purposes. We leave this task for future research, as it falls beyond the scope of the present paper.

<sup>47</sup>All words spoken by each party within a given subphase were aggregated prior to computing the sentiment index.

Figure 10: Sentiment and ratio digital/classical for OTP and defense (*Yekatom & Ngaiissona*)

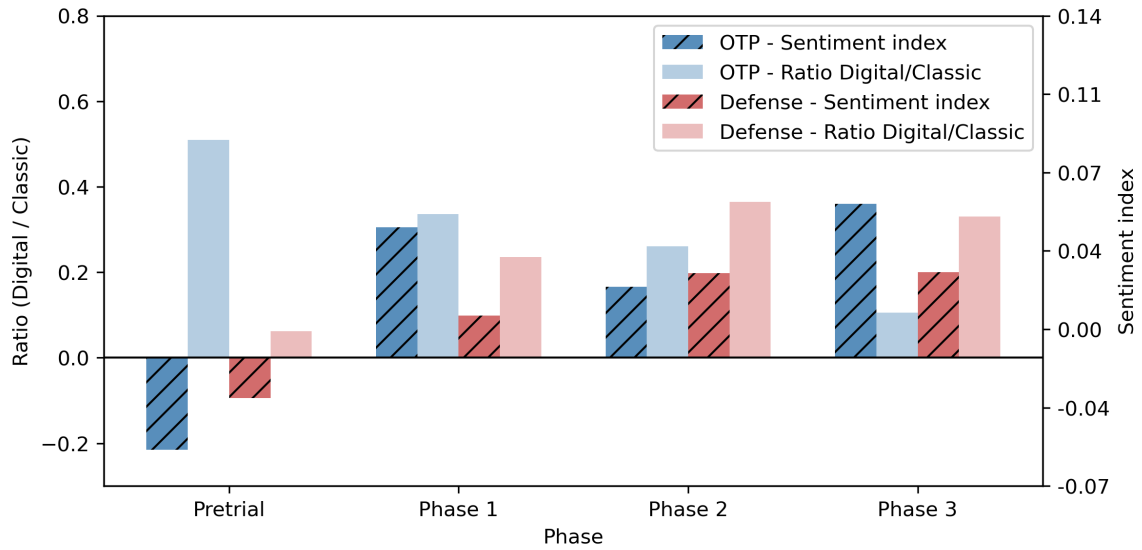
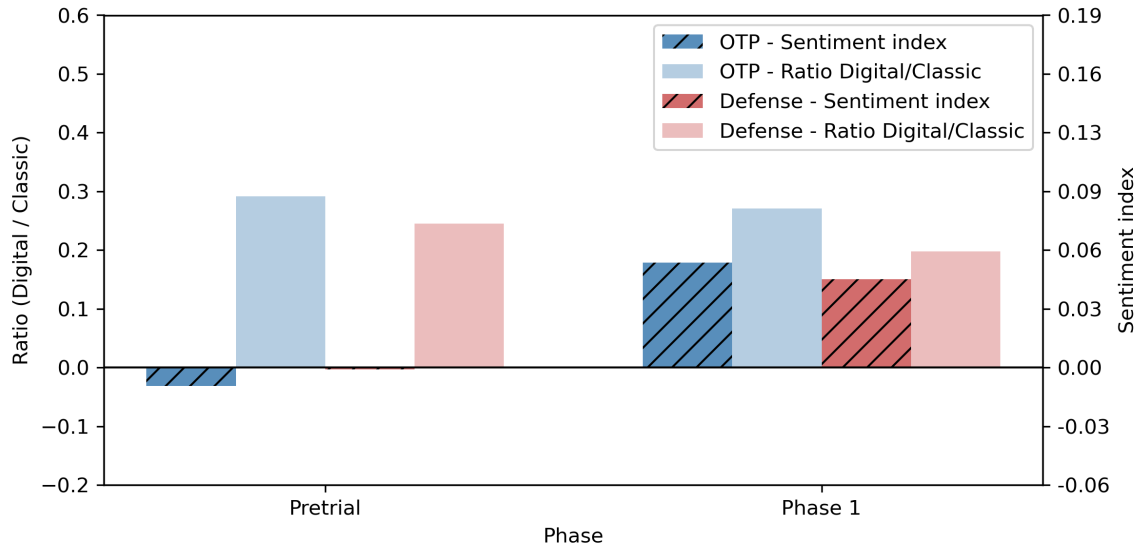


Figure 11: Sentiment and ratio digital/classical for OTP and defense (*Gbagbo & Blé Goudé*)



Figures 9, 10 and 11 show that the sentiment indexes of both the OTP and the defense are close to zero during the pre-trial phase, which is consistent with the limited role of oral argumentation at the confirmation of charges subphase, in contrast with the trial hearings (during which evidence are presented in an adversarial setting). Moreover, these figures suggest a positive correlation between the sentiment indexes of the OTP and those of the defense (with the OTP generally displaying slightly higher sentiment values across most subphases). One possible interpretation of this positive

correlation is that each party adjusts its tone in response to the other (*i.e.*, the harsher one party’s intervention, the harsher the opposing party’s reaction).<sup>48</sup> However, we do not observe any clear correlation between, on the one hand, the sentiment indexes of the OTP and the defense and, on the other hand, the ratio of references to digital *versus* classical evidence. A possible explanation is that the sentiments expressed by both parties are more closely tied to the topic and tone of the exchange (and perhaps the substance of the evidence presented) than to the type of evidence (digital or classical).

## 5 Conclusion

In this paper, we examine the references to digital *versus* classical evidence at the International Criminal Court (ICC). We define digital evidence as information transmitted over the Internet, whereas classical evidence refers primarily to witness testimony. Despite its usefulness, particularly when access to the territory is impossible, the probative value of digital evidence is often questioned or limited due to the difficulty of formally identifying and authenticating its source. There is also a substantial risk that such information may be inaccurate, for instance because of disinformation or misinformation. Moreover, the use of digital evidence raises practical challenges at the investigation stage, including the storage, verification, and analysis of large volumes of video material.

Our objective is to explore the extent to which digital evidence is used in court, and how the prosecution refers to it alongside classical evidence to meet the different standards of proof at the pre-trial and trial stages (Section 3). We also investigate whether the Prosecution and the Defense refer to these two types of evidence differently throughout the trial (Section 4). To pursue these objectives, we constructed a database containing the pre-trial and trial hearing transcripts of three main cases: *Gbagbo & Blé Goudé*, *Al Hassan* and *Yekatom & Ngaïssona*. These transcripts record the statements made by the various participants during the hearings (including judges, registry staff, members of the Office of the Prosecutor, defense counsel, legal representatives of victims, and witnesses).

In Section 3, we develop a simple theoretical model of evidence collection by the Office of the Prosecutor (OTP) and, using a list of predefined keywords (see Tables 8 and 9), we statistically examine references to evidence in the hearing transcripts. Within the theoretical framework, we show that as the standard of proof becomes more stringent, the quantities of both digital and classical evidence gathered by the prosecutor generally increase. The relative increase in each type of evidence depends on the relative marginal costs of collecting each type of evidence and on the degree of complementarity between them. Our empirical analysis indicates that references to both types of evidence increase between the pre-trial and trial phases, the latter being characterized by a higher standard of proof.<sup>49</sup> Estimates from a negative binomial regression further show that the increase in references between the two phases is relatively larger for classical evidence.

In Section 4, we examine the use and presentation of digital *versus* classical evidence by party, focusing specifically on the OTP and the defense. We show that substantial differences exist in the

---

<sup>48</sup>Given the small sample size and the exploratory nature of this analysis, these results should be interpreted with caution.

<sup>49</sup>Other differences exist between the pre-trial and trial stages. As reflected in the number of transcripts, the pre-trial phase is considerably shorter than the trial phase, and the importance of oral presentation is significantly more limited.

extent to which each party relies on the two types of evidence. These differences can be explained by the respective roles that each party plays during the trial. The results suggest that, in the *Al Hassan* case, the OTP uses digital evidence relatively more intensively than the defense, whereas this difference nearly disappears in the *Yekatom & Ngaiïssona* and *Gbagbo & Blé Goudé* cases. It is also noteworthy that the *Yekatom & Ngaiïssona* case features numerous references by the defense to the social media platform *Facebook*. We further observe that, for both parties, the final subphase of the trial (defense hearings) is associated with fewer references to digital evidence than the first subphase (OTP hearings). Finally, we use natural language processing, and specifically sentiment analysis, to measure the emotional tone of the OTP and the defense in these cases. Our results suggest a positive correlation between the sentiment indexes of the two parties. However, we do not find any evidence of a correlation between the sentiment index and the relative use of digital evidence across subphases and parties.

There are several limitations to our work. A first limitation concerns the timing of the crimes in relation to the evolution of digital social platforms and digital communications. The hearing transcripts examined in this paper relate to crimes that occurred more than a decade ago. For instance, the *Al Hassan* case involves crimes committed in 2012. Since then, the use of digital platforms has become far more widespread. The recent Russian invasion of Ukraine, for example, has generated extensive digital recordings documenting Russian abuses, which have undoubtedly played a role in triggering ICC investigations. However, any eventual trials relating to the situation in Ukraine are unlikely to take place in the near future. A second limitation relates to our reliance on predefined lists of keywords to conduct the textual analysis. An alternative approach would be to manually identify references to classical and digital evidence within a subsample of transcripts and then train a classifier on that sample. To fully exploit such an approach, however, one would ideally need access to a large database of relatively homogeneous cases. By contrast, our dataset includes only three cases, each of which is very specific, as we chose to focus exclusively on the most recent ICC cases that have completed the full trial stage. This small number of heterogeneous cases constitutes a third limitation, insofar as it allows only for limited external validity of our empirical results.

# Appendix

## Appendix A: Additional tables for section 2

Table 7: Proportion of pages redacted by phase and by case

	Pre-trial (%)	Trial (%)
Gbagbo & Blé Goudé	0	10.73
Al Hassan	6.82	17.32
Yekatom & Ngaissona	4.06	7.45

Table 8: Proportion per 1000 words of keywords related to classical evidence for each case

Keyword	Al Hassan	Gbagbo & Blé Goudé	Yekatom & Ngaissona
Testimonial	0.0038	0.0065	0.0071
Testimony	0.5486	0.4388	0.4540
Witness	8.3238	4.8013	4.7993
Overall	8.8762	5.2466	5.2605

*Note:* The plurals of each relevant word have been taken into account.

Table 9: Proportion per 1000 words of keywords related to digital evidence for each case

Keyword	Al Hassan	Gbagbo & Blé Goudé	Yekatom & Ngaissona
Facebook	0.0207	0.0068	0.1640
Instagram	0.0000	0.0000	0.0000
Linkedin	0.0000	0.0000	0.0000
Snapchat	0.0000	0.0000	0.0000
Telegram	0.0000	0.0002	0.0002
Tiktok	0.0000	0.0000	0.0000
Twitter	0.0097	0.0000	0.0004
Whatsapp	0.0017	0.0000	0.0071
Flickr	0.0000	0.0000	0.0000
Youtube	0.0013	0.0199	0.0013
Viadeo	0.0000	0.0000	0.0000
Digital	0.0114	0.0065	0.0047
Video	0.8976	0.7604	0.6258
Footage	0.0882	0.2235	0.1087
Internet	0.0190	0.0186	0.0078
Hach	0.0177	0.0000	0.0002
Broadcast	0.0793	0.0940	0.0629
Social Media	0.0080	0.0046	0.0129
Social Network	0.0042	0.0007	0.0042
Online Broadcast	0.0000	0.0000	0.0000
Video Footage	0.0325	0.0654	0.0367
Digital Evidence	0.0008	0.0005	0.0004
Open Data	0.0000	0.0000	0.0000
Open Source Video	0.0000	0.0000	0.0000
Overall	1.1921	1.2012	1.0374

*Note:* The plurals of each relevant word have been taken into account.

Table 10: Reference to classical evidence in hearing transcripts

	n	%
No reference	12	1,6
Reference	744	98,4
Total	756	100

Table 11: Reference to digital evidence in hearing transcripts

	n	%
No reference	144	19
Reference	612	81
Total	756	100

## Appendix B: The optimal levels of classical evidence.

The Lagrange function resulting from (1) is:

$$L(C, D) = \alpha C + \beta D + \lambda(f(C, D) - \bar{x}) \quad (4)$$

From the first-order conditions, we have:

$$\begin{aligned} \frac{\partial L}{\partial C}(C, D) &= \alpha + \lambda \frac{\partial f}{\partial C}(C, D) = 0 \\ \frac{\partial L}{\partial D}(C, D) &= \beta + \lambda \frac{\partial f}{\partial D}(C, D) = 0 \end{aligned} \quad (5)$$

The constraint in (1) is binding ( $\lambda > 0$ ), because otherwise the OTP can always decrease  $\alpha C + \beta D$  by decreasing either  $C$  or  $D$  without violating that constraint. Thus, choosing  $C$  and  $D$  such that  $f(C, D) > \bar{x}$  cannot be optimal.

From (5) and the binding constraint, the optimal levels of classical evidence ( $C^*$ ) and digital evidence ( $D^*$ ) are characterized by:

$$\begin{cases} \frac{\partial f}{\partial C}(C^*, D^*) = \frac{\alpha}{\beta} \\ f(C^*, D^*) = \bar{x} \end{cases} \quad (6)$$

### Effect of an increase in the standard of proof on $C^*$ and $D^*$ .

By differentiating the first equation of (2) with respect to  $\bar{x}$ , we get:

$$\frac{\partial C^*}{\partial \bar{x}} \left[ \underbrace{\beta \frac{\partial^2 f}{\partial C^2}(C^*, D^*) - \alpha \frac{\partial^2 f}{\partial C \partial D}(C^*, D^*)}_{\equiv X} \right] + \frac{\partial D^*}{\partial \bar{x}} \left[ \underbrace{\beta \frac{\partial^2 f}{\partial C \partial D}(C^*, D^*) - \alpha \frac{\partial^2 f}{\partial D^2}(C^*, D^*)}_{\equiv Y} \right] = 0 \quad (7)$$

Similarly, by differentiating the second equation of (2), we get:

$$\frac{\partial C^*}{\partial \bar{x}} \frac{\partial f}{\partial C}(C^*, D^*) + \frac{\partial D^*}{\partial \bar{x}} \frac{\partial f}{\partial D}(C^*, D^*) = 1 \quad (8)$$

We can write the system of equations formed by (7) and (8) in matrix form:

$$\begin{pmatrix} X & Y \\ \frac{\partial f}{\partial C}(C^*, D^*) & \frac{\partial f}{\partial D}(C^*, D^*) \end{pmatrix} \begin{pmatrix} \frac{\partial C^*}{\partial \bar{x}} \\ \frac{\partial D^*}{\partial \bar{x}} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (9)$$

Under assumption (ii), we can show that the determinant of the first matrix, denoted  $H$  in the following, is negative (it is a necessary condition of the minimization program (1)).

Applying Cramer's rule to (9) yields:

$$\frac{\partial C^*}{\partial \bar{x}} = \frac{\begin{vmatrix} 0 & Y \\ 1 & \frac{\partial f}{\partial D}(C^*, D^*) \end{vmatrix}}{H} = -\frac{Y}{H} \quad (10)$$

From which it follows that:

$$\text{sign} \left( \frac{\partial C^*}{\partial \bar{x}} \right) = \text{sign} (Y) > 0 \quad (11)$$

And:

$$\frac{\partial D^*}{\partial \bar{x}} = \frac{\begin{vmatrix} X & 0 \\ \frac{\partial f}{\partial C}(C^*, D^*) & 1 \end{vmatrix}}{H} = \frac{X}{H} \quad (12)$$

From which it follows that:

$$\text{sign} \left( \frac{\partial D^*}{\partial \bar{x}} \right) = \text{sign} (-X) > 0 \quad (13)$$

**Condition for which the quantity of classical evidence increases relatively more with the standard.**

We compare (10) and (12). For a small increase in the standard, the quantity of classical evidence increases more than the quantity of digital evidence if:

$$\begin{aligned} \frac{\partial C^*}{\partial \bar{x}} > \frac{\partial D^*}{\partial \bar{x}} &\Leftrightarrow X + Y > 0 \\ &\Leftrightarrow \beta \frac{\partial^2 f}{\partial C^2}(C^*, D^*) - \alpha \frac{\partial^2 f}{\partial D^2}(C^*, D^*) + (\beta - \alpha) \frac{\partial^2 f}{\partial C \partial D}(C^*, D^*) > 0 \end{aligned} \quad (14)$$

## Appendix C: Additional tables for section 3

Table 12: References (per 1000 words) to classical and digital evidence by (phase  $\times$  case)

	Classical evidence	Digital evidence
(Pre-trial) $\times$ (Gbagbo & Blé Goudé)	2.67	0.64
(Pre-trial) $\times$ (Al Hassan)	0.92	0.85
(Pre-trial) $\times$ (Yekatom & Ngaïssona)	3.04	0.70
(Trial) $\times$ (Gbagbo & Blé Goudé)	6.16	1.29
(Trial) $\times$ (Al Hassan)	9.73	1.28
(Trial) $\times$ (Yekatom & Ngaïssona)	6.23	1.04
Overall	6.96	1.16

Table 13: References (count) to classical and digital evidence

	<i>Dependent variable:</i>	
	Classical evidence	Digital evidence
	(1)	(2)
Trial (ref. Pre-trial)	0.879*** (0.109)	0.527** (0.234)
Gbagbo and Blé Goudé (ref. Al Hassan)	-0.452*** (0.058)	-0.001 (0.126)
Yekatom and Ngaïssona (ref. Al Hassan)	-0.433*** (0.055)	-0.205* (0.120)
Constant	-5.534*** (0.114)	-7.183*** (0.245)
Observations	756	756
Log Likelihood	-3,911.449	-2,723.004
$\theta$	2.823*** (0.151)	0.601*** (0.033)
Akaike Inf. Crit.	7,830.898	5,454.008

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## Appendix D: Methodology of the hearing transcript cleaning process

1. A cleaning process had been applied on each transcript to remove headers, footers, and page line numbers.
2. We performed some textual preprocessing on the transcripts to capture each speaker's words during the hearings.
3. We applied frequency analysis methods to tabulate the occurrences of our keywords. Note that to ensure comprehensive coverage of words related to both digital and classical evidence, we carefully selected specific terms and their synonyms for each category. To enhance this process, we leveraged the *Thesaurus API*<sup>50</sup>, which automates the search for synonyms. To ensure semantic comprehensiveness, we employed the Thesaurus API not merely as a synonym generator but as a systematic extension tool. For each core keyword (*e.g.*, “video”, “witness”), the API provided a list of related terms, which we then manually curated to retain only those relevant within the legal and evidentiary context. This hybrid approach of automated expansion and manual filtering enhanced the reliability of our word occurrence analysis by accounting for lexical variations, while minimizing noise from irrelevant synonyms. Note that a team of lawyers familiar with the ICC's procedures collaborated on the keyword selection process. Furthermore, certain phrases were excluded to avoid misclassification of words, such as “The witness testifies via video link” for the term “video” and “Mr. Witness” or “Ms. Witness” for the word “witness”. Note that our counting methodology potentially involves counting multiple references to the same piece of evidence. For example, if the same video or witness is referenced 10 times, we count the word 10 times. This is not a problem for our empirical analysis since i) the same phenomenon can occur for both classical and digital evidence; ii) we are interested in variations (between phases and between parties) in references to evidence, therefore if this phenomenon is fairly uniform across procedural stages and parties, this will not cause any bias in the measured variations.
4. We matched the word count for each speaker in both the pre-trial and trial transcripts to their respective roles in the case.

---

<sup>50</sup>From API Ninjas: <https://api-ninjas.com/api/thesaurus>

## Appendix E: Additional tables for section 4

Table 14: References to classical evidence by parties

	<i>Dependent variable:</i>			
	Count of classical evidence by party per transcript			
	OTP (1)	Defense (2)	Victims (3)	Judge (4)
Gbagbo and Blé Goudé (ref. Al Hassan)	-0.314*** (0.073)	-0.338*** (0.071)	0.283* (0.145)	-0.334*** (0.036)
Yekatom and Ngaïssona (ref. Al Hassan)	-0.445*** (0.068)	-0.584*** (0.064)	-0.542*** (0.119)	-0.168*** (0.032)
Victims stage (ref. OTP stage)	0.174 (0.296)	-0.407* (0.219)	0.359 (0.258)	-0.235** (0.100)
Defense stage (ref. OTP stage)	0.215** (0.088)	-0.061 (0.080)	-0.096 (0.154)	0.061 (0.040)
Constant	-4.246*** (0.053)	-4.192*** (0.052)	-4.318*** (0.091)	-4.089*** (0.026)
Observations	660	658	535	663
Log Likelihood	-2,403.870	-2,621.112	-662.779	-2,342.103
$\theta$	2.507*** (0.159)	2.679*** (0.167)	2.329*** (0.396)	14.136*** (1.262)
Akaike Inf. Crit.	4,817.740	5,252.224	1,335.558	4,694.207

*Note:*

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 15: References to digital evidence by parties

	<i>Dependent variable:</i>			
	Count of digital evidence by party per transcript			
	OTP (1)	Defense (2)	Victims (3)	Judge (4)
Gbagbo and Blé Goudé (ref. Al Hassan)	0.142 (0.170)	-0.073 (0.176)	35.578 (3,238,572.000)	0.820*** (0.184)
Yekatom and Ngaïssona (ref. Al Hassan)	-0.201 (0.161)	-0.099 (0.159)	33.291 (3,238,572.000)	0.629*** (0.171)
Victims stage (ref. OTP stage)	-1.066 (0.920)	-0.590 (0.544)	3.038* (1.610)	-1.774*** (0.671)
Defense stage (ref. OTP) (ref. OTP stage)	-0.980*** (0.226)	-0.408** (0.200)	1.680 (1.374)	-0.745*** (0.228)
Constant	-6.125*** (0.126)	-6.373*** (0.132)	-43.011 (3,238,572.000)	-7.672*** (0.146)
Observations	660	658	535	663
Log Likelihood	-1,405.320	-1,521.801	-43.023	-1,014.078
$\theta$	0.470*** (0.037)	0.452*** (0.034)	0.097* (0.055)	0.541*** (0.055)
Akaike Inf. Crit.	2,820.641	3,053.602	96.045	2,038.156

*Note:*

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

## References

- Aronson, J. D. (2018). The utility of user-generated content in human rights investigations. New Technologies for Human Rights Law and Practice, pages 129–48.
- Ash, E., Chen, D. L., and Galletta, S. (2022). Measuring judicial sentiment: Methods and application to us circuit courts. Economica, 89(354):362–376.
- Bull, J. and Watson, J. (2019). Statistical evidence and the problem of robust litigation. The RAND Journal of Economics, 50:974–1003.
- Cao, Y., Ash, E., and Chen, D. L. (2020). Automated fact-value distinction in court opinions. European Journal of Law and Economics, 50(3):451–467.
- Christmann, R. and Kirstein, R. (2023). You go first!: coordination problems and the burden of proof in inquisitorial prosecution. European Journal of Law and Economics, 56(2):403–422.
- de Arcos Tejerizo, M. (2023). Digital evidence and fair trial rights at the international criminal court. Leiden Journal of International Law, 36(3):749–769.
- Demougin, D. and Fluet, C. (2008). Rules of proof, courts, and incentives. The RAND Journal of Economics, 39(1):20–40.
- Emons, W. and Fluet, C. (2009). Accuracy versus falsification costs: The optimal amount of evidence under different procedures. The Journal of Law, Economics, & Organization, 25(1):134–156.
- Frankenreiter, J. and Livermore, M. A. (2020). Computational methods in legal analysis. Annual Review of Law and Social Science, 16(1):39–57.
- Freeman, L. (2018). Digital evidence and war crimes prosecutions: The impact of digital technologies on international criminal investigations and trials. Fordham International Law Journal, 41(2):283–336.
- Froeb, L. M. and Kobayashi, B. H. (2001). Evidence production in adversarial vs. inquisitorial regimes. Economics Letters, 70(2):267–272.
- Galanter, M. (1974). Why the haves come out ahead: Speculations on the limits of legal change. Law & Soc’y Rev., 9:95.
- Gilligan, M. J. (2006). Is enforcement necessary for effectiveness? a model of the international criminal regime. International Organization, 60(4):935–967.
- Hellwig, K. (2021). The potential and the challenges of digital evidence in international criminal proceedings. International Criminal Law Review, 22(5-6):965–988.
- Kaplow, L. (2012). Burden of proof. The Yale Law Journal, pages 738–859.
- Laux, J. (2018). A new type of evidence? cyberinvestigations, social media, and online open source video evidence at the icc. Archiv des völkerrechts, 56(3):324–360.
- Lecorps, Y. (2021). On optimal enforcement in international crime setting. European Journal of Law and Economics, 51(2):285–296.
- Lester, B., Persico, N., and Visschers, L. (2012). Information acquisition and the exclusion of evidence in trials. The Journal of Law, Economics, & Organization, 28(1):163–182.

- Mungan, M. C. (2011). A utilitarian justification for heightened standards of proof in criminal trials. Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft, pages 352–370.
- Obidzinski, M. and Oytana, Y. (2019). Identity errors and the standard of proof. International Review of Law and Economics, 57:73–80.
- Posner, R. A. (1999). An economic approach to the law of evidence. Stan. L. Rev., 51:1477.
- Rizzolli, M. (2016). Adjudication type i and type ii errors. Forthcoming in the Encyclopedia of Law and Economics.
- Rizzolli, M. and Saraceno, M. (2013). Better that ten guilty persons escape: punishment costs explain the standard of evidence. Public Choice, 155(3):395–411.
- Sutter, D. et al. (2006). The deterrent effects of the international criminal court. In Conferences on New Political Economy, volume 23, pages 9–24. Mohr Siebeck, Tübingen.
- Talley, E. L. (2013). Law, economics, and the burden(s) of proof. In Research Handbook on the Economics of Torts, pages 305–329. MA: Edward Elgar.