Doctorate of Business Administration (DBA)

in Management and Technology

Legal Finance Analytics:

a data-driven proposal of asset pricing litigation risk

applied to international investment arbitration.

Madrid, 17th May 2024

Author: Jaime Aparicio García

Director: Dr. José Portela González

To my father

"Un coeur de père est le chef-d'oeuvre de la nature"

(attributed to Abbé Antoine Prévost (1697-1763))

Acknowledgements

To Dr. José Portela González and Dr. Carlos Bellón Núñez-Mera, for their guidance.

To my colleagues of the DBA - from whom I have learned so much.

To Professor Juan Fernández-Armesto, Mr. Brian J. Crowe (rest in peace), and Mr. Donald B. Workman - the greatest responsibilities can be met with the highest of the ethical standards.

To my dear F. Pedro Gil Garbisu, in memoriam.

To my brother Dr. Jorge Aparicio - a relentless researcher and practitioner in the civil engineering domain, for his inspiration and unconditional support.

To my parents, Dionisio and Ruth - without them I would not have been able to make this journey.

To my wife Virginie and our beloved three children – my heart beats in their chests.

Abstract

Claimants and third-party funders face a challenge to value litigation risk. Fair valuation through future discounted cash flows is highly dependent on subjective parameters. Demonstrating the precision of management's application of the process to determine certain assumptions used in the measurement of the fair value of Capital Provision Assets (CPAs) has become increasingly critical for the growing Litigation Finance industry.

This thesis proposes a data-driven methodology, based on existing available historical data, for two models that estimate the distribution of probability of the expected return and the time to the obtain the outcome, conditioned to winning an international arbitration award. Understanding expected return of past positive cases facilitates claimants and investors evaluate whether pursuing arbitration could be viable from a budget planning perspective, allocating resources effectively and provide a financial, litigation risk and portfolio management tool. Visibility on time to obtain the outcome is critical for calculating the expected return, given the significant opportunity costs of dedicating resources to an arbitration process. Estimating the duration and potential outcome of an arbitration case helps in managing expectations and developing effective and informed case strategies.

We find there is a lack of a standard methodology for calculating the fair value of litigation risk, and this research could contribute to fill that gap, providing an additional benchmarking tool to enhance stakeholder communication at financial reporting vis-à-vis internal decision making, or external audits and financial regulators.

Through the analysis of a sample of awards retrieved from a public Investor-State Dispute Settlement (ISDS) dataset, we have used statistical metrics to calibrate the difference between what the claimant had requested and actually obtained. A correlation between the amount received, comprising principal, interest and, whenever awarded, costs, has been analysed vis-à-vis selected legal regressors based on formal and substantive law, namely alleged and found breaches of international treaties in awards that have been positive to the claimant.

Subsequently, we have employed a counterfactual scenario to determine the internal rate of return (IRR) of a Third-Party Funder (TPF) that would have financed 100% of the costs, and received 40% of the proceeds of the damages award.

The goal of the methodology exercise is two-fold: (i) to achieve an additional objective benchmark for fair valuation of existing litigation cases in a relatively illiquid market; and (ii) to propose a historical-data based forecast for new cases, estimating the potential compensation and time-to-award considering variables known to the plaintiff beforehand: claimed amount, projected costs, and those legal regressors that have been found as statistically significant in our research. The robustness of the model has diminished when attempting to predict the time-to-award duration vis-à-vis the prediction of the damages award.

Academic literature has pursued legal judgment prediction through several methods, mainly focused on predicting the outcome with accuracy. We have focused on the characterization of the quantum and the time to reach a positive outcome in international arbitration. Legal Analytics and Natural Language Processing has been enhanced by Artificial Intelligence (AI) techniques, allowing delivery of outputs in a very short period of time with increasingly acceptable results. This proposal of litigation risk asset pricing model in the investment arbitration field leverages on available data and a common framework of institutional arbitration rules and international treaties. The novelty of the model proposed is to embed past financial and legal information as objective data. The data-driven method is therefore agnostic of variables such as the type of claimant, the venue of the process, the tribunal composition, arbitration rules, or treaty applicable to the dispute. This proposal addresses the need to reduce the subjectivity of the assessment of litigation risk and provides an additional quantitative tool to the qualitative considerations that are still required.

A probability distribution function of internal rate of returns and a method to classify and rate legal cases can become an additional asset management tool within the litigation finance industry and lead to the categorisation of litigation risk as an ad hoc asset class in the alternative investment landscape.

INDEX

3. Proposed methodology – Legal Finance Analytics applied to Third Party Funding (TPF) in International Investment Disputes

A- Preliminary Risk-Reward considerations

Multiple vs Percentage Approach

Internal Rate of Return

The Hit Ratio

Calibrating Risks and Rewards

B- Proposed Methodology Applied to Investor-State International Arbitration Disputes

The Data Set

Modelling

Data extraction - Retrieval Augmented Generation (RAG) and Prompting

C- Selection of regression variables

Legal Features and Financial Variables

Ratios applicable to the Claimant

Ratios applicable to a Third-Party Funder

Proxies and Caveats

4. Results

5. Discussion

6. Conclusions and suggested next steps

7. References

Appendix 1 - Legal Judgment Prediction Literature Review.

Appendix 2 – Case Study Data

2.1 Total population (n=260)

2.2 Selected sample - positive outcome (n=94)

Appendix 3 – Histograms

Appendix 4 – Correlation graphics and matrix

Appendix 5 – OLS Models

5.1 – OLS regression for the Principal

5.2 – OLS Regression for the Years

List of tables

List of figures

AI - Artificial intelligence

API - Application Programming Interface

ATE – After The Event

AuM – Assets Under Management

BERT - Bidirectional Encoder Representations from Transformers

BIICL - British Irish Institute of International and Comparative Law

BIT – Bilateral International Treaty

B2B – Business to Business

CPAs - Capital Provision Assets

CIETAC - China International Economic and Trade Arbitration Commission

CoC – cash on Cash (also MOIC Multiple of Invested Capital)

DBA - Damages Based Agreement

DCF – Discounted Cash Flow

DL- Deep Learning

DNJ – Denial of International Justice

ECT - Energy Charter Treaty

FET - Fair and Equitable Treatment

FFSS - Financial Statements

FPS – Full Protection and Security

FTA – Free Trade Agreement

FV – Fair Value

GenAI - Generative Artificial Intelligence

GLUE - General Language

Understanding Evaluation

GPT - Generative Pre-trained Transformer

GRETL- Gnu Regression Econometrics and Time-series Library

IA – Investment Arbitration

IAC - International Arbitration Centre

ICA – International Court of Arbitration

ICC – International Chamber of Commerce

ICSID – International Center for Settlement of Investment Disputes

IIT – International Investment Treaty

IIA - International Investment Agreement

ILFA – International Legal Finance Association
IRR - Internal Rate of Return
ISDS – Investor State Dispute Settlement
IT – Information Technology
ITA – Investment Treaty Arbitration
LLM - Large Language model
LitFin - Litigation Finance
LFA - Litigation Funding Agreement
ML – Machine Learning
MOIC - Multiple of Invested capital or CoC (Cash on Cash)
MFN - Most-Favoured-Nation
NAFTA- North America Free Trade Agreement
NLP - Natural language processing
NPV – Net Present Value
OLS - Ordinary Least Squares
PDF – Portable Document Format
PCA – Permanent Court of Arbitration
Q&A - question and answering
RAD - Representative Actions Directive
RAG – Retrieval Augmented Generation
RLHF - Reinforcement learning from human feedback
RNNs - Recurrent neural networks
ROIC - Return On Invested Capital
SCC - Arbitration Institute of the Stockholm Chamber of Commerce
SIAC - Singapore International Arbitration Centre
SME – Small and Medium Enterprise
STD – Standard Deviation
TPF - Third-Party Funding or Third-Party Funder
TPLF - Third-Party Litigation Funding
UC – Umbrella Clause
UNCTAD - United Nations Conference on Trade and Development
UNCITRAL – United Nations Commission on International Trade Law
VC - Venture Capital
WAL - Weighted average life

1. Introduction

A- Motivation

Financing a case before a Court or a Tribunal might be the only way for an individual or a corporation to uphold their rights, resolve a dispute and make the perpetrator of a wrongdoing accountable; preserving the rule of law is a pillar of a modern society. Rule of law without effective and efficient justice is no justice.

When a third party approaches the dispute requires different inputs to evaluate what are the merits of the case, and what are the monetary consequences of winning or losing. There is a high degree of subjectivity in these assessments, normally reliant on the intuition and know-how of the legal advisor and damages´ expert, after a deep understanding of the facts and evidence.

Moved by the need to address the access to justice principle, we question whether new approaches or, more systematically, methodologies leveraged on the treatment of data, might contribute to measure, monitor and manage litigation risk, defined in this research as the probability of a plaintiff of succeeding in a claim.

Stakeholders, such as claimants, defendants, their legal advisors, litigation funders, insurance companies and policy-makers could address litigation finance not only as a means to access to justice, but also as a risk management tool to assess the litigation risk promptly, anticipating a positive or adverse outcome with a certain degree of certitude and, as a consequence, inter alia, reducing the number of non-meritorious claims brought to court, shortening the length of the claims, and reducing the costs for the parties, in benefit to the arbitration and judicial systems.

Case assessment methodology is essential to allow calculable predictions on litigation investments, distinguishing “litigation funders from gamblers” (Jonas von Goeler (2016)). In the end, “only claims that hold up on a quantitative and qualitative basis are funded” (Nieuwveld, L. B., & Sahani, V. S. (2016)).

New means of extraction of relevant data, Natural Language Processing (NLP), Large Language Models (LLMs) and the application of statistical methods can contribute to model, explain and predict said litigation risk.

The mix of litigation finance, legal research and legal analytics could contribute to a nascent Legal Finance Analytics discipline. The contribution of this research is a methodology that can be used for asset management purposes, reducing the degree of subjectivity in the assessment for fair valuation purposes can be useful for investment or hedging decision-making. The methodology could contribute as an additional tool to fair valuation for accounting and auditing purposes.

To the knowledge of this author, the novelty and contribution is the combined qualitative and quantitative analysis of legal - both formal and substantive – features and financial variables as regressors in addressing the financial outcome in an international arbitration dispute. The methodology could be extended, *mutatis mutandis,* to other litigation disciplines beyond arbitration in international Investor-State investment disputes. Rating legal cases and breaking down the concept of probability of success impacts investment decisions and, ultimately, allows the categorisation of litigation risk as an asset class.

This study introduces a portfolio approach methodology focused on the subsegment of international investment arbitration, utilizing historical data for risk benchmarking. While recognizing that our modelling cannot guarantee accuracy or predictability, the methodology offers a potential objective proxy. The aim of this research is to offer an additional tool for benchmarking Investor-State international arbitration cases, with the goal of minimizing subjectivity in the aforementioned fair value assessment.

B- Objectives

The general objective of this research is to propose a novel methodology for assessing litigation risk in international investment arbitration.

The specific objectives of this thesis are three-fold. First, to review the state-of-the-art of litigation finance, and in particular applied to international arbitration disputes, as well as the influence of Artificial Intelligence (AI) in the legal domain.

Second, to elaborate a statistical model that analyses quantitative and qualitative data from International Arbitration awards that have been successful to the claimant. The analysis of positive awards is driven by the availability of public information. By retrieving legal and financial information through AI-based search tool, ChatPDF, namely breaches of international treaties and figures related to costs and damages awards, we aim at calculating, for instance, the Multiple On Invested Capital (MOIC) of a claimant or what would have been the Internal Rate of Return (IRR) for a third-party investor that bears 100% of the costs and receives 40% of the proceeds.

Finally, a tool has been designed, based on past information in order to predict the amount and time to receive an award, based on features known before an arbitration commences, such as what treaty breaches are alleged, the amount claimed and the costs to face an international arbitration dispute.

We believe the methodology in this research, focused on international arbitration, could be transferrable to other types of disputes, as these are leveraged on observable input and data, thus reducing the level of subjectivity, and contributing to address litigation risk as an asset class.

Calculating fair value of litigation risk is complex, with a mix of quantitative and qualitative features. We have not found in the literature a standard methodology for valuating this risk.

This thesis proposes a data-driven methodology, based on existing available historical data, for two models that estimate the distribution of probability of the expected return and the time to the obtain the outcome, conditioned to winning an international arbitration award.

Understanding expected return of past positive cases facilitates claimants and investors evaluate whether pursuing arbitration could be viable from a budget planning perspective, allocating resources effectively and provide a financial, litigation risk and portfolio management tool. Expected returns can also provide leverage in settlement negotiations.

Visibility on time to obtain the outcome is critical for calculating the expected return, given the significant opportunity costs of dedicating resources to an arbitration process. Estimating the duration and potential outcome of an arbitration case helps in managing expectations and developing effective and informed case strategies.

This research could contribute to fill that gap, providing an additional benchmarking tool to enhance stakeholder communication at financial reporting vis-à-vis internal decision making, or external audits and financial regulators.

C- Thesis Structure

This research is divided in two blocks: the former comprises Chapter 2, which covers the state-of-the-art, and Chapter 3, which outlines the methodology of the study. The second block consists of Chapter 4, which presents the results, Chapter 5, which delves into the discussion, Chapter 6, covering conclusions and next steps, and Chapter 7, compiling the references.

More specifically, Chapter 2, Literature Review, outlines the different themes this research verses upon: Litigation Finance in Investment Arbitration (2.A), Third Party Funder (TPF) assessment criteria and methodologies (2.B) and a survey of Legal Tech and Legal Analytics in Legal Research (2.C), with a particular focus on legal judgment prediction.

Chapter 3, Proposed methodology – Legal Finance Analytics applied to Third Party Funding (TPF) in International Investment Disputes, refreshes some basic preliminary risk-reward concepts for an investor in litigation finance (3.A), such as Multiple on Invested Capital (MOIC) or the Internal Rate of Return (IRR). In 3.B we describe the proposed methodology applied to Investor-State International Arbitration Disputes: data extraction, the relevance of Retrieval Augmented Generation (RAG) and prompting, categorization and modelling. Regression variables are elaborated in 3.C, where legal features and financial variables are identified and selected. Additionally, we calculate relevant ratios from both the claimant's and the funder's perspectives.

The Results of the study, chapter 4, include the key statistics of the sample for both an investor and a funder, as well as a critical analysis and interpretation of the regressors and coefficients of the Ordinary Least Squares (OLS) regression model. Limitations are also highlighted here.

Chapter 5 intends to elaborate further on the sensitivities of the variables of the model to check and balance its potential predictive nature, and last chapter, number 6, intends to summarise the findings and suggest a few courses of action that extend beyond the boundaries of our current knowledge, which we have endeavoured to push.

2. Literature Review

A - Litigation Finance in Investment Arbitration

Investment Arbitration

Given the lack of a global universal investment-arbitration system, there are two main distinctions (Van Boom, W. H. (2011)):

(i) *Institutional* vs *ad hoc* arbitration: the former is overseen by international bodies such as the World Bank's International Centre for Settlement of Investment Disputes (ICSID), the ICC International Court of Arbitration (ICA) and the Arbitration Institute of the Stockholm Chamber of Commerce (SCC). In the later, United Nations Commission on International Trade Law (UNCITRAL) arbitration rules apply.

(ii) More relevant is the difference between *commercial* arbitration, which could be initiated by either private party, and *investment* arbitration: an instrument exclusively at a private investor´s disposal against a State, whereby the ready-framework would be triggered by a referral clause to a Bilateral International Treaty (BIT), an International Investment Treaty (IIT) or to national legislation.

This thesis will focus on institutional international arbitration and investment arbitration where foreign investors assert rights against States. As a consequence, international commercial arbitration between private parties in international commercial transactions or inter-state disputes are not the object of this study. Additionally, the research will side with the interests of these claimants - private investors and their prospect funders - versus those of the States as respondents or defendants.

The international investment regime is built on a network of over 3,500 bilateral investment treaties and a few regional and plurilateral Free Trade Agreements (FTAs) (Langford, M., Behn, D., & Lie, R. (2020)). According to the United Nations Conference on Trade and Development (UNCTAD), and its World Investment Report published last July (United Nations Conference on Trade and Development. (2023)), in 2022, for the third consecutive year, the number of effective treaty

terminations (58) exceeded that of new (15) International Investment Agreements (IIAs), resulting in a IIA universe of 3,265, including 2,584 currently enforceable.

Figure 1. Stock of IIAs signed and in force, 1959–2022 (By date of signature). United Nations Conference on Trade and Development. (2023).

Additionally, in 2022, claimants filed 46 new Investor-State Dispute Settlement (ISDS) cases under IIAs, bringing the total count of publicly known cases to 1,257.

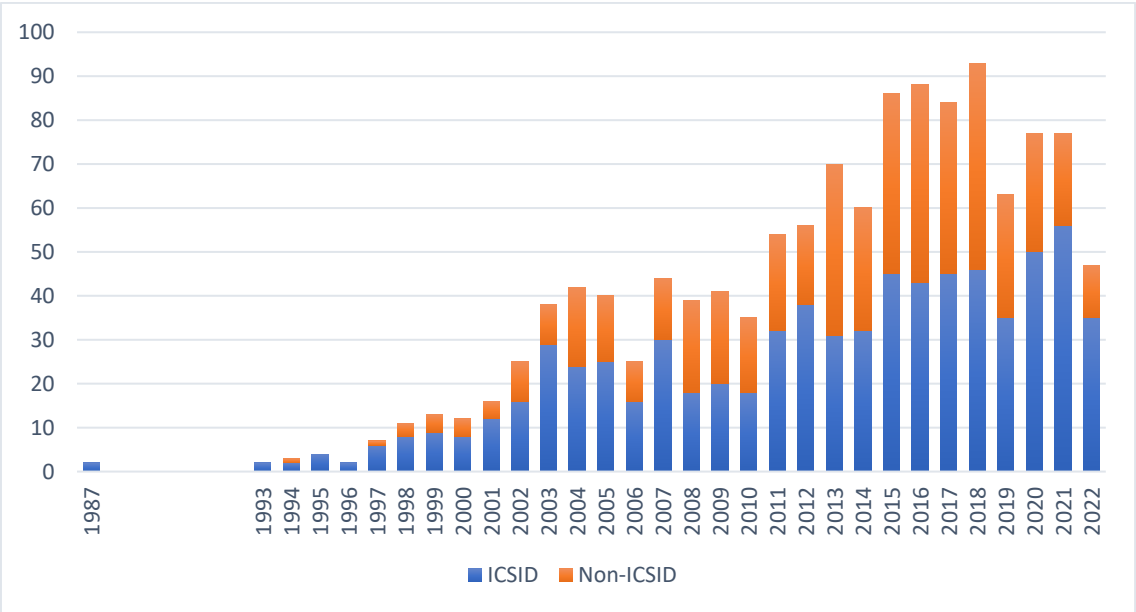


Figure 2. Treaty-based ISDS cases, 1987–2022. United Nations Conference on Trade and Development. (2023).

These 1,257 known ISDS cases, shown in Figure 2, have been updated and increased to 1,303 as of 31 July 2023 (United Nations Conference on Trade and Development. (n.d.)). It is important to highlight that, out of the 1,303 cases, there is an *undisclosed number of cases* that are being financed by third parties.

Investment disputes, not only treaty-based, according to Kluwer data base, as of April 2023, amount to 3,342. Treaty-based ISDS cases would therefore represent approximately 4 out of 10 of the total investment disputes.

<table>
<tr><td>2,608 concluded (78%)</td><td>924 concluded (71%)</td><td rowspan="3">260 in favour of claimant (28%)
344 in favour of respondent (37%)
173 settled (19%)
124 discontinued (13%)
23 neither party (liability no damages) (3%)</td></tr>
<tr><td>715 pending (21%)</td><td>357 pending (27%)</td></tr>
<tr><td>18 other (1%)</td><td>22 unknown (2%)</td></tr>
<tr><td>3,341 (100%)
Total Investment Disputes</td><td colspan="2">1,303 (100%)
924 concluded (100%)
International Investment Treaty-based</td></tr>
</table>

Table 1. International Investment Treaty-based ISD cases as of 31.7.2023 (center and right columns source: UNCTAD Navigator, accessed January 2024). Total Investment Disputes (left column, Arbitration Kluwer data base, April 2023)

This research has focused on the claims that have been granted in favour of the claimant, driven by the availability of information, as will be explained in the methodology section.

Relevant for the insight within the IIAs data relevant to this research, we would like to highlight the work of Hodgson, M., Kryvoi, Y., & Hrcka, D. (2021). Sponsored by the British Irish Institute of International and Comparative Law (BIICL) and Allen & Overy (A&O), Hodgson et al have undertaken a thorough analysis out of a sample of several hundred concluded ISDS cases, where key variables and statistics have been extracted namely, *inter alia*, winners/losers, costs involved, time to resolution, and damages award obtained. This document will be referred to in this research as "the A&O report (2021)", whose takeaways for the purpose of this research are summarised below**:**

(1) Outcome: Out of 433 cases subject of analysis, 44% were won by the claimant, 30% by the respondent State and 24% were lost in issues related to jurisdiction, whilst 2% were settled. This breakdown contradicts the prior global data, where cases are more frequently won by the States. Like Hodgson et al, we will also deep dive further on the subsegment of the cases won by the investor.

(2) Costs and access to funds: the median costs in an ISDS proceeding for investors is USD 3.8 million (M) - mean costs is USD 6.4 M. For respondent States, the median and mean costs are lower: USD 2.6 M and USD 4.7 M, respectively. Further granularity, out of 329 cases analysed (234 before June 2017 and 95 from June 2017 to May 2020), can be found in Table 2.

Amount claimed (AC) in USDM	AC < 50	250 >AC >50	1,000 >AC >250	AC>1,000
% of # of cases	34%	34%	32%	
Cost of claimant (median in USDM)	1.5	4.7	7.1	11.8
Cost for defendant (median in USDM)	1.2	2.7	4.2	6.2

Table 2. Cost analysis dependant on the amount in dispute. A&O report (2021).

As could be expected, the larger the amount under dispute, the larger the costs incurred.

(3) Length: Out of 400 investor vs State cases (ICSID and UNCITRAL) and 70 ICSID annulments, the mean length is 4.6 years (ICSID) and 4.2 years (UNCITRAL). The median length is similar: 3.8 years and 3.9 years, respectively. Generally, the higher the amount in dispute, the longer the proceedings (A&O report (2021), page 33). Table 3 summarizes the above statements:

Outcome	Costs (median)	Length
	Claimant | State	ICISD | UNCITRAL
44% won by claimant	*amount in dispute >USD 1 B*	Median
	USD 11.8 M | USD 6.2 M	3.8 years | 3.9 years
30% won by respondent		
	amount in dispute >USD 250 M	Mean
24% lost in jurisdiction	USD 7.1 M | USD 4.2 M	4.6 years | 4.2 years
2% settled	*100 M< amount in dispute >USD 250 M*	
	USD 4.7 M | USD 3.7 M	
	50 M< amount in dispute >USD 100 M	
	USD 3.7 M | USD 1.6 M	
	amount in dispute <USD 50 M	
	USD 1.5 M | USD 1.2 M	
Sample 433 cases	**Sample** 329 cases	**Sample** 400 cases and 70 annulments

Table 3. Outcome of Investor-State proceedings (concluded as of May 2020), Average cost by size of claims (234 before June 2017 and 95 to May 2020) and average duration. A&O report (2021), costs (pages 29-30) and length (page 33).

According to the A&O report (2021), most investment treaties and international investment agreements provide guidance on costs, being shaped by the rules of arbitration. Notably, despite the report not disclosing the number of cases that have been financed by third parties, the authors highlight that funding impact the "legitimacy and sustainability of the ISDS system". It also echoes UNCITRAL working group statement where *"some States have expressed concerns that costs and damages awarded in investor-State disputes have become excessive, causing difficulty for those with limited financial resources"* (United Nations Commission on International Trade Law. (n.d.-b)). On the other hand, the report also flags that investors with relatively modest claims could be precluded from pursuing legal action and their access to justice, undermined by the high costs and long duration of ISDS proceedings.

Litigation Finance and Market Size

A funder, who has no pre-existing interest in a litigation, acquires the rights or finances a dispute – of etymological Latin origin, *lis* - when it bears the costs of the petitioning party (or plaintiff) associated with initiating and/or maintaining a claim. This party, not originally involved in the dispute, is referred to as a Third-Party Funder, and the transaction as Third-Party Funding (we will refer to both as ´TPF´).

These costs are mainly comprised of lawyers' and experts' fees as well as Court expenses. Financing ends when collection takes place, either from a firm judgment or an out-of-Court settlement. The claimant could also desist from litigating. The funds are disbursed in exchange for a share of the proceeds in case of success of the claim. Should failure take place, the funder has no recourse against the claimant nor a right to reimbursement – thus the ´non-recourse´ financing nature of TPF.

The abovementioned costs could also include the contingency for adverse costs or security for costs, which could be transferred to a fourth party, such as an insurer, in exchange for a premium (after the event or ´ATE´ insurance premium). The adverse costs element will be excluded in this research study, which is a significant exclusion: contingency, premium or liability payment could play an important role in the decision to claim or invest, particularly relevant in some jurisdictions and in case of multiple defendants. Allocation of legal costs vary from jurisdiction to jurisdiction, from predetermined ranges of amounts, mainly in continental Europe, ´each pays his own´ rule applicable in the USA ´loser pays all´ in the UK.

Numerous academic papers address the Third-Party Funding (TPF) discipline. We would like to highlight, *inter alia,* the works of Goeler, J. V. (2016) and van Boom, W. H. (Ed.). (2016).

Common law jurisdictions have historically banned third-party funding under prohibitions against maintenance and champerty. The former is the practice of providing material support to other parties involved in litigation, not limited to a financial point of view. Champerty is a species of maintenance, when the support is given in exchange for something of value contingent on the outcome (Sebok, A. J. (2010); Solas, G. M. (2019)).

In Australia the law was relatively clear about the legality in bankruptcy situations in the 1980s and 1990s (Abrams, D. S., & Chen, D. L. (2012). As from 2006, with the judicial relaxation of the maintenance rules, funding spreads rapidly to the United Kingdom and the United States, and has recently expanded to Singapore, Hong Kong, China, Latin America, and Europe (Waihenya, J. (2021) cit. Sherry Xing Chen & Kirring Hough (2019)). Some regulations or judicial decisions mention TPF to be allowed or prohibited; for instance, in the United States, Florida, New York, Ohio and Texas embrace TPF whilst in others such as Alabama, Colorado, Kentucky and Pennsylvania, TPF is either restricted or considered unlawful (Popp, A. T. (2019)). Arbitration institutions are also taking steps to provide for TPF within their rules such as the Singapore International Arbitration Centre (SIAC) and the China International Economic and Trade Arbitration Commission (CIETAC) (Waihenya, J. (2021)).

Self-regulation has taken place by the members of the industry, through the Code of Conduct of the Association of Litigation Funders based in London, published in November 2011 by the Civil Justice Council, an agency of the UK Ministry of Justice. Worth mentioning are recent developments in the UK which, further to the Jackson Review on civil litigation costs (Jackson, R. M. (2010)) in November 2008 and significant market practice since, a recent Supreme Court reasoning (R (on the application of PACCAR Inc and others) v Competition Appeal Tribunal and others, 2023)) has linked the enforceability of litigation funding agreements (LTA´s) to statue-defined Damages Based Agreements (DBA´s). This would represent to encapsulate litigation funding as an extension of a law firm´s contingency fee agreement and an exception to the champerty prohibition (Steinitz, M. (2014)).

The first attempt to formally define and legislate TPF as a discipline comes from the EU European Parliament resolution of 13 September 2022 with recommendations to the Commission on Responsible private funding of litigation (2020/2130(INL)). From now onwards, it will be referred to as ´the EU TPLF draft legislation´, which has been strongly contested, for the reasons explained later on, by the TPF sector in particular through their Legal Finance Association (LFA) based in Washington. The LFA report, *Resourcing the rule of Law in Europe* (International Legal Finance Association. (2023, June)) was published last July 2023 (´the LFA report´). EU

TPLF draft legislation intends to address litigation in a broad sense, setting examples such as "collective redress, as well as arbitration, insolvency proceedings, investment recovery, anti-trust claims and others". Additionally, it states *"while TPLF is virtually non-existent in Europe, it is a booming phenomenon in investment arbitration that multiplies the number and the volume of claims of private investors against States"*. These last two statements – regarding the non-existence in Europe and the increase in volume caused by TPF - should be calibrated and challenged in the light of empirical evidence.

The EU TPLF draft legislation defines a third-party funding agreement as follows: *"an agreement in which a litigation funder agrees to fund all or part of the costs of proceedings in exchange for receiving a share of the monetary amount awarded to the claimant or a success fee, so as to reimburse the litigation funder for the funding it provided and, where applicable, cover its remuneration for the service provided, based wholly or partially on the outcome of the proceedings.*
This definition covers all agreements in which such a reward is agreed, whether offered as an independent service, or achieved through a purchase or assignment of the claim."

In principle and in a broad sense, it can be stated that Third-Party Funding is, to date, an "unregulated" discipline, in the sense that it lacks formal *ad hoc* statutory definition (i.e. enacted by a legislative body) – with the caveats explained above – and is experiencing increasing regulatory attention as the market increases.

Market Size

The estimated annual size of the Litigation Finance market is uncertain (Strom, R. (2020, June 11)) and difficult to calibrate in a like-for-like comparison. Figures can refer to amounts committed by the investors or to the amount actually invested (assets under management or AuM).

Depending on the sources and methodology, these figures vary: from USD 16 billion (B) in 2022 expected to grow to USD 24.3B by 2028 driven by *"increasing awareness of litigation funding"* (Rationalstat LLC. (2023, August 9)), to close to the

USD 40 B mark (Bloomberg News. (2020, June) cit. Brown Rudnick law firm figures), or expected to surpass the USD 57.2 B threshold by 2035 (Research Nester. (n.d.)).

The difficulty of assessing these figures is driven, *inter alia*, by the mismatch between the amounts in dispute or claimed by the plaintiff and the actual awards, the different perception of the probability to win or lose by both parties - leading, for instance, to the defendants not provisioning the amount claimed in their financial statements - impact of time (interest could be a significant portion of a damages award in a lengthy processes or high-inflation jurisdictions), as well as the confidentiality obligations existing to or requested by the Court or Arbitration Tribunals. Finally, insolvency-related cases are embedded in illiquid and opaque situations within the equity and the gargantuan debt capital markets. The Table 4 shows the estimated committed balance sheet capital of the largest US and UK players, which amount to approximately USD 18 B (Litigation Finance Insider (2023, September 25)).

Ranking	Funder	USD millions*
1	Burford	8.000
2	Omni Bridgeway	1.745
3	Harbour	1.541
4	Therium	1.400
5	Bench Walk	1.200
5	Longford	1.200
5	Parabellum	1.200
6	Curiam Capital	1.100
7	Augusta Ventures	842

**Includes balance sheet capital*

Table 4. Litigation funders committed capital. League Table. Litigation Finance Insider (2023, September 25)

Westfleet Advisors. (2023), with data as of 30 June, 2022, refer to USD 13.5 B Assets Under Management for 44 funders with “substantial participation” in the U.S. commercial litigation finance market.

For the time being, litigation financing is not tracked as Non-Bank Financial Intermediation in the latest Financial Stability Board report. In principle, the published non-recourse commitments seem to be small vis-à-vis the size of the

shadow banking sector (USD 217 trillion allocated to Non-Bank Financial Intermediaries (Board F.S. (2023)).

What seems undeniable at this juncture is that the Litigation Finance sector is a multibillion-dollar industry (i.a. Stroble, J. J., & Welikson, L. (2020)) that has received increased attention in the last two decades, bringing arguments both for and against the practice. When siding with one or the other side, one should always be mindful that a litigation funder is likely to refuse to lend money to plaintiffs with the weakest cases (Abramowicz and Alper, 2013); as a consequence, frivolous claims would not be brought to court.

Litigation Finance in Investment Arbitration

Jurisprudence, academic literature and news articles relating to third-party funding in most jurisdictions largely focus on domestic litigation funding, which represents the majority of third-party funding instances worldwide (*Nieuwveld, L. B., & Sahani, V. S. (2016)).*

There are varied definitions of Third-Party Funding under legal instruments such as international treaties – for instance, article 2 of the EU-Vietnam FTA, or the International Treaty between Canada and the EU. Additional institutional definitions are found in UNCITRAL Working Group III and the Code of Conduct for Litigation Funders of the Association of England and Wales. Scholarly definitions (*inter alia* Sweify, M. F. (2023)) complete this myriad of sources.

TPF has a tradition in the maritime sector through protection and indemnity clubs; in the 1960s the Investor-State Dispute Settlement (ISDS) system allowed foreign investors to bring claims against host States based on Bilateral Investment Treaties, or other international agreements, creating the possibility of large monetary settlements, incentivising funders to take a stake in investment arbitration (i.a. Waihenya, J. (2021); Alexander, C. S. (2023)).

TPF in international arbitration is still not regulated and domestic litigation rules cannot always be automatically applied in arbitration (Frignati, V. (2016) cit. Nieuwveld, L. B., & Sahani, V. S. (2016)). For instance, jurisdictions such as UK, some parts of the USA and Hong Kong consider international arbitration a

completely different system from traditional court litigation; whilst others such as Australia and Singapore apply the same rules in both litigation and arbitration. Finally, in the remaining jurisdictions, including most of Europe, Asia, the Middle East, and Africa, regulation of the phenomenon of TPF is totally absent, both in court litigation and in arbitration.

This landscape is intended to be changed in Europe by the previously mentioned EU TPLF draft legislation: *"whereas, while TPLF is virtually non-existent in Europe, it is a booming phenomenon in investment arbitration that multiplies the number and the volume of claims of private investors against States*".

We challenge the accuracy of this statement for the following reasons:

(a) there seems not to be publicly available information to assess – e.g., out of the current open 1,303 cases in dispute - what percentage are being financed by third parties, nor the number of cases dismissed after due diligence by one or several privately-owned funds; and

(b) the reasons for initiating an investor-State dispute could be varied, and not necessarily due to the existence of TPF; for instance: (i) some cases could be originated by the disputed adherence to certain international treaties (e.g. the Energy Charter Treaty is being exited by several European countries (DW, 2023)), difficult to distil whether the increase in disputes are or could be a consequence of the increase of economic activity; (ii) claimants in sectors such as oil and gas might not need financing, given their financial resources and/or access to capital at a lower cost; (iii) country-specific circumstances have led to numerous cases increasing the awareness and appetite for litigation funders, e.g. alleged expropriations in certain jurisdictions (Latin America accounted for 21 of the 49 new cases in the first half of 2023 as per the 2023 ICSD Annual report (Kluwer Arbitration, 2024), or the paradigmatic case of Spain's change of law regarding renewables in the 2010s with over 50 claims requesting over EUR 8 billion (Prabhu, A. (2023, February 15)), only to mention those with abundant public information.

Despite this “boom in arbitration” - Dafe, F., & Williams, Z. (2021) refer to 2018 as the highest number of new cases registered in ICSID quoting UNCTAD 2019 report figures -, there is no empirical evidence of a direct correlation between greater access to finance with the willingness of a claimant to initiate a dispute whilst dedicating own or a third party´s resources.

There are recent considerations of making TPF available to the respondent States, which might face significant costs against resourceful corporations. As previously stated, this research will focus on the perspective of the claimant or its prospect funder, and on the arbitral justice as an appropriate forum to secure a fair process from both a procedural and a substantive perspective (Sweify, M. F. (2023)).

The risk transfer to a third party – potentially triggering moral hazard – as a cause to increase the number of disputes is challenged by the fact that the third-party funder does not in turn have unlimited resources, and would require to analyse the merits of the claim in depth, so as to preserve and remunerate the capital from its investors.

Moral Hazard Considerations

According to the Oxford American Dictionary, moral hazard can be defined as the “lack of incentive to guard against risk where one is protected from its consequences e.g., by insurance”. Transferring the risk to a third-party funder raises moral hazard for the plaintiff e.g. by not cooperating if risk is transferred. The underlying principle is that the justice system should address only meritorious cases, at least in theory.

We have already pointed out how historically, the maritime business embraced the non-recourse loan as an insurance against damages to the contents of a ship. The sale of a ´litigious credit´ is also present in Lex Anastasiana (506 AD)*:* "he who gave amounts to receive claims should not obtain more than what he gave in exchange".

Once and only if a claim has been contested, should the claimant have assigned the credit to a third party in exchange for an amount - allegedly equal or lower to the expected amount -, the respondent may liquidate its debt by reimbursing the claimant said exact amount. This practice was not only meant to avoid abuses to those claimants, given their weak bargaining power based on their need for actual

or future proceeds, but also to shorten the process (Spanish Supreme Court. (2008)). Regulation thereof has disappeared in jurisdictions such as Italy and Portugal, whilst it survives in France (article 1,699 Civil Code), Spain (article 1,535 Civil Code) and other countries such as Egypt, Philippines, and some LatAm countries (Moya Fernández, A. J. et al (2016)).

Ethics and legal implications in litigation finance have been flagged i.a. by Goral, R. (2015); Nieuwveld, L. B., & Sahani, V. S. (2016); Frignati, V. (2016); Popp, A. T. (2019); and Bedi, S., & Marra, W. C. (2021). The reference to "the possible future bundling, securitisation and trading of legal claims" would require a deeper and lengthier debate, outside the scope of this research. It is nevertheless important to bring forward the drawbacks that relate to:

(i) The potential influence of the funder in the process, as the possible *conflict of interest* of the funder and the client or with the lawyer of the client or with the defendant, as well as between the funder and the arbitrators, who must remain independent and impartial; and
(ii) "The potential use of the judicial system for financial speculation", including the potential "increase of new frivolous cases or the possible discouragement of settlements in favour of fighting a larger recovery".

Conflicts of interest have been addressed by the International Bar Association (IBA), which issued revised Guidelines on Conflicts of Interest in International Arbitration in November 2014require arbitrators to disclose connections to funders when relevant to the case (Nieuwveld, L. B., & Sahani, V. S. (2016)). The ICSID and UNCITRAL working groups are inclined to support disclosure of TPF as well as its impact on security for costs whilst the International Chamber of Commerce obliges parties - enforceable since January 2021 – obliges to disclose the existence of TPF and the identity of such funder (Waihenya, J. (2021)). The IBA guidelines have been revisited in February 2024 (International Bar Association. (2024)), confirming the principles and disclosure trend set up one decade ago.

In general terms, there seems to be consensus around acknowledging the need to disclose the existence and identity of the funder, albeit not as much as peaceful

conclusion regarding the terms of the litigation finance agreement, since critical information could transpire from the prior due diligence; for instance, commercial terms could be an indicator of the robustness of the merits of the case - a higher cost of capital for the claimant could mean additional risk and therefore lower merits.

Empirical data on behavioural implications is in many instances handicapped by the confidential nature of the business. Having said this, facing the arguments against litigation finance - mainly related to ethical and moral considerations- are confronted by numerous arguments in favour.

To start with, financing has been undertaken *de facto* by law firms' contingent or conditional fee arrangements, although these arrangements exclude covering for other costs, such as experts´ fees or Court expenses, which could be significant and necessary to initiate or maintain the claim. Other risk transfer instruments in the legal insurance industry vary from securing adverse costs liabilities to the actual subrogation of the insurer in the position of the claimant.

Large capital providers such as banks and insurance companies have often stayed away because of the legal uncertainty associated with litigation funding (Molot, J. T. (2009).

Van Boom, W. H. (Ed.). (2016) points out that, additional and beyond access to justice for the claimant, the interest of TPF by the financial industry is self-explanatory: not to drain the claimant´s liquidity, additional comfort on cross-checking the merits of a case by a third-party and reducing the potential costs of enforcement of a favourable judgment. Additionally, investors see TPF as an uncorrelated asset to the stock or bond markets; disregarding unmeritorious cases would benefit *both* parties and the system. Needless to add, for the claimant lawyers´ perspective, TPF reduces the risk of non-payment.

Access to finance to defendants through TPF is also a trend, although it is acknowledged that evaluating and pricing defences would be more difficult than funding claims (Stadler, A. (2016) cit. Molot, J. T. (2010)). In any case, the respondent could also benefit from a third-party analysis – Abramowicz and Alper (2015) have suggested a claim-screening regime by virtue of which "claims below a

certain probability of success *cannot* be brought, while claims above a certain probability of success *cannot* be defended against".

A recent award is worth mentioning in the case Petersen and Eton Park v. Argentina and YPF (September 2023), where the Court states that the role of the funder, Burford Capital, should be considered as financing : "*The relevant question is what the Republic owes Plaintiffs to compensate them for the loss of the use of their money, not what Plaintiffs have done or will do with what they are owed. The Republic owes no more or less because of Burford Capital's involvement. (…) If Plaintiffs were required to trade a substantial part of their potential recovery to secure the financing necessary to bring their claims, in Petersen's case because it was driven to bankruptcy, and litigate their claims to conclusion against a powerful sovereign defendant that has behaved in this manner, this is all the more reason to award Plaintiffs the full measure of their damages."*

This interpretation backs those positions that TPF is, simply put, an additional and alternative solution to better serve justice (Affaki, G. (2013) and avoid debates with allegations, such as TPF bringing up "more and riskier claims against States" (Davitti, D., & Vargiu, P. (2023)), not based, at least to date, on specific empirical evidence.

Beyond the access to justice principle, cases with lower probabilities of success will most likely not reach the court, alleviating the justice system (less costs and time in unmeritorious claims). The net impact – the decrease and increase of cases - is nevertheless difficult to evaluate, given the lack of publicly available data of those cases that, after a study, do not reach Court. In general terms, the funders analyse more cases than they finally fund, as will be explained later on.

Litigation finance has been a phenomenon that has increased in the last fifteen years, with particular strength in Australia, Singapore, and the ´common law´ countries such as US, UK and Canada. The potential repercussions vis-a-vis consumer rights have, *inter alia*, inspired the EU TPLF draft legislation. Its preparatory work (Voss, J. (2022)) – the Voss report, also deems excessive the returns from the funders, mentioning *"in Europe, rates of returns for litigation funders*

may be up to 300% or even 3.000%" - without further reference to the period of the investment - which has led to the proposal of a cap for the funders return to 40% of the proceeds.

It is worth mentioning the reaction through forward-looking requests of the International Legal Finance Association (ILFA), headquartered in Washington, through their report (International Legal Finance Association. (2023, June). The ILFA has suggested: (i) *"to ensure that any proposals on regulating legal finance are based on evidence which demonstrates a need for EU intervention, in line with the principle of subsidiarity"*; (ii) to consult with key consumer rights groups, the European Innovation Council and SMEs Executive Agency on the impact of *curbing legal finance on emerging technologies and breakthrough innovations;* and (iii) to await the full transposition and implementation of the collective Redress Directive in all member states to gather evidence on the legal financial model of investment and the interaction between consumer rights and business-to-business (B2B) litigation.

The extension period for the Directive to be enforceable has been prudently granted.

The EU Draft TPLF legislation, by adding conduct obligations to litigation funders and to their ecosystem, added to the implementation of a supervisory scheme of the Litigation Finance industry in Europe, will most likely increase the cost and regulatory burden to existing players, may preclude new litigation funders from entering the market, as well as potentially transferring the ultimate risk of the protection to the taxpayers. Whilst the size of the shadow banking industry, at least at present, does not seem to represent a systemic threat, some degree of checks seems in order, motivated not only by preserving the rights to consumers, but also the impact in reputation and contingent off balance-sheet commitments of defendants.

This research will defend that Litigation finance is a non-recourse financing that should be available to both claimant and defendants and not be discriminated from any other type of financing. Any complaint about the excessive returns allegedly perceived by a litigation funder should be considered in the context of the associated risks, and the time to the reward. The litigation finance practice, married to Artificial

Intelligence techniques, when carefully addressed in terms of methodology to explain and/or predict, can contribute to bringing transparency to an increasingly sophisticated market.

B - Assessment Criteria by a Litigation Funder

Following Jonas von Goeler (2016), litigation funders' assessment criteria range from underwriting guidelines or informal methods to complex probabilistic evaluation financial models that compare different estimated outcomes of the case and damages award probabilities calculated over the stages of the proceedings. The latter approach is not unusual in the asset finance discipline, for instance the reserve-based lending in the oil and gas sector. Such quantitative assessment would be complemented by a ´softer´ qualitative assessment criteria. The quantum is benchmarked against the expected investment - budget or costs - and the time to the judgment, award or early settlement, in the light of the merits of the case. The minimum projected quantum applied to arbitration varies between litigation funders, ranging between USD 1 million (´M´) and USD 25 M or more. Actions requesting specific performance, injunctive relief, or declaratory judgment are not fundable unless they lead or can be translated to a specific monetary relief.

Under examination are facts, evidence (existing or to be discovered / disclosed), procedural framework, applicable legal base and/or case law. Additional factors highlighted by the previous author are the solvency and degree of risk aversion of both the claimant and the respondent, the motives to bring the claim, including the rationale of previous unsuccessful attempts at obtaining funding, the parties' respective incentives to settle, the legal teams involved, the arbiters or internal variables related to portfolio management considerations and investment policy (duration, amount, diversification by nature of the claim or geography…) and the enforceability of the award.

In this research, we will refer to these as external factors, exogenous to the merits of the case. We will also assume De Morpurgo (2011) assumptions on the parties in TPF: they are rational and risk-neutral.

The award expected by the litigation funder could be in the form of a multiple of the investment (´the multiple approach´) or a percentage of the recovery of the proceeds (´the percentage approach´), whichever is larger. As a rule of thumb, litigation funders look for a ratio of investment costs to expected recovery of around one to ten (1:10), with a minimum ratio of one to four (Rowles-Davies, N., & Cousins, J. (2014)).

The most significant cost items include the legal fees of the claimant's counsel – which could comprise or include a contingency or conditional fee - together with expert fees, arbitrator and administrative costs, as well as potential adverse costs. The latter could be substituted by costs for obtaining legal insurance or directly granting security for these costs. This premium could be expensive for some funders and could be not recoverable as per the applicable law and jurisdiction.

The operating costs of the funded party – for instance, working capital of an SME or a start-up with an allegedly infringed patent portfolio – are in principle excluded, unless expressly agreed.

Jonas von Goeler (2016) also points out that unresolved claims and litigation funding is also challenging from an accounting – and auditing – perspective for both claimant and the defendant: when a gain or a provision needs to be recorded in the Profit and Loss account. This has a particular impact of fair valuation of investments of listed litigation funders. According to PWC, auditor to Juridica Investments Limited (2013 Annual Report) – which filed for bankruptcy in 2017 -, investments in claims are categorised as contractual interests held at fair value through profit or loss; these fair values may differ materially from the realisable values, given the inherent uncertainty associated with the valuation of such non-current assets and the absence of a liquid market. The auditing firm approach is described, where *"a range*

of settlement proceeds are assigned to different scenarios and probabilities of outcome including Monte-Carlo method analysis, leading to expected outcomes and other statistical data used to calculate the future valuation of each contractual interest. A discount rate is then applied to the future value to determine the current fair value".

Using large datasets from the Stanford Intellectual Property Litigation Clearinghouse, a risk analysis model for intellectual property litigation has been built and tested based on "*prior factors"* – i.e. factors that do not model directly the merits of the case but instead focus on past information that may influence the outcome of the current case (Surdenau et al (2011)), concluding that *'IP litigation is a problem fit for forecasting'*.

Worth citing once more to Jonas von Goeler "whatever the degree of financial and mathematical sophistication, a litigation funder's case assessment is ultimately more art than science, (…) little different from other businesses whose activity is centred around employing specialized processes and methods to create subjective probabilistic value judgments relating to uncertain future events, be it in the investment business – cit. Molot, J. T. (2014) - or in the area of financial intermediaries, notably credit rating agencies".

On a more theoretical level, the extent to which litigation outcomes are predictable is increasingly attracting attention from both law and economics literature (McShae et al (2012), Surdenau et al (2011)), which will be discussed with more granularity later on in this research.

Accounting Standards and Burford Capital Fair Valuation

Accounting Standard Classification (ACS) 450-20 and International Accounting Standards (IAS) 37 provide guidance on the recognition and measurement of provisions and contingencies. From a respondent´s perspective, a contingency is recognized if the following criteria are met: (i) it is probable that a liability has been incurred at the date of the financial statements and (ii) the amount can reasonably be estimated. If a range of the estimated loss exists and no amount within the range is more likely, the lower range would apply.

From a claimant´s perspective, gains are not recognised until they are realised. Legal costs are expensed as incurred. From the perspective of the funder, these assets are considered as financial investments accounted at fair value.

Burford Capital, with a GBP 2.6 B market capitalization as of 22nd of April 2024, is considered the largest, by capital commitments, global finance and asset management firm focused on law (see Table 4). Capital Provision Assets (CPAs) are accounted at fair value using an income approach (Accounting Standards Codification Topic 820-Fair Value Measurement ("ASC 820")), by virtue of which "*fair value (is) based on estimated, risk-adjusted future cash flows, using a discount rate to reflect the funding risk of deploying capital for funding capital provision assets".* Said approach "*requires management to make a series of assumptions, such as discount rate, the timing and amount of both expected cash inflows and additional fundings, and a risk-adjustment factor reflecting the uncertainty inherent in the cash flows primarily driven by litigation risk, which changes as a result of observable litigation events. These assumptions are considered Level 3 inputs"* (Burford Capital (2023a)).

Burford Capital measure and report financial instruments at fair value. Fair value represents *"the price that would be received to sell an asset or paid to transfer a liability (an exit price) in an orderly transaction between market participants at the measurement date"* (2022 Annual Report, page 96). The financial instruments are classified and disclosed based on the observability of inputs used in the determination of fair values as follows:

Level 1—quoted prices (unadjusted) in active markets for identical assets or liabilities that the reporting entity can access at the measurement date;

Level 2—inputs other than quoted prices included within Level 1 that are observable for the asset or liability, either directly or indirectly; and

Level 3—unobservable inputs for the asset or liability.

As of 31 December 2022, the proportion of the fair value of total assets based on subjectivity ("level 3") is 96,1%, increased to 97.5% as of 30th September, 2023.

USD ´000	Level 3	Total Assets	%
31.12.2022	3.833.393	3.988.496	96,1%
30.09.2023	4.943.781	5.072.938	97,5%

Table 5. Proportion of Level 3 assets in fair value to Total Assets as of December 31, 2022 and September 30, 2023 (Burford Capital, 2022 Annual Report, page 115, and Burford Capital Quarterly Report, September 2023, page 25).

As of June 30, 2023 Burford Capital´s interim report elaborates further: *"The preparation of the Group's condensed consolidated financial statements requires management to make estimates that affect the reported amounts of assets and liabilities at the date of the condensed consolidated financial statements, the disclosure of contingent assets and liabilities at the date of the condensed consolidated financial statements and the reported amount of revenues and expenses during the reporting periods. Such estimates include, among others, the valuation of capital provision assets, which requires the use of Level 3 valuation inputs, and other financial instruments, the measurement of deferred tax balances (including valuation allowances) and the accounting for goodwill. Actual results could differ from those estimates, and such differences could be material"* (Burford Capital. (2023-b).

The value of the cases at the beginning and at the end of the period are accounted for. These values might have experienced deployments (investments) and realisations (divestments) across the different investment options: ´single case´ (related to one counterparty) or aggregated in a portfolio. For instance, the circa

USD 3.8 B Total Capital Provision Assets (CPAs) as of 31 December 2022 increased to circa USD 5 B as of 31 December 2023 according to the Letter to Shareholders by the Managing Board (Burford Capital. (2024b, March 14)). The Total Capital Provision Assets (CPAs) as of 2022 year-end are broken down by type: 57% portfolio, including portfolio with equity risk, 21% single cases, and 22% funds, joint ventures and equity method investments (Burford Capital 2022 Annual Report, page 117).

"*The amount of unrealised gain or loss in the relevant period is added to or subtracted from, as applicable, the asset or liability value in their condensed consolidated statements of financial position*" (fair value adjustment, Burford Capital Annual Report, 2022, page 12). As per their release to the market in May 2023 *"a cash flow forecast is developed for each capital provision asset based on the anticipated capital commitments, damages or settlement estimates, and our contractual entitlement. (…) Capital provision assets are recorded at initial fair value, which is equivalent to the initial transaction price for a given capital provision asset, based on an assessment that it is an arm's length transaction between independent third parties and an orderly transaction between market participants."* (Burford Capital Limited. (2023, May 16)).

The 2022 Annual Report further explains the methodology used for valuation: *"Using the cash flow forecast and a discount rate, an appropriate **risk adjustment factor** is calculated to be applied to the forecast cash inflows to calibrate the valuation model to the initial transaction price. (…).* Each reporting period, (i) the cash flow forecast is updated based on the best available information on damages or settlement estimates and it is determined whether there has been an objective event in the underlying litigation process which would change the litigation risk and thus the risk-adjustment factor associated with the capital provision asset; and (ii) the updated risk-adjusted cash flow forecast is then discounted at the then current discount rate to measure fair value.

The abovementioned *"objective events"* could include, among others (…) *"in arbitration cases, where there are limited opportunities for appeal, issuance of a tribunal award; or an objective negative event at various stages in the litigation*

process. Each reporting period, the updated risk-adjusted cash flow forecast is then discounted at the then current discount rate to measure fair value" (Burford Capital Annual Report, 2022, page 47).

A closer look into Burford 30th of September 2023 results (Burford Capital Quarterly Report, September 2023), show the breakdown of USD 3.9 B Level 3 assets and liabilities, net, comprised of cost plus unrealised fair value.

Level 3 assets and liabilities (USD ´000 30.9.23)		Cost	Unrealized	Fair Value	%
Positive case milestone factor	44%	869.330	654.031	1.523.361	40%
Negative case milestone factor	6%	128.815	- 76.617	52.198	1%
No case milestone	47%	932.546	48.170	980.716	26%
YPF-related assets case	3%	57.850	1.227.918	1.285.768	33%
	100%	1.988.541	1.853.502	**3.842.043**	100%

Table 6.A Breakdown of Level 3 assets and liabilities, net. Principal value technique: Discounted Cash Flow (Burford Capital (30 September, 2023), page 30).

Note the impact of a single case (YPF-related), which can drive a large change in the valuation of the company: 3% of the cost can account for one third (1/3) of the FV. The FV of USD 1.2 B as of 30th September, 2023 increases to USD 1.4 B as of 31st December, 2023 (Burford Capital. (2024c, March 14).

Fair Value (USD)	Discount Rate (weighted average)	Duration Years (weighted average)	Adjusted Risk Premium
3.842.043	7,9%	3,5	29,8%

Table 6.B Weighted Average of the Discount Rate, Duration and Adjusted Risk Premium (Burford Capital (30 September, 2023) page 30).

This portfolio of assets and liabilities, net, has (i) a "weighted average discount rate" on the expected future cash flows of 7.9% - embedding the impact of the underlying market rates have had across all currencies and tenors and the potential impact of "inflationary pressure in the period"; (ii) "weighted average duration" of 3.5 years and (iii) an "**adjusted Risk Premium" metric** of 29.8% (weighted average, applied to unrealized fair valuation). As part of their fair value methodology, *"this metric is a*

risk adjustment (haircut) *applied to the potential proceeds due to us in the event of a successful litigation outcome due to the remaining litigation risk".* Worth mentioning the minimum and maximum values of the above weighted averages: discount rate ranges between 6.3% and 8.0%; the tenor between 0.3 and 7.5 years whist the adjusted Risk Premium ranges from 0% to 100% (Burford Capital Quarterly Report, September 2023, pages 30 and 43).

The funder´s Managing Board seems *"not much concerned with nominal interest rates, given that they would translate into higher nominal returns on cash on hand, leaving the cost of debt on uninvested capital roughtly the same"*, whilst *"most jurisdictions either apply floating rates ot ligitation judgments or use fixed rates generally above market rates"* (Burford Capital. (2024b, March 14), page v).

We will also deep dive and focus on how the latter parameter, the adjusted Risk Premium, is reported. In page 29 of the quarterly report elaborates further on the valuation policy for capital provision assets: ranges of percentages are applied against the risk adjustment factor to more than 70 discrete objective litigation events – also referred to as "litigation milestones" - across five principal different types of litigation in order to calculate the adjusted Risk Premium. The range for each event is 10%, typically marking assets at the middle of that range unless there are specific factors that cause the valuation committee to select a different point in the range or, exceptionally, outside that range. To decide which percentage to apply to a given asset, the committee considers the kind and degree of legal, procedural or other investment-specific circumstances.

We extract in Table 6.C the key unobservable inputs used to value their capital provision assets and the applicable ranges and weighted average by relative fair value for such inputs.

´Milestone factors´ to reevaluate the pricing relate to (i) significant ruling or other event (prior to trial court judgment), (ii) trial court judgment or tribunal award, (iii) appeal judgment, (iv) exhaustion of as-of-rights or all appeals, (v) asset freeze, (vi) settlement and (vii) portfolios with multiple factors.

"In a small number of instances, a secondary sale of a portion of an asset or liability, the market evidence is factored into the valuation process to maximize the use of relevant observable inputs" (Burford Capital Limited. (2023, May 16)). As a consequence, until there is a trigger, for instance, a judgment or award or a secondary sale, assumptions need to be made based on ´unobservable inputs´. Uncertainty in the cash flows is primarily driven by the inherent litigation risk, subject to further precision on a case-by-case basis.

Case Milestone factor	30.09.2023	Weighted average
(i) Significant ruling or other objective event prior to court judgement	123.456	from 5% to 40%
(ii) Trial court judgment or tribunal award	202.590	from 4% to 60%
(iii) Appeal judgment	108.394	from 60% to 80%
(iv) exhaustion of as-of-rights appeals	95.661	N.A.
exhaustion of all appeal	134.845	N.A.
(v) Asset Freeze	26.279	from 20% to 20%
(vi) Settlement	18.046	from 20% to 80%
(vii) Portfolios with multiple factors	813.935	from 1% to 100%
Other	155	
Total positive case milestone fair value	**1.523.361**	
(i) Significant ruling or other objective event prior to court judgement	18.427	from -10% to -60%
(ii) Trial court judgment or tribunal award	17.435	from -10% to -60%
(iii) Appeal judgment		from -100% to -100%
(vii) Portfolios with multiple factors	16.336	from -3% to -60%
Total negative case milestone fair value	**52.198**	
No case Milestone	**980.716**	
YPF-related case	**1.285.768**	
Total fair value (USD ´000) weighted aver	**3.842.043**	

Table 6.C Ranges of the case milestone factors (positive and negative) applied to the weighted average Fair Value (USD ´000) as of 30.09.2023. Burford Capital (30 September, 2023) page 30.

The milestone factors, as shown in Table 6.C, are embedded within disclosed ranges; e.g. from 5 to 40% (significant ruling), from 60 to 80% (judgment or award), from 4 to 60% (appeal) or from 1 to 100% (portfolios with multiple factors).

Further granularity is given in terms of how these weighted averages of risk premium are built, pondered by the milestone factors – in either ´negative´ or ´positive´ cases. These factors are dynamic, and vary from quarter to quarter (see Table 7), connected to the particular values of the cases and assets they are assigned to, and the events alongside these proceedings. Quarter by quarter, not all the amounts are repriced. For instance, as of 30 September 2023, close to half (47%) of the fair valuation of the portfolio does not experience a positive or a negative milestone factor.

	Case Milestone Factor	Positive case		Negative case	
	Weighted average as of	31.12.22	30.09.23	31.12.22	30.09.23
(i)	Significant ruling or other objective event*	20%	22%	-13%	-40%
(ii)	Trial court judgment or tribunal award	53%	52%	-56%	-59%
(ii)	Appeal judgment	67%	72%	-80%	-100%
(iv)	Exhaustion of as-of-rights appeals		80%		
	Exhaustion of all appeald		100%		
(v)	Asset Freeze	20%	20%		
(vi)	Settlement	76%	58%		
(vii)	Portfolios with multiple factors	14%	20%	-50%	-31%
	Other		100%		

**prior to trial court judgment*

Table 7. Case Milestone Factors (weighted average) for negative and positive cases as of 31 December 2022 and 30 September, 2023. Adjusted Risk Premium ranges applied to level 3 assets and liabilities (Burford Capital (2023a and b)).

According to E&Y LLP, auditor of 2022 Burford Capital (Annual Report, page 61) stated *"(…) While the potential range of outcomes for the assets is wide, our fair value estimation is our best assessment of the current fair value of each asset or liability. Such an estimate is inherently subjective, being based largely on management's estimate of forecasted cash flows, an assigned discount rate, and an assessment of how individual events have changed the possible outcomes of the asset and their relative probabilities and hence the extent to which the fair value has altered. The aggregate of the fair values selected falls within a wide range of reasonably possible estimates. In our management's opinion, there is no useful alternative valuation that would better quantify the market risk inherent in the portfolio and there are no inputs or variables to which the values of the assets are correlated other than interest rates which impact the discount rates applied".*

Additionally, it is added "*auditing management's judgments and assumptions used in the valuation of the CPAs involved complex auditor judgment, due to the selection of the valuation methodology utilized, as well as the significant estimation uncertainty associated with the unobservable inputs described above and the forward looking and judgmental nature of the estimated future cash flows. To test the fair value of the Capital Provision Assets (CPAs), we performed audit procedures that included, among others, evaluating the valuation methodology used by management in determining the fair value of the CPAs, against the requirements of FASB Accounting Standards Codification Topic 820 – Fair Value Measurement."*

For the time being, we would like to highlight that the accounting from a funder´s perspective depend on specific contractual terms and there is an acknowledged significant degree of subjectivity involved in the assumptions. There is however an methodology to classify and monitor fair valuation for reporting purposes.

According to Burford Capital, on SEC reporting: *"in our forthcoming form 20-F we will be reporting a material weakness in our internal control over financial reporting and that our disclosure controls and procedures were not effective in the aftermath of the adoption of our new valuation policy. That material weakness and corresponding controls determinations relate to the lack of documented available evidence demonstrating the precision of management´s application of the process to determine certain assumptions used in the measurement of the fair value of capital provision assets. To be clear, this is an issue of internal documentation of a management process, no material accounting errors were identified as a result"* (Burford Capital. (2024a, March 14)).

Whilst it is understood that no accounting errors have been identified, the statement highlights the importance to track and document the process and records management assumptions used in the measurement of the fair value. They also rightly acknowledge and warn, as a potential disclaimer vis-a-vis investors, *"valuation uncertainty with respect to the fair value of our capital provision assets"* and add *"the estimates include, among others, the valuation of capital provision assets, which requires the use of Level 3 valuation inputs".*

This research proposes a portfolio approach methodology within the subsegment of international investment arbitration based on historical information. Albeit cannot guarantee accuracy in or ensure the prediction, the methodology could lead to an objective (observable) proxy, an additional tool to benchmark fair valuation in Investor-State international arbitration cases contributing to reduce the degree of subjectivity and manage risks.

Expected Value and Probability of Success

Hereinafter we select those theories addressing the expected value of a claim that are relevant for our research. For Cooter, R. D. (1991) – see also Cooter, R., & Ulen, T. (2011) in a dynamic context -, the Expected Value (EV) of a trial can be represented as described below:

$EV = p(S) \times Q + p(L) \times 0 - costs$
where
p(S)= Probability of Success;
p(L)= Probability of not succeeding or Losing;
Q = amount of claim requested or expected quantum; and
costs= costs of the trial (legal, experts and administrative fees)
Since the probability of losing where p(L) = (1- p(S)) and the amount to be received in such case would be zero, the equation could be simplified as
$\boldsymbol{EV = p(S) \times Q - Costs}$

Table 8. Expected Value (EV) of a claim formula (Cooter, R.D. (1991), Cooter, R., & Ulen, T. (2011)).

De Morpurgo (2011), based on Shavel basic formula for litigation (Kaplow, L., & Shavell, S. (2002)), elaborated a basic economic model on TPF defining the conditions under which suits are brought, representing them as follows:

$\boldsymbol{Costs < \lambda R}$
where
C are the costs to litigate; and
λ the probability to obtain a given return (R).

Table 9. Formula for litigation (Kaplow, L., & Shavell, S. (2002))

Kidd, J. (2011) states that any litigant must "weigh the costs of pursuing a case against the perceived benefits of prevailing, adjusted for the probability of success. If the expected returns are greater than the cost, then the case will proceed". To the latter definition "greater than the cost", we would add "taking into consideration the value of money over time".

Solas (2019) elaborated on the conditions under which the funder and the claimant, assuming rational behaviour and symmetry of information, decide to enter into a Litigation Finance (or Funding) Agreement (LFA) under either the American or the English rules on cost allocation.

From the lexicon used by lawyers, Fore, J. (2019) proposes the following lexicon on the probability to win a case:

	Probability
Almost certain	from 90 to 100%
Very likely / Very probable	from 75 to 90%
Likely / Probable	from 60 to 75%
More likely than not	from 50 to 60%
Unlikely / improbable	from 20 to 50%
Very unlikely / very improbable	from 10 to 20%
Almost no chance	from 0 to 10%

Table 10. Numerical estimates of verbal probabilities from reviews of empirical research (Fore, J. (2019))

As pointed out by Dhami, M. K., & Mandel, D. R. (2022), the recipients of these probabilities in cognitive sciences would prefer numeric precision for decision making. For instance, in the medical domain, they highlight (cit. Wiles, M. D., Duffy, A., & Neill, K. (2020)) that patients interpret risk probability terms as referring to greater risk than clinicians. In this research we will aim to get a deeper understanding of the probability of success from a less subjective perspective – based on data. Before that, it would be useful to review the criteria used to date and how it has been captured by academia from the Litigation Finance industry.

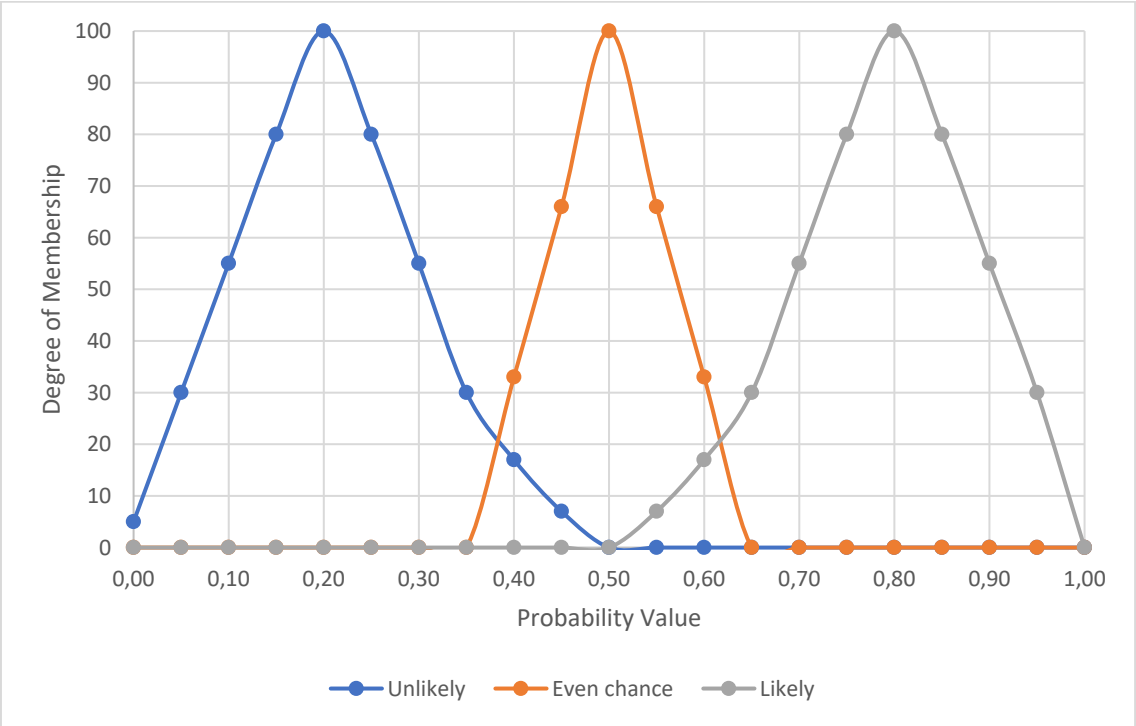


Figure 3. Trends in cognitive Sciences (Dhami, M. K., & Mandel, D. R. (2022)). Hypothetical membership functions for three probability terms ("unlikely", left, "even chance", center and "likely", right).

Probability of Success in Investment Arbitration

Litigation funders usually work with percentages, either related to positive award outcomes or likelihood of settlement (Jonas von Goeler (2016)). An ICC France study found that, on average, litigation funders only accept between 5% and 10% of cases received (Veljanovski, C. (2011), corroborated by independent research and anecdotal evidence.

A claim's probability of success essentially results from "the description of core elements of facts, law and evidence". Litigation funders have to be selective about the cases to fund because choosing 'losers' would erode their capital and result in failure (Shepherd, J. M., & Stone, J. E. (2015)).

Reportedly, litigation funders tend to require a probability of about 60%-70% or more: (i) *"at least 60%"* (Cremades, B. M. (2013) cit. Affaki, G. (2013)) ; (ii) *"70% or more"* (Veljanovski, C. (2011) and Scherer, M. et al (2012)); and (iii) "*Funders generally look for claims that have about a 75% chance of being successful"* (Seidel,

S. (2013)). In practice, the 60% is the most commonly used – see for instance, Hodges OBE, C., Peysner, J., & Nurse, A. (2012), Rowles-Davies, N., & Cousins, J. (2014) (cit.) and Osborne Clarke, n.d. (2016).

According to Jonas von Goeler in 2016 - echoing Nieuwveld, L. B., & Sahani, V. S. (2016) and Hodges, C., Vogenauer, S., & Tulibacka, M. (Eds.). (2010) – unlike the insurance business, where premiums and underwriting policies have standardised risks for decades, the litigation funding industry has not reached standard contract terms, "*particularly true with regard to the funding of international arbitrations, a comparatively young industry segment in which, at this point, only a handful of litigation funders worldwide show significant activity"*. The same authors highlight the challenges of arbitration detrimental to the predictability of a case's outcome in comparison with litigation: more flexible, often discretionary approach towards jurisdiction (by nature, multijurisdictional), merits, and procedure; absence of formal precedent and the limited transparency of arbitral decisions resulting from confidentiality. They claim that in advocating for a balance between the risks and returns of investing in international arbitration, finding the right answer is *"anything but simple"*, requiring special expertise.

Risk-Reward metrics

A Litigation Finance Agreement (LFA) defines the outcome and specifies the return structure depending on the level of success; as previously pointed out, the proceeds need to consist of cash or subject to monetisation *(*Jonas von Goeler (2016)). Litigation funders typically seek to recover their investment to date (outlay), *plus* the higher of either a multiple of that outlay - we have referred to as ´the multiple approach´ - or a percentage of the proceeds - ´the percentage approach´.
Fee structures may contemplate several scenarios that factor internal rate of return and risk-reward over time (e.g. through a return and time grid).

The multiple varies from 2.5 to 4 times its up-front investment (cit. Cremades Jr., *Spain Arb. Rev.* (2012)); Veljanovski, C. (2011) ranges this multiple threshold from 1.5 to 6 times the capital invested. Jonas von Goeler (2016) cited Seidel's work from the ICC Dossier Third-Party Funding (p. 16, 27) with a funder´s investment at USD 5 M and a projected amount of about 100 M, with a return of three times the investment i.e. 2x plus recovery of the investment or 3 times.

The percentage ranges between 20% and 40% of the recovery (Seidel, S. (2013), Pinsolle, P. (2011)); from 30 to 60% (Abrams, D. S., & Chen, D. L. (2012)); “from 20% to 40% of the recovery, occasionally 50% or more” ((Veljanovski, C. (2011)), between 15% and 50% (Khouri, S., & Hurford, K. (2012)), between 25% and 30% (Hodges et al); between 10 and 45% (cit. Cremades Jr., *Spain Arb. Rev.* (2012)); between 20 and 30% (Kohlmeier, T. (2017)) or from 10% to 40% (Osborne Clarke, n.d.).
Minimum guaranteed returns are usually set at the amount of the outlay to ensure the recovery of the investment whilst some funders use capped returns at around 50% of the recovery in order to keep the interests of the funded party aligned.

Litigation funders´ returns are meant to be comparable to those demanded by equity investors in venture capital and project finance contexts (Ishikawa, T. (2021, June 25)), although more empirical evidence is required for this particular benchmark.

The investors of a litigation funder aim to receive above three times the capital invested; this is said to generate an “investment return on the portfolio of 15% to 20% over five to six years" (Seidel, as cited in ICC Dossier TPR, 16,27; cited by Jonas von Goeler, 2016). The same author states that international arbitration tend to be priced higher than commercial litigation claims to compensate for higher loss rates, due to the “comparatively low predictability of outcomes and low settlement rates”.

Steinitz, M. (2013) addresses how the funders must assess the value of the investments for their investors. She considers, based on economic theory, that Venture Capital (VC) contracts can provide a model for litigation-funding contracts, being both business models characterised by “extreme uncertainty, information asymmetry and agency problems”. Litigation claims would then be mirrored to VC investments in start-ups with R&D projects. In her review of the history of legal claims as assets, she advocates for the *real options theory*, which is able – unlike the Discounted Cash Flows or Net Present Value (´NPV´) models – to capture the investor´s flexibility to adapt, with new information, to market (i.e. court) developments. Following Cornell, B. (1990), the interest in real options theory vis-à-vis Net Present Value comes from the ability to capture unexpected market developments: the more uncertainty, the more value. The more uncertain the outlook of the litigation option, the more valuable is this flexibility.

Rhee (2006) referred to a ´random walk´ (coined name in 1905 by Karl Pearson) stating that “like stock prices, perceived case values ´fluctuate´ based on new information (…) had there been a (more liquid market in legal claims), their price would trade within a broad range, moving stochastically upon the disclosure of new information and events”.

On return adjusted to risk, we side with those that claim the possibility to reach an approximation to an ´objective´ framework of probability of success referred to a legal case based on historical information. Aware that accuracy in the forward-looking prediction cannot be guaranteed since we leverage on past information, an objective (observable) proxy is sought.

Finally, it would be important to highlight that Burford Capital seems to confirm that the Multiple On Invested Capital (MOIC) and the Internal Rate of Return (IRR) are the critical metrics form a risk reward perspective: *"we believe that deployments and realizations, as well as two key performance metrics derived from those cash outflows and inflows, internal rate of return and return on invested capital are the best measures of the success of our business"* (2022 Annual Report, page 2).

C - Legal Tech and Legal Analytics in Legal Research

Artificial Intelligence (AI) is playing a role in shaping economic and financial sector developments and is perceived as an engine of productivity and economic growth through efficiency, improved decision-making processes and the creation of new products and industries (Ghiath Shabsigh and El Bachir Boukherouaa, 2023).

Another subset of artificial intelligence is Deep learning (DL), an advanced ML, where computers are able to automatically extract, analyse and understand the useful information from the raw data (Chauhan, N. K., & Singh, K. (2018)).

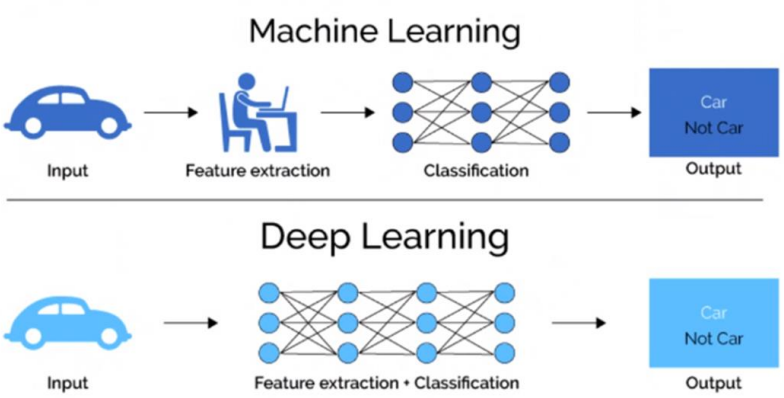


Figure 4. Differences between Machine Learning and Deep Learning.

Feature engineering in DL is the process of putting domain knowledge into the creation of feature extraction to reduce the complexity of the data and make patterns more visible to learning algorithms. The process in DL requires more resources - budget, time and data - than in ML (Lawtomated. (2019, April 18)).

Generative AI (GenAI) is a specific subset of Machine Learning (ML) technologies, which are able to create new content based on Large Language Models (LLMs), neural network–based models trained on massive amounts of data, including text and documents.

The LLMs are at the core of the technology applied to the legal discipline. The friendly interface with the user through Chat Generative Pre-Trained Transformer platforms, such as ChatGPT, has gained significant adoption. Launched on November 30th 2022 by Open AI, ChatGPT-3.5 gained more than 100 million active users across the globe at a much faster pace than any other innovation platform to

that date. ChatGPT-4 was launched in March 2023, managing 175 billion parameters performing Natural Language Processing (NLP) tasks, including question and answering, text summarization and machine translation.

Currently, data analytics research and AI initiatives are embedded into the public or private domains. Research institutions focus on the concept of "access to justice" with legal assistance tools that aid the public to identify or navigate legal information, provide assistance with questions, contract analysis, legal document generation, and outcome prediction[1].

The legal profession have used analytics to search and research, extract data, and beyond in terms of writing memos or depositions, and developing predictive tools capable of determining how courts will rule including the odds of winning a case (Dahan, S. (2020)).

Some of the abovementioned solutions and tools are summarized[2] in Table 11 below:

[1] Some examples of the former are CodeX (Standford Law), Cyberjustice (University of Montreal), SMART Law (HEC Paris) and the Conflict Analytics Lab (CAL at Queen's University).
[2] List is not exhaustive.

Dahan, S. (2020) *LexMachina*[3](´LexisNexis' group) *Blue J Legal*[4] *Ross*[5] Luk, M. (2023) ChatGPT-4 based *Casetext*[6] *Gracenote*[7] *Harvey*[8] Anthropic´s Cloude based *RobinAi*[9]	Ortolani, P., Janssen, A., & Wolters, P. (Eds.). (2022) *MyOpenCourt*[10] *Solomonic*[11] *Docket Alarm*[12] *Docket Navigator*[13] *Monitor Suite and Westlaw Edge*[14] *Arbilex*[15] Kohlmeier, T. (2017) *Elevate.law*[16] *Predictice.com*[17] (civil law)

Table 11. Selection of Commercial Legal Solutions that incorporate new technologies into the Legal Domain.

This thesis advocates that technology and new tools for data extraction and analysis could contribute to bridge the gap between the legal and financial disciplines.

Further to research by Park, S. H et al (2021), Legal Tech, defined as modern technologies and IT solutions that can be used to provide some types of legal services, has emerged as an important research topic for the legal and IT industries.

[3] https://lexmachina.com/ last access in March 2024.
[4] https://www.bluej.com last access in March 2024.
[5] https://www.rossintelligence.com last access in March 2024.
[6] https://casetext.com/cocounsel/ last access in March 2024.
[7] https://gracenote.ai last access in March 2024.
[8] https://www.harvey.ai last access in March 2024.
[9] https://www.robinai.com last access in March 2024.
[10] https://myopencourt.org last access in March 2024.
[11] https://www.solomonic.co.uk last access in March 2024.
[12] https://www.docketalarm.com/ last access in March 2024.
[13] https://brochure.docketnavigator.com/ last access in March 2024.
[14] https://legal.thomsonreuters.com/en/products/monitor-suite and https://legal.thomsonreuters.com/en/products/westlaw-edge, last access in March 2024.
[15] https://www.arbilex.co/welcome last access in March 2024.
[16] https://elevate.law/ last access in March, 2024.
[17] https://predictice.com/ last accessed in March, 2024.

Data analytics can be defined as a procedure of creating value by processing, analysing, and interpreting raw data through approaches and techniques such as AI, Machine Learning (ML) and data mining. Data analytics legal tech has emerged as a research topic, where, according to their view, half of the relevant papers have been published by a first author from Asia, followed by studies from Europe (31%) and North America (11%).

China´s pole position in these new technologies should be highlighted: the world's first ´Internet Court´ was established in Hangzhou in August 2017, followed by two additional ones in Beijing and Guangzhou in 2018 (Guo, M. (2021)).

Legal Artificial Intelligence focuses on applying especially natural language processing (NLP) to benefit tasks in the legal domain.

According to Zhong et al. (2020), legal professionals often think about how to solve tasks from rule-based and symbol-based methods, while NLP researchers concentrate more on data-driven and embedding methods. Since retrieving and understanding legal documents take time, ML and DL techniques and neural models are being applied to Legal AI. "Deep learning and neural networks excel at exactly the tasks that symbolic AI struggles with" according to Ben Dickson (2019). In other words, whilst symbolic AI processes symbols inferring and deducting concepts and their relationships from rules and axioms, the great advantage of neural networks is that they can induce from unstructured data.

The majority of the resources in the legal field are presented in text forms, such as judgment documents, contracts, and legal opinions. Therefore, most Legal AI tasks are based on NLP technologies.

The Eruption of Generative Artificial Intelligence

Shabsigh, G., & Boukherouaa, E. B. (2023) comprise Generative Artificial Intelligence (GenAI) within a specific subset of AI–Machine Learning (AI/ML) technologies, distinguished by their ability to create new content.

At the heart of GenAI are Large Language Models (´**LLMs**´), which are neural network–based models trained on massive amounts of data, including text and documents, and capable of producing understandable and meaningful text or human languages.

LLMs enable a wide range of applications across sectors, including legal. LLMs are machine learning models that are good at understanding questions or requests and generating human language. Those models operate by ingesting vast quantities of data, for training purposes, to discern statistical patterns, including the relationships between words and the contextual significance of each word within a sentence. With this knowledge, the models can predict word sequences sequentially, one word at a time.

The key discovery of the LLM was the transformer architecture that was introduced by Vaswani and others (2017), their key innovation being the introduction of the self-attention feature. This mechanism allows the model to select the key words in the input to pay attention to and deem relevant rather than using the entire input equally.

Transformers are a newer and more powerful type of neural network architecture. Designed to process sequential data without using recurrent connections, they use an attention mechanism to learn the relationships between parts of the input sequence. This makes them more efficient and easier to train than Recurrent Neural networks (RNNs). Transformers are better at handling long sequences of data and are also well suited for a variety of NLP tasks. They are easy to scale and easier to train than RNNs.

The computer does not understand words or text; hence, all input words must be converted to vectors before the computer can perform all statistical patterns and mathematical modulization of each step.

The limitations are also highlighted by the same authors, *inter alia*: understanding context, lack of common sense, training data could be biased, control and safety concerns, significant computational resources, inability to verify facts and hallucination (models invent incorrect statements). Potential solutions comprise finetuning and prompt engineering, and integration with knowledge graphs.

Release date	Company	LLM Platform	No. Parameters (billions)	Performance GLUE score*	Main Application
October 2018	Google AI	BERT	0.34	86.5	NLP
June 2022	Google AI	Jurassic-1 Jumbo	1.75	94	NLP
June 2022	Beijin Academy of AI	WuDao 2.0	1.75	94.2	NLP
November 2022	Open AI	GPT-3	0.175	80.3	NLP
March 2023	Open AI	GPT-4	0.175	93.3	
March 2023	Anthropic	Claude	0.137	92.2	
June 2023	Inflection	Pi	0.137	92.8	

** benchmark for natural language understanding*

Table 12. LLM Platforms (NLP processing and understanding). Shabsigh, G., & Boukherouaa, E. B. (2023)

Other LLM platforms beyond Open AI to highlight are Wu-Dao 2.0 (Beijing Academy of Artificial Intelligence) and Jurassic-1 Jumbo by Google Ai, both of them released in June 2022 and managing 175 trillion parameters. The latter substituted prior LLM BERT released in October 2018, which managed 380 million of parameters. Lai, J. et al (2023) have provided an overview of the large language models, specifically in the Legal discipline.

Size and quality of training data, along with fine-tuning, are differentiating factors, although Chinchilla (from Google´s DeepMind) showed that a smaller model trained on more data could outperform significantly larger models trained on less data. In other words, there is an optimal ratio between model size and training. The ecosystem we are facing are powerful multimodal models with general tool-using capabilities that can connect with several API platform of tools to perform multiple tasks (Luk, M. (2023)).

A software bot is an interface that connects users to services (Lebeuf, C. R. (2018)). In the lexicon (Lexico Dictionaries. (n.d.)), a chatbot is defined as “a computer program designed to simulate conversation with human users, especially over the Internet” (Adamopoulou, E., & Moussiades, L. (2020)).

Chat Generative Pre-Trained (GPT) has been released and introduced by OpenAI in November 2022. The primary goal of designing ChatGPT is to generate natural language text for a wide range of applications (Kashefi, A., & Mukerji, T. (2023))

ChatGPT is helping the academia. Pre-trained Large Language Models have the potential to enhance science research and academia through inter alia evidence review, synthesis, text classification, content creation, information extraction and assisted programming.

Following (Panda, Subhajit (2023)), ChatGPT Application Programming Interface (API) uses natural language processing (NLP) to understand user queries and provide accurate responses. A common format for digital resources is the Portable Document Format (PDF). PDFs allow for easy distribution of information and can preserve the layout and formatting of a document, making them popular for academic articles, research reports, and other types of content. However, one common problem with PDFs in library systems is their limited interaction capabilities. ChatGPT API enhances the PDF interaction tool ChatPDF by enabling natural language processing, context-awareness, and easy annotation capabilities.

According to ChatPDF website "*In the analyzing step, ChatPDF creates a semantic index over all paragraphs of the PDF. When answering a question, ChatPDF finds the most relevant paragraphs from the PDF and uses the ChatGPT API from OpenAI to generate an answer"* (ChatPDF. (n.d.)).

Although academia agree on the potential, the latter is somehow limited, at least at present: GPT-4 can perform at a level comparable to well-trained law student annotators which could reduce costs (Savelka, J., Ashley, K. D. et al (2023)). Savelka, J. et al (2023) found that direct application of GPT-4 supplies high quality explanations on the surface, however detailed analysis uncovers the limitations in terms of factual accuracy, augmentation leading to improved quality eliminating the issue of hallucination. Tan, J. et al (2023) concur that "*ChatGPT is not yet accurate enough to provide legal information directly to laypeople"* in the ´access to justice´ goal.

Intelligence Assistance (Schilder, F. (2023)) and experimental platforms are progressing in this front i.a. Westermann, H. et al (2023) in connection with Online Dispute Resolution; Guha, N. et al (2023) assessing how several LLMs interact with different type of tasks and questions); and the extensive survey on LLMs by Wang (C., Liu, X. et al (2023)).

The assessment of the legal capabilities have been addressed by Fei, Z. et al (2023) benchmarking multilingual, Chinese or law-specific LLMs oriented on memorization, understanding, application of legal knowledge. Their challenges are well pointed out by Kaddour, J. et al (2023).

The use and the opportunity of using AI in online arbitration has been flagged in the past not only in the process of reaching an award, but also on the drafting of the content of the award itself – see for instance Paisley, K., & Sussman, E. (2018), Rhim, Y. Y. et al (2023), Shalaby, A. G. (2023) and Hussain, M. A. (2023).

We will focus in the next on how AI has been used for predicting legal judgments, aimed by the stakeholders to take informed decisions.

Legal Judgment prediction

Legal information is written in a mostly unstructured and specific language. In order to build high quality legal decision predictions, the methodologies mostly relate to language analysis and (automatic) extraction features from legal texts, taking advantage of Natural Language Processing as well as (supervised) Machine Learning and/or Deep Learning (neuronal networks) techniques.

According to Alexander, C. S. (2023), litigation outcome prediction is an active scholarly research area with a variety of data sets, modelling approaches and performance measures. It is worth sharing some of her conclusions in the following paragraphs. Citing Kevin Ashley (2019), she traces back the history of the field to the work of two academics (Mackaay, E., & Robillard, P. (1974)) who used a machine learning algorithm called k-nearest neighbours in the 1970s to forecast the outcome of Canadian real estate tax disputes. In the United States, academic interest has focused on decisions by the US Supreme Court, federal appellate courts, federal district courts, immigration court, state trial courts and administrative agencies. Case types studied include employment, asylum, tort and vehicular, and trade secret misappropriation. Other scholars outside the United States have developed outcome prediction tools focused on the European Court of Human Rights, the International Criminal Court, French appeals courts, the Supreme Court of the Philippines, lending cases in China, labour cases in Brazil, public morality and freedom of expression cases in Turkey's Constitutional Court, and Canadian employment and tax cases.

Charlotte Alexander (2023) also points out the limitations in academic methods to predicting court rulings: (i) relying on court descriptions of case facts and legal citations, which may be biased as they are often written by judges or their clerks with prior knowledge of the ruling, leading to predictable outcomes and undermining the validity of the predictions; (ii) limited ability to generalise, due to the narrow selection of cases used for training and testing predictive models, mainly due to inaccessible or missing court data and (iii) prioritization of prediction over explanation; despite a growing trend towards explainable predictions, where

interpretations are provided, achieving explainability remains a challenge: *"the predictive jump is left to the user, who decides whether to adopt the approximation as a prediction or to distinguish it from the case at hand*" (Alexander, C. S. (2023).

We have selected a list of precedents in legal analytics, summarized below.

Year	Author	Sample size (#)		Topic or sentences source
2024	Pereira, J. et al (2024)	122	Brazil	Brazilian Audit Courts
2024	Sun, J., Huang, S., & Wei, C. (2024)	241.434	China	Chinese Legal Judgement Prediction competition dataset
2023	Schilder, F. (2023)	17.215	France	French Court of Appeal
2023	Shalaby, A. G. (2023)	7.482	China	Private lending
2023	Sherry Xin Chen (2019)	538	Canada	Unemployment law (worker or contractor status)
2022	Lage-Freitas, A. et al (2022)	4.403	Brazil	Brazilian Court decisions
2022	Sert, M. F. et al (2022)	338	Turkey	Turkish Constitutional Court (public morality)
2022	Sokhansanj, B. A., & Rosen, G. L. (2022)	10.462	US	PTAB (Patent Trial & Appeal Board)
2021	Niklaus, J. et al (2021)	85.000	Switzerland	Federal Supreme Court (trilingual corpus)
2021	Mumcuoğlu, E. (2022)	93.340	Turkey	Turkish Constitutional Court and Courts of Appeal
2021	Rhim, Y. Y. et al (2023)	184.125	US	State of Connecticut Judicial branch civil cases
2021	Savelka, J. et al (2023)	16.024	Int	WIPO cases
2020	Bertalan, V. G. F., & Ruiz, E. E. S. (2020)	782	Brazil	Sao Paolo State Higher Court
2020	Medvedeva, M., Vols, M., & Wieling, M. (2020)	1.942	Europe	European Convention oF Human Rights
2020	Strickson, B., & De La Iglesia, B. (2020)	4.959	UK	UK Court judgments
2019	Chalkidis, I. et al (2019)	11.500	Europe	European Convention of Human Rights
2019	Li, S. et al (2019)	1.367.654	China	Supreme People´s Court of China (criminal cases)
2019	Yang, W. et al (2019)	185.228	China	Supreme People´s Court of China
2018	Barros, R. et al (2018)	10.000	Brazil	Brazilian labour court
2018	Elnaggar, A. et al (2018)	20.000	Germany	Eur-Lex corpus
2018	Hu, Z. et al (2018)	383.697	China	China Judgments online (charge prediction criminal cases)
2018	Kowsrihawat, K. et al (2018)	1.207	Thailand	Thai Supreme Court
2018	Long, S. et al (2019)	100.000	China	Divorce decisions China Judgements online
2018	Virtucio, M. B. L. et al (2018)	8.132	Philippines	Philippines Supreme Court decisions (criminal cases)
2018	Zhong, H. et al (2018)	113.536	China	Supreme People´s Court (charge prediction criminal cases)
2017	Katz, D. M. et al (2017)	28.084	US	US Supreme Court (case outcome)
2017	Popp, A. T. (2019).	254.021	US	Asylum court decisions
2017	Shulayeva, O. (2017)	50	UK&I	Principles & facts citations British & Irish Legal Institute
2017	Sulea, O. M. et al (2017)	126.865	France	French Supreme Court (case ruling, law area and timing)
2016	Aletras, N. et al (2016)	584	Europe	European Convention oF Human Rights
2012	McShane, B.B. et al (2012)	5.898	US	Securities fraud litigation
2012	Pérez López, J. Á.et al (2012)	864	US	US Federal employment cases

Table 13. Legal Judgment Prediction. Snapshot. Note: This table was created by the author based on a review of the legal judgment prediction literature.

Further detail can be found in Appendix 1 (Table 37), where the attempts from different authors, jurisdictions (i.a. US, China, Europe and LatAm) and approaches (Machine learning or a combination of ML/DL and sample sizes) can be seen. Some of this research has spun off into commercial products, discussed in the next section.

For instance, Medvedeva, M. et al (2020) concluded in their study that using information obtained in a simple and automatic manner, their models are able to predict decisions correctly in about 75% of the cases, improving the chance

performance of 50%. Crucially, they emphasise that the discussion of predicting judicial decisions verses solely to the available data and methodology, not claiming the ability to predict a new case.

Branting, L. K. et al (2021) and Dias, J. et al (2022) remind that data-centric approaches have in general lower transparency and explanatory capability. Researchers working on AI and Law recognise that it would be necessary to justify the predictions in terms that could be understood by lawyers and laypeople (Atkinson, K., Bench-Capon, T., & Bollegala, D. (2020)).

The praxis has suffered a disruptive shock from the accessible and available pre-trained Large Language models (LLMs) such as GPT, causing a shift in data mining and ML. Retrieval-Augmented Generation (RAG) models, which combine pre-trained parametric and non-parametric memory for language generation, are being architected and continuously improved through prompting for the purpose of prediction (see for instance i.a. Lewis, P. et al (2020), Wei, J et al (2022) and Liu, P. et al (2023)). Legal reasoning and legal prompt engineering have emerged to the point of passing demanding pass exams in US (with 70% score) or Japan (Zhang, D. et al (2022)). Reinforcement learning from human feedback (R–HF) - passive and active/proactive is increasing these capabilities.

When selecting the features, Branting, L. K. et al (2021) concluded that if the induced (i.e. predicting) tags could be accurately projected *from (manually) annotated cases onto the remainder of the cases*, the resulting corpus could be useful for prediction and explanation. We once more agree with Charlotte Alexander (2023) that no single dominant approach has emerged, whilst direct comparison among studies can be difficult given different datasets and performance measures. We have attested that “predictive performance ranges from relatively modest marginal classification accuracy to high F1 scores”. The challenges and inconveniences flagged by the reviewers are aimed to be faced or mitigated in this research, avoiding ex ante self-discovery, and enhancing transparency and explainability.

The existing state-of-the-art methods in legal judgment prediction (see Appendix 1 for the detail) use several machine learning or deep learning algorithms to *classify* the existing selected judicial corpa, aiming to find out whether the decisions are positive or negative. Techniques vary from language models based on artificial neural networks random forest classifier or regression trees using supervised machine learning algorithms – applied to facts and charges, or to find out whether the decision was unanimous or not. We simply observe what are the characteristics of the decisions (awards) that have been positive in terms of alleged and found breaches. The same could be undertaken for negative awards. We deep dive into the existing information and apply econometrics to find out trends, conscious that past behaviour is not guarantee of future performance.

Relevance of this research´s contribution vis-à-vis state of the art

To our knowledge, there is no standard methodology to approach fair valuation in litigation risk; we propose to reduce the subjectivity based on the analysis of historical legal and economic data, leveraging on the tradition of linking economic analysis and law (Postner, R. A. (2014)), Providing a framework through interdisciplinary studies for rational decision making when considering disputes may contribute to an efficient pursuit of a lawsuit and the allocation of financial resources (Merlone, U., & Lupano, M. (2022)).

As pointed out, substantive work has been directed towards the probability of success of the outcome. Within the Expected Value (EV) of a claim formula, instead of focusing on the probability of success, we have further analysed, based on historical data, the quantum that a claimant would obtain within the investment arbitration domain, and how long would it take to reach that award. Ortolani et al. (Eds.). (2022) anticipate the potential of technology, which could be applied to the amount of compensation claimed and awarded in similar past cases.

The importance ranges from financial planning, decision making management as well as a risk and portfolio management perspectives.

The beneficiaries are claimants, litigation funders and their investors, as well as other stakeholders from law firms, insurers. To provide additional tools to assess litigation risk and its financial consequences can also help policy makers to architect a more efficient judicial system and preserve access to justice with less and meritorious cases.

Data Mining, Text Analytics and Investor-State Arbitration

From sentiment analysis to synthetic data generation and automatic or robo-advisory, quantitative techniques, leveraged by new technologies, are becoming increasingly important for asset managers in finance (Luk, M. (2023), JP Morgan (Kolanovic, M., & Krishnamachari, R. (2017)).

In principle, the legal community is sceptical about the lack of human element in decision-making in moral or ethical dilemmas, although a gradual use of AI assisted arbitration could be regulated (Waqar, M. (2022)).

Traditionally, legal texts have been analysed qualitatively or, more recently, have been converted to data for purposes of quantitative analysis manually through the process of hand coding. This treating “law as data” has been studied in the past by, inter alia, Rissland et al. (2003) and Bench-Capon et al. (2012).

We have followed closely Franck, S. D., & Wylie, L. E. (2015) who, addressing ITA-law files supplemented and checked against other resources including the UNCTAD and ICSID websites, extracted information from over 500 awards out of approximately 1000 cases in the database. Awards are defined as elements of cases that resolve dispositive key legal issues—under the ICSID or New York Convention, or the UNCITRAL Model Law - creating a dispositive and binding determination on a substantive phase, namely decisions on (1) jurisdiction; (2) merits; (3) quantum of damages; (4) allocation of costs; or (5) settlement agreements or other orders indicating a dismissal or discontinuance. They remind that the objective of an IIA is to entice inbound foreign investment and to protect a

state's own investors abroad while minimising the risk of state liability, promising foreign investors that they will receive certain basic treatment, including the right to freedom from expropriation without proper compensation, from discrimination, and guarantees of fair and equitable treatment. If a tribunal awarded an investor at least USD 1, it is considered an "investor win" (claimant) or the tribunal failed to award damages to the investor (i.e. an award of USD 0), namely a "respondent win". They also measure examined investors' "relative success rate", or the percentage awarded to investors relative to their asserted claim; higher values reflecting higher levels of success. They coded over 200 of the above 500 cases, most recent amounts of awards taken into account, with adjustment to currency (USD) and inflation. They explore a multiple regression approach for their analysis taking into account "extra-legal factors", such as arbitrator or venue, or identity of the investor – whether individual or company – or the experience of the legal counsel.

Marrow et al (2020) consider arbitration as a suitable candidate for AI driven programs. Anecdotally, stylometry has been studied in order to predict the authorship of investment arbitration awards (Langford, M., Behn, D., & Lie, R. (2020)). Paisley, K., & Sussman, E. (2018) anticipated that AI offers the potential of predicting results in advance, including chances of success, range of damages, timings and costs. The stakeholders in arbitration "aim to make a more detailed calculation of the return of investment on the time and cost relating to initiating the arbitration" (Morgan, C. (2020)).

Frankenreiter, J., & Livermore, M. A. (2020), point out that "*certain techniques condense textual information into one or a small number of variables and then carry out more familiar statistical analyses based on the transformed, lower-dimensional data*". In their overview, they start with *"applications of law as data in studies that focus on quantitatively investigating causal relationships using tools such as regression analysis, with numerical representations of the law typically feature as independent variables in a regression"*.

Wilinski, P., & Durbas, M. (2021) point out companies such as Arbilex[18] that operate in jurisdictions (particularly in the US) where the data on which predictions are based is public, stable, and predictable. These services in the arbitration market provide tools quantifying case risks using machine learning and game theory inspired models. They also remind there are drawbacks, mindful that one of the pioneers, CourtQuant, failed recently due to *"the conservative nature of the legal sector, including litigation funders, when it came to using predictive techniques; and the tough market for legal tech startup funding during the current pandemic*" (Artificial Lawyer (2020, October 7))." A recent drawback has also been suffered by Gavelytics (Legaltech News (2022, June 29)).

Following Alschner and Charlotin (2018) cit. by Ortolani, P., Janssen, A., & Wolters, P. (Eds.). (2022), AI can be used from online case management platforms to the application of data mining techniques to arbitral awards, from the use of smart contracts automating the enforcement of arbitral awards to the electronic processing of mass arbitration claims. Legal tech appears as an "exogenous shock", offering new tools to the arbitration community. For instance, *"technology could be applied to (…) the amount of compensation claimed and awarded in similar past cases. A predictive machine learning algorithm can then be trained to create a relationship between case characteristics and outcomes. While traditionally information about a case needed to be extracted by human analysts, today's data mining methods allow extracting at least some of that information as metadata directly from the full text of decisions".* Automated text extraction methods benefit from the broadly similar formatting of an arbitration decision: introductory part, factual and procedural backgrounds, the parties' arguments, the tribunal's reasons, and the conclusion. Algorithms can adapt to the different formatting conventions to extract possible sets of relevant metadata. Moreover, they add, several academic, governmental and commercial databases have collected metadata on awards through laborious manual efforts.

[18] https://www.arbilex.co/ last accessed in March, 2024.

To remark, amongst the forward-looking statements of Burford Capital Financial Statements in 2022 and 2023 year-end, not present prior years, can be found "*the inaccuracy or failure of the probabilistic model and decision science tools including artificial intelligence (´AI´) tools, we use to predict the returns on our legal finance assets and in our operations"*.

In this research, (i) awards have been accessed from external websites (UNCTAD[19] redirecting to Italaw[20] , *Jusmundi*[21], *Investment Arbitration Reporter*[22] and ***Kluwer Arbitration***[23] database licensed from November 2022 to November 2023); (ii) the relationship between case characteristics - legal and financial features - and its impact for a positive outcome has been studied, agnostic of exogenous variables such as venue, sector or Tribunal composition; (iii) statistics and econometry has been applied to the relationship between amount claimed and amount awarded - since computational tools have not been as fruitful as expected for data gathering and (iv) a methodology for a probabilistic model has been proposed – a predictive algorithm without the machine learning element, enhancing explainability.

Journals

Three journals to highlight in Legal Tech discipline: two academic peer-reviewed journals of reference are ***Journal of Empirical studies*** (Wiley Online Library. (n.d.) and ***Jurimetrics*** (American Bar Association. (n.d.)). Jurimetrics is the application of quantitative methods, specially statistics and probability, to law (Loevinger, 1963).

This the official journal of the American Bar Association Science & Technology Law Section is published quarterly, and it is the oldest and most widely circulated peer-reviewed journal in its field.

[19] https://investment-policy.unctad.org/investment-dispute-settlement last access in March, 2024.
[20] https://italaw.com/ last access in March, 2024.
[21] https://jusmundi.com/ last access in March, 2024.
[22] https://www.iareporter.com/ last access in March, 2024.
[23] https://www.wolterskluwer.com/en/solutions/kluwerarbitration last access in March, 2024.

A third journal, ***Artificial Intelligence and Law*** (Springer. (n.d.)), seeks to address "the development of formal or computational models of legal knowledge, reasoning, and decision making".

To our knowledge, we have not encountered a peer-reviewed article that includes a methodology for predicting, based on past information, the expected amount and the years to an award or, more broadly speaking, a judgment.

3. Proposed methodology – Legal Finance Analytics applied to Third Party Funding (TPF) in International Investment Disputes

After reviewing the academic state-of-the-art and a publicly listed litigation finance practitioner, we will deep dive into Investor-State International Arbitration Disputes of our research case.

Through Legal Finance Analytics, in this research, it is aimed to find out the relationship between the econometric and legal features of an award.

We have purposely not focused on variables such as venue of the arbitration, the composition of the tribunal, or the sector of the dispute. These are considered "exogenous" variables that have no influence on the outcome on our study. As opposed to said approach, we have explored whether there is a relationship between the merits of the case and the award obtained, in other words between the "legal variables" – formal and substantive breaches of an international investment agreement, both alleged and found - and the amount obtained in a positive outcome. We are therefore considering and analysing law - or judicial and more precisely arbitration precedent - as data.

We will remind some preliminary considerations on risk reward from a litigation finance practice: what is the difference between the multiple and risk-reward approach, and the importance of measuring risk-reward in terms of Multiple On Invested Capital (MOIC) and Internal Rate Return (IRR).

A- Preliminary Risk-Reward Considerations

Multiple vs Percentage Approach

Let´s assume, by way of example, the investment committee of a funder analyses a case with a USD 10 M of damages award or ´expected quantum´. In order to finance the case, a USD 1 million (´M´) investment is needed. This figure would represent the ´budget´, which could be broken down, indicatively and *ex ante*, as USD 0.7 M as expense *forecast* in legal fees, USD 0.2 M for damages´ experts fees and USD 0.1 M in-court fees and other fees and expenses (e.g. witnesses travels, translations, etc.). We will also assume that all the budget will be disbursed and no additional funding will be required, i.e. total commitment and disbursed amounts equal to USD 1 M.

Let´s assume said percentage is 40% and that the positive outcome for the claimant has been met, i.e. USD 10 M have been awarded as compensation. Each party would have paid its costs, so no additional amount has been granted beyond the compensation.

In principle, a funder would expect (i) a return of 3 times the investment, plus the recovery of the investment (´multiple approach´) or (ii) a percentage of the proceeds (´percentage approach´), *whichever is higher*. In the example, and for the sake of argument, both approaches would lead to the same share for the funder i.e., USD 4 M (USD 3 M plus recovery of USD 1 M) or USD 4 M (40% of USD 10 M).

The table below shows an order of magnitude in case of several scenarios of expected quantum: from USD 5M to USD 100 M for every USD 1 M, USD 5 M or USD 7 M of investment - budgets 1, 2 and 3 -, respectively.

Figures in USD M		5,0	10,0	50,0	100,0
Budget 1	**1,0**				
Investment plus a multiple of	3x	4,0	4,0	4,0	4,0
%	40%	2,0	4,0	20,0	40,0
Budget 2	**5,0**				
Investment plus a multiple of	3x	20,0	20,0	20,0	20,0
%	40%	2,0	4,0	20,0	40,0
Budget 3	**7,0**				
Investment plus a multiple of	3x	28,0	28,0	28,0	28,0
%	40%	2,0	4,0	20,0	30,0

Table 14.A Expected reward of a TPF (in USD M) using the multiple or the percentage approach in case of a USD 1 M, USD 5 M or USD 7 M investment and a USD 5 M, USD 10 M, USD 50 M or USD 100 M quantum. Own data.

When the quantum is high, the multiple approach would benefit the claimant in detriment of the funder. In case of low damages award, the multiple approach would harm the claimant vis-à-vis the funder. These scenarios are referred to in the outlined boxes.

Consequently, claimants would prefer the percentage approach to the multiples approach when the compensation is low. In order to incentivise the funder in awards where high damages awards have been granted, an additional mechanism can be sought to share the upside, for instance, a percentage over the proceeds, should the award exceed a certain threshold.

Example above considers a 3x plus recovery of investment or 4x. As outlined in the risk-reward section, that multiple can vary, upwards or downwards. The importance of using one or another could be very relevant for IRR purposes, when the value of money in time is considered.

Variations can apply in practice (e.g. a grid where returns are expressed in multiples or percentages increasing over time). As a ´rule of thumb´, a funder would expect a minimum of ten million in damages for every one million of investment (1:10 rule).

Caps and floors - e.g. to ensure at least the recovery of the investment - can apply depending on the bargaining power of the parties and the enforceable legislation in the jurisdiction of reference.

According to **Burford Capital, *"Return on Invested Capital ("ROIC")*** *from a concluded asset is the absolute amount of realizations from such asset in excess of the amount of expenditure incurred in funding such asset divided by the amount of expenditure incurred, expressed as a percentage figure. ROIC is a measure of our ability to generate absolute returns on our assets.* ***Some industry participants express returns on a multiple of invested capital ("MOIC") instead of a ROIC basis.*** *MOIC includes the return of capital and, therefore, is 1x higher than ROIC. In other words, 70% ROIC is the same as 1.70x MOIC*" (2022 Burford Capital Annual Report, page 9).

Internal Rate of Return

The above table does not take the value of money into account, which enters into play with the concept of IRR.

We take from Burford Capital (2002) the definition of **Internal Rate of Return (IRR),** defined as "*a discount rate that makes the net present value of a series of cash flows equal to zero and is expressed as a percentage figure".* They add: *"We compute IRR on concluded (including partially concluded) legal finance assets by treating that entire portfolio (or, when noted, a subset thereof) as one undifferentiated pool of capital and measuring actual and, if necessary, estimated inflows and outflows from that pool, allocating costs appropriately. IRRs do not include unrealized gains or losses."*

Time, a critical element when calculating IRR, also experiences a certain degree of estimation when addressing realisations. They also define Weighted Average Life ("WAL") of a legal finance asset the one that "*represents the average length of time from deployment and/or cash outlay until we receive a cash realization (actual or, if necessary, estimated) from that asset weighted by the amount of that realization or deployment, as applicable. In other words, WAL is how long our asset is outstanding on average."*

Table below shows the IRR for a Third-Party Funder (TPF) per USD 1 M investment, with the following three assumptions: (a) the *expected quantum* will not be reached, i.e. the amount awarded will be lower than the amount of the claim requested,

suffering a discount or ´haircut´ (the ´hit ratio´ would equal to 50%), (b) a TPF assumes 100% of the budget - investment - and receives a compensation of 40%; and (c) according to the funder´s criteria, an investment would take place if and only if the expected IRR is above 20%.

USD M Investment	Amount Claimed	Haircut %	Expected Quantum	Claimant(s) share	TPF share	TPF return (x) Cash on cash	IRR (%) 3 years	6 years	8 years	10 years
1	100	50%	50	30	20	20	171%	65%	45%	35%
1	50	50%	25	15	10	10	115%	47%	33%	26%
1	20	50%	10	6	4	4	59%	26%	19%	15%
1	16	50%	8	4,8	3,2	3,2	47%	21%	16%	12%
1	12	50%	6	3,6	2,4	2,4	34%	16%	12%	9%
1	8	50%	4	2,4	1,6	1,6	17%	8%	6%	5%
1	5	50%	2,5	1,5	1	1	0%	0%	0%	0%
1	2	50%	1	0,6	0,4	0,4	0%	0%	0%	0%

Table 14.B Expected IRR scenarios of a TPF for a 3, 6, 8 and 10 years time horizon (columns) with a varying expected Quantum (rows). Own data.

The chart above shows expectations *prior* to investment, that is, the *actual* award could be below the expected award. According to the table and assumption (c), the decision to invest *ex ante* (*id est*, before commitment takes place) would not have taken place in the area with red numbers, reflecting the importance of expected quantum and time to cash.

Value of money normally decreases over time, and so does IRR. This reduction is exponential; in simple terms, doubling the time-to-award from investment would represent a more than double decrease in the IRR. In some instances - e.g. with a 3.2x cash-on-cash return expected by the funder - the criteria would only be met if the time span to actual reward is below a certain threshold; in the example as shown in the box, beyond 6 years. This is highlighted in the box.

In case an additional criterium was introduced for a TPF to invest, e.g. that both a 20% IRR *and* a time to return below 6 years needs to be met – the additional four

scenarios, at the right top corner of the table, also highlighted in a box, would not be met and would be also red-coloured, despite IRR above the threshold.

Meeting the criteria of a funder for a particular claim would lead to an ´investment grade´ status. The uncertain time-to-award, *inter alia*, is the basis to assert that litigation finance is an uncorrelated asset vis-à-vis stocks, bonds or real estate; as insurance risks are uncorrelated to other market risks (Molot, (2009)).

The "Hit Ratio"

The example in the above table considers this ratio as 50% for illustrative purposes. This research on international arbitration analyses, based on empirical evidence, that difference between the amount originally claimed and finally obtained.

The Hit Ratio can be expressed as follows:

Hit ratio = Amount obtained / Amount claimed

Table 15. The hit ratio formula, calculated as the ratio of the amount obtained to the amount claimed, refers to the principal amounts only. Own data.

The amount claimed is normally above the amount obtained, when referred to the principal amount requested, i.e. excluding both interests and, if awarded to the claimant, also the costs to be reimbursed by the defendant.

The "hit ratio" captures, in our view, relevant information: a high ratio mean that the amount requested was too optimistic, too high, or that only a small portion of what was requested was deemed to be fair by the Tribunal, or to both. The rationale for the former being too high could mean that the ´lost profit´ (*lucro cesante*) element was too high versus actual damages.

Calibrating Risks and Rewards

TPF is a useful tool to improve access to justice and through which the funded party can better allocate its financial resources. In principle, from a financial perspective, the interests of claimant and Third-Party Funder are aligned (Merlone, U., & Lupano, M. (2022)). This does not mean that they may have different perceptions of risk going forward rather than, within a rational behaviour, both would prefer having lower outflows or investments and higher inflows or rewards, and sooner rather than later.

If the TPF, instead of receiving 40% of the award, received a multiple over its USD 1M investment, several scenarios would need to be addressed in the documentation of the Litigation Finance Agreement, for instance: if the actual amount is below USD 1M and the funder – is the funder still entitled to recover 100% of the investment ?; if the actual amount is just above USD 1M and the funder was entitled to recover 100% of the investment - how should the excess over the investment be recovered in the priority or waterfall (50-50% or differently) ?; if the expected amount is well above let´s say the 3x plus investment recovery, could the funder request a % on the excess to share the upside ?

In practice, the bargaining power of the parties, claimant and funder, play a significant role to solve these questions. The return figure could also be converted from a multiple to a percentage of the actual proceeds.

The percentage is becoming more common as it facilitates to address floors - minimum guaranteed return - and caps – maximum potential return – scenarios. In fact, EU draft TPLF legislation limits the funder´s award to 40% of the proceeds on the higher end of the award, whilst does not preclude to receive that same percentage from a lower amount. According the Proposal for a Directive of the European Parliament and of the Council on the Regulation of Third-Party Litigation Funding, *"absent exceptional circumstances, where a litigation funding agreement would entitle a litigation funder to a share of any award that would dilute the share available to the claimant and the intended beneficiaries to 60% or less of the total award, including all damages amounts, costs, fees and other expenses, such an*

agreement shall have no legal effect". Albeit this cap to a funder´s remuneration is unprecedented and very relevant, the discussion around this damages award share cap would exceed the scope of this research.

In the above tables, it has been assumed that the amount committed equals the amount disbursed, and that there is no amount allocated to the premium to insure against the eventuality of losing or paying the other party´s legal costs. Which party assumes to pay for legal costs is relevant, taking into account that it would further erode IRR for the investor in case of losing. In America, each side usually pays its own legal costs. In Europe, the loser usually pays most of the winner´s legal costs (´each pays his own´ in the United States, versus European rule of ´loser pays all´ (Cooter, R. D., & Rubinfeld, D. L. (1989), Hederer, C. (2016)). These costs could be significant in case of multiple defendants.

In international arbitration and under ICSID rules, the tribunal has a high degree of discretion to allocate these costs (legal fees, fees and expenses of the arbitrators, administrative fees).

For the purposes of this research, we will consider the percentage approach for both downside and best-case scenarios from a funder´s perspective. The funder would still need to be vigilant on timing and its risk-reward, i.e. IRR considerations.

B- Proposed methodology applied to Investor-State International Arbitration Disputes

International Investment Agreements (IIA) disputes may be brought to (i) an international arbitration pursuant to the dispute resolution mechanisms in the IIA; (ii) a domestic court if the IIA is part of the relevant municipal law system; or (iii) a contractually agreed forum, in the case of an investor-state contract Newcombe, A. P., & Paradell, L. (2009). The two latter, domestic or contract IIA litigation, have been uncommon, because IIAs provide their own effective system of dispute resolution. According to these authors, the complexity of investor-state dispute resolution concern both jurisdictional issues i.e. scope of the arbitration agreement and procedure or law regulating the arbitration process and the validity and enforce-ability of the award; as well as to the substance of the dispute: content of the rights and obligations that the investor seeks to enforce.

This research focuses on international arbitration. The parties abide to an arbitration system and its rules, governed by the international Bilateral or International Treaty (BIT) or International Agreement (IA) – in some instances another international rule - intended to protect and promote investments. The respondent State foregoes ordinary jurisdiction rights towards the resolution of the dispute to a third party international organized body. The private party seeks a prompt and expeditive resolution, whilst the State assumes that an adverse decision of the Tribunal for the country could imply in some instances a significant amount in terms of compensation.

The Data Set

As per United Nations Conference on Trade and Development website (United Nations Conference on Trade and Development (n.d.-a)), with data as of 31 July 2023, out of the 1,303 known ISDS treaty-known cases, 357 are still pending a resolution, 22 are unknown and 924 have been concluded.

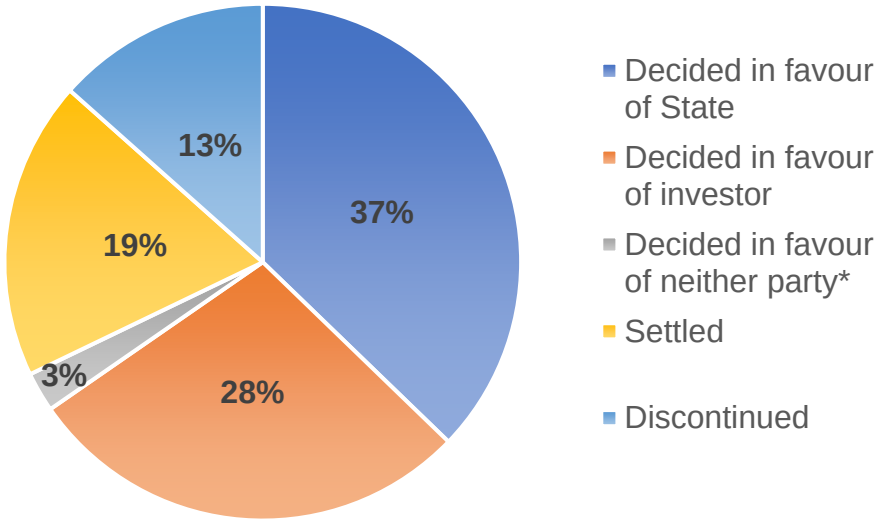


Figure 5- Breakdown of the 924 cases (in percentage) concluded proceedings as per UNICTAD (July 2023). United Nations Conference on Trade and Development. (n.d.-d)

	# cases	%
Decided in favour of Claimant	**433**	**47%**
Decided in favour of investor	260	28%
Settled	173	19%
Decided in favour of State	**491**	**53%**
Decided in favour of State	344	37%
Discontinued	124	13%
Decided in favour of neither party	23	2%
Total	**924**	**100%**

Table 16. Breakdown of the 924 cases concluded proceedings as per UNICTAD (July 2023). United Nations Conference on Trade and Development. (n.d.-d)

Out of the 924 known concluded cases, 344 have been favourable to the State (37%), whilst 282 were favourable to the claimant (31% of the total): 260 damages have been awarded and in additional 23 the State was found liable, but no damages were found, therefore siding, *de facto*, with the State.

From the available information online it cannot be inferred whether the 173 settled or 124 discontinued cases were or not favourable to the State or claimant. Having said this, in the table above it is assumed that settlements imply some kind of positive outcome for the claimant, whilst discontinued or decided in favour of neither party is not a positive outcome for the plaintiff who is seeking relief. By the definition of ´in favour of the claimant´ we follow the UNICTAD classification, whereby a monetary compensation has been granted to the claimant.

An example of the database lay-out can be found below:

Desisions rendered	
Award dated 18 August 2021	
Desisions rendered	
Claimed by investor	11.00 mln EUR (12.50 mln USD)
Awarded through settlement	15.50 mln USD

Table 17. Example of information extracted from UNICTAD website related to case (ICSID Case No. ARB/19/5) (United Nations Conference on Trade and Development. (n.d.-c).

Out of the 260 awards in favour of the claimant, extracting information from the website database, we have analysed a subsample of 94 awards (36%).

The awards have been picked randomly, being the driver those awards that have the information that is been sought for. A scorecard per award has been recorded, with the following information:

(i) The amount claimed or requested by the plaintiff at inception; where several figures have been asked for, for instance, when requesting a primary claim, auxiliary or subsequent requests have been placed in case the former is or are not granted, the smaller amount requested has been considered, where available. These above amounts normally refer to the main compensation figure, and also request a method to calculate the interest due.

(ii) The amount of interest added to the principal; when the calculation was possible and the information available, the figure is input; whenever an index and a spread over that index has been granted, an estimation has taken place based on publicly available information. The period

comprises from the start of the damages to the moment the award is disbursed; we have undertaken a few assumptions that will be elaborated later on.

(iii) The amount of costs, and whether these have been awarded to the claimant, should the respondent State have been obliged to reimburse the claimant.

Additionally, from a legal perspective, it has been accounted (YES / NO, being Y=1 and N=0):

(iv) Whether there were jurisdictional issues - time, investment or investor protection - precluding the claimant from continuing the dispute; as the awards within the sample relate to positive outcomes, none of these were adverse to the claimant.

(v) Has there been any kind of discrimination against the claimant: the respondent State not complying with a minimum standard, not supplying full protection and security, nor ensuring a fair and equitable treatment or denying international justice; breaching the Most-favoured Nation or the Umbrella Clauses.

(vi) Has there been a lack of compensation, that is, if the IIA breach at stake has led to some kind of expropriation, whether direct or indirect.

Finally, from a financial perspective:

(vii) What has been the amount received by the claimant, principal, interests and costs, when awarded, and what was the multiple on invested capital, i.e. the relationship between what has been received and the costs assumed by the claimant.

(viii) Should a funder had borne 100% of the expenses and received 40% of the proceeds, what would have been its IRR, assuming the timespan from the moment of inception of the damage to the time of the award. These assumptions will be considered throughout this research and in the limitations section of the research.

Award files have been downloaded – documents in pdf format - from three main databases: Italaw (International Arbitration Law - ITA.) and UNCTAD Investment Policy Hub, and Wolterskluwer arbitration module licensed from 1 November 2022 to 31 October, 2023. Data can be traced to both the awards and UNCTAD website and navigator (United Nations Conference on Trade and Development. (n.d.)).

The research has been undertaken in 2 phases:

Phase I – A limited sample (n=22) randomly chosen from the subset of 260 positive awards for the claimant. We have also analysed negative awards (n=22) of a subset of 344 awards negative to the claimant. The focus on the former has been driven the available information from a financial perspective.

The awards have been downloaded to ChatPDF and prompted through subsequent versions in order to obtain answers to the questions. Prompting is referred to in this research as rephrasing the questions, so as to improve the version of the template that provides improved output. The evaluation of the improvement has taken place through manual supervision.

The guiding principle of the search was exploring trends within the data points and inferring possible correlations between the different variables, knowing beforehand that the outcome would be positive for the claimant, as per the sample chosen.

Phase II – Once the output sought and the scorecards have been defined, the sample has been extended to 94 of the 260 awards positive to the claimant. The selection criteria has been random: wherever the final award included those variables that we wanted to complete, irrespective of the venue, tribunal composition, arbitration rules or IIA allegedly breached.

We have undertaken a test to whether the sample is or could be representative of the observed population.

From a technology perspective, in our research ChatPDF connected to ChatGPT-3.5 has been used to handle PDF documents. We also made a test with Ai-PDF plugged in Chat GPT-4.0. Since no material improvement was observed, in both

cases manual intervention and monitoring has been required given the lack of accuracy of the results, due to hallucination.

In this second phase, additional breaches, available from UNICTAD data base have been identified and introduced as potentially relevant, namely, the breach for umbrella clauses and most favoured nation.

The counterfactual scenario has been built with the above output, corresponding to 89 awards (down from 94) – whenever there was sufficient information: what would have been the position of a funder that bears 100% of the costs and receives 40% of the proceeds. We have also explored what % of the awards would have met a theoretical investment criterion of “IRR above 20% and time to award of six years or less”.

A regression model, based on statistically significative variables, has been designed to calculate the principal and years.

Finally, further discussion has taken place in terms of simulation analysis seeking to predict, based on historical information, what would have been the principal and years to award if two variables are known: the amount of claim requested and the costs, information accessible to the claimant through their main suppliers - law firm and damages expert.

Modelling

The International Centre for Settlement of Investment disputes (ICSID) is headquartered in Washington D.C., United Nations Commission on International Trade Law (UNCITRAL), a body of the United Nations General Assembly, in Vienna, whilst the Permanent Court of Arbitration (PCA) in The Hague. Their arbitration rules involve several phases and provide a flexible framework of rules to conduct the arbitration proceedings. Building a model has been helped by the relative standardisation of their governance.

In general, the awards include (1) a request for Arbitration including factual and legal background and relief sought by the claimant; (2) a response to the request by the defendant based on factual and legal arguments; (3) the constitution of the tribunal and appointment of arbitrators (normally three; two appointed by each party and a decision-making President); (4) the preliminary hearing to discuss procedure and initial issues; (5) exchange of statements and presentation of evidence and legal arguments; (6) hearings and presentation of evidence, closing arguments and any post hearing submissions and, finally, after considering evidence and arguments, (7) drafting the award with the merits of the dispute, any damages or relief awarded and allocation of costs.

Focus has been driven towards specific areas that, whilst being somehow standard across cases, could have a connection with a positive outcome. Following Franck, S. D., & Wylie, L. E. (2015), "*despite some literature focusing on jurisdictional decisions, little quantitative research has explored final outcomes or otherwise identified variables reliably associated with results. Identifying these factors, however, is vital to informed debates about the normative design of dispute resolution."* Their analysis based on regression models did not support claims that ITA outcomes were random. Rather, even with variation, "investor identity" and "counsel expertise" were the most reliable predictors of ITA outcomes. (…). They also conclude that *"focusing on evidence-based insights, relying on data, and minimizing emotive reactions that induce nonreplicable intuition would best serve the debate"*.

Following that thread, focus has taken place on the data captured within the awards by the Tribunal. This document is the conclusion of the Tribunal after perusal of numerous documents, hundreds or thousands of pages exchanged during an extensive period of time. We are assuming that the venue of the dispute, the composition of the tribunal, investor identity (whether a company or an individual) or the sector are irrelevant in our analysis.

The process undertaken in this research for modelling has been three-fold:

(i) Identifying the **legal features** in the sampled awards: the merits of the case have been targeted through the IIA breaches alleged *and* found by the Tribunal.

In a broad sense, an arbitrator or a judge considers how facts and evidence are embedded into applicable law. The Tribunal *"subsumes the facts within the rule" ("subsume el supuesto de hecho en la norma"*). In this statement, we construe "rule" in a broad sense, whether private contract, law and/or precedents. In common law systems, judicial precedents have a greater weight than in civil law systems, where codified laws and statutes are primary sources of law. This process is undertaken under the principle of *audiatur et altera pars*, the arbiter decides upon listening the arguments of both parties.

From a formal perspective the following questions are answered; has the claim been filed in due course? are the investor and investment protected under the relevant treaty? Once the jurisdictional matters have been cleared, the merits are considered further from a substantive law perspective: has the foreign investor been discriminated against and received fair and equitable treatment? has the foreign investor been compensated in case of expropriation?

(ii) Identifying the key **financial features** of the award. What have been the costs of the proceedings, what has been the amount of the damages award or compensation – both sought and obtained – including interest? have the costs been also be awarded?; how long has it taken until an award has been granted and relief obtained?

ChatPDF has been used to extract data from the international investment disputes awards under direct supervision. The UNCTAD website has also been addressed. To our knowledge it is the first time that prompting is being applied to investment arbitration and to litigation finance. Having said this, manual intervention has been necessary given the lack of accuracy of the output from ChatPDF. One of the key findings of our study from a technology perspective is that there is room for improvement.

(iii) **Cost-benefit or risk-reward analysis.** Once the information is retrieved, we have analysed the output, calculating the MOIC for the claimant and for the funder.

Finally, assuming the latter has accounted for 100% of the costs and 40% of the proceeds in case of a positive outcome, we have analysed, as a counterfactual scenario, what would have been its IRR, and what would have been the cases that comply with the theoretical criterion for a funder of minimum 20% Internal Rate of Return and less than 6 years to cash the investment.

Within our set of data and during **Phase I**, the awards have been considered against the 17 questions, divided in three main areas: legal-formal (related to jurisdiction), legal-content, both in in blue, and financial (in yellow). These questions are summarized in the Table below:

#0	Introductory	**Rules applicable to the arbitration**
#1	Introductory	**Rules applicable to the claim** (e.g. BIT (bilateral international agreement), IIA (international investment agreement))
#2	Introductory	Chairman
#3	Introductory	Other (2) Members of Tribunal
#4	Legal-Formal (jurisdiction)	**Limitation – was the claim time barred**
#5	Legal-Formal (jurisdiction)	**Protected investor – is the investor protected**
#6	Legal-Formal (jurisdiction)	**Protected investment – is the investment protected**
#7	Legal-Substantive	**Minimum Standard / Full Protection and Security standard (FPS)**
#8	Legal-Substantive	**Fair and Equitable Treatment (FET) Standard**
#9	Legal-Substantive	**Denial of international justice**
#10	Legal-Substantive	**Expropriation-- direct or indirect**
#11	OUTPUT	**OUTCOME – in favour of Claimant / Respondent**
#12	Financial	**Quantum claimed or claim originally requested**
#13	Financial	**Costs of proceedings**
#14	Financial	**Actual award – principal**
#15	Financial	**Actual award – interest**
#16	Financial	**Total quantum received or actual award (including or not costs)**
#17	Financial	**Estimated time from damage to award (time of accrual of interest)**

Table 18. Introductory (#0 through #3), Legal (#4 through #10) and economic (#11 through 17) Features.

Some of the aforementioned variables have been tagged as ´introductory´ and are deemed, in our approach, irrelevant to the outcome. For instance, the composition of the members of the Tribunal, which is assumed to be impartial and independent, or whether the decision has been unanimous or there has been a dissenting opinion from one of the members of the tribunal; or the particular rules applicable to the arbitration or the claim - which we assume are homogeneous across ICSID and

UNCITRAL - not to be confounded with the rules of the investment agreement under review (e.g. a bilateral agreement between two countries).

The extracted features, traceable to the award, are the source of the input for the model, that are summarised below.

Procedural law | Formal	Substantive Law | Content
Jurisdiction *Ratione temporis* Limitation *Ratione personae* *Ratione materiae*	Minimum Standard/Full Protection and Security Standard (´FPS´) Fair and Equitable Treatment (´FES´) Access to international justice Expropriation (direct / indirect) Most Favoured Nation (´MFN´) Umbrella Clause (´UC´)

Table 19. Research categorization – Legal features, IIA breaches.

Left column in Table 19 refer to legal features related to jurisdiction, whilst the right column are the legal categories in connection with the merits (breaches alleged and found).

Economic Variables	Ratios
Expected damages (claim request) Costs Principal and interests Actual damages (total award)	**RATIO 1** Hit ratio (Claimant) **RATIO 2** Multiple On Invested Capital (MOIC) for the claimant
	RATIO 3 Multiple On Invested Capital (MOIC) for the TPF. **RATIO 4** Internal Rate of Return (IRR) for the TPF.

Table 20. Research categorization – Financial variables and ratios.

Table 20 comprise the economic data of an award: actual damages equal principal, plus interest and costs, if when awarded.

Ratio 1 or ´hit ratio´ relate to how much the claimant has actually received vis-à-vis what was requested as principal, without including interest. Ratio 2 captures the Cash on Cash (CoC) or MOIC for the claimant. Ratios 3 and 4 are relevant to the funder: what would have been the Internal Rate of Return should the funder have born all the costs incurred by the claimant and received a percentage of 40% of the compensation – percentage driven by the EU TPLF draft legislation. For the purposes of Ratio 4, the variable timelapse (or period) has been factored in. The following section will elaborate further on the categorization.

Data extraction - Retrieval Augmented Generation (RAG) and Prompting

Through ChatPDF - embedded into ChatGPT 3.5, version available as of the date of the analysis and issue of this report – a limited number of international investment dispute awards have been studied – an original sample of 22 with questioning in several iterations (9) changing length, syntaxis and semantics, in order to obtain a more satisfactory outcome (*id est* more accurate responses).

1	#Arbitration rules # <what are the rules applicable to the arbitration proceedings? International Center for Settlement of International Disputes (ICSID), UNCITRAL or other rules?>
2	#BIT (bilateral international agreement), IIA (international investment agreement) and rules applicable to the claim # <This field can include the name of two countries, and the date when was signed. This agreement and these rules should be applicable to the case and be enforceable at the time of the award. What is the bilateral International Treaty (BIT) or International Investment Treaty (IIT) applicable in the award? If no treaty is mentioned or applicable, could other applicable rules apply, such as legislation regarding "foreign investments"? >
3	#Members of the tribunal# <What are the names of the Chairman and the other members of the tribunal? This can be found in the document and on the first page>
4	#Limitation# < Answer YES if, according to the final decision of the tribunal, the claim has being filed in due course by the claimant or the claims are not time barred. Otherwise, answer NO >

5	#Protected investor# <Answer YES if, according to the final decision of the Tribunal, the claimant or claimants are considered protected investor(s) under the treaty or the rules applicable to the claim. Otherwise, answer NO >
6	#Protected investment# <Answer YES if, according to the Tribunal in their final decision, the investment is protected under the treaty or international agreement. Otherwise, answer NO >
7	#Full Protection and Security standard (FPS)# <Answer YES if, according to the tribunal in their final decision, the defendant has breached its obligation to provide the claimant with full protection and security. Otherwise, answer NO>
8	#Fair and Equitable Treatment# <Answer YES if, according to the tribunal´s final decision, the respondent failed to provide the claimant with FAIR AND EQUITABLE TREATMENT (FET). Otherwise, answer NO>-
9	#Denial of international justice# <Answer YES if the tribunal, in their final decision has stated that the claimant has been denied international justice by the respondent. Otherwise, answer NO>
10	#Expropriation# <Answer YES if the Tribunal in their final decision has stated that the claimant has suffered expropriation. Otherwise, answer NO. Additionally, If there has been expropriation, has it been DIRECT or INDIRECT. >
11	#OUTCOME# <Read and focus on the ´award´ section. What was the outcome of the arbitration proceeding between the claimant and the respondent? Who did the Tribunal rule in favour of? Was the claimant successful or, on the contrary, did the tribunal accept the defendants' arguments in the final decision? If so, first word of the answer should be CLAIMANT. Has the claim been rejected by the tribunal in their final decision and, as a consequence, the respondents have been successful? If so, first word of the answer should be RESPONDENT and should answer the following question: What was the reason given by the tribunal in order to deny compensation to the claimant?>

Table 21. List of questions through prompting (version 9) related to introduction (#1 through #3), legal content (#4 through #10) and outcome (#11)

The award begins with the Tribunal declaring itself competent in terms of jurisdiction: has the claim been filed in due time and place, by the right counterparty on the right

object (*ratione personae, ratione temporis, ratione materiae*). Formal issues are resolved at this juncture, such as time limitation, or if investment or investor are protected according to bilateral treaties or international agreements. In a positive outcome award, these first steps have been considered positive for the claimant, and the tribunal continues to evaluate the case on merits.

The output of the questionnaire from ChatPDF has been compared with the original sample (the award in pdf version), ensuring that the answers to the questions were accurate. Input to our model has been cross checked to the UNCTAD website, when and if available.

The original sample has been extended later on to complete the 94 awards sample. Responses have been double-checked as a quality assurance test regarding the retrieval and extraction of relevant information, given the infancy of the programme and the limitations highlighted in the literature review. When the information was not available, a reasonable proxy has been defined.

These assumptions relate mainly to lack of historical interest rates not available for the period of reference, where a reasonable estimation has been made. All the numbers within the data set have been traced.

The settlement part of the award, within the last paragraphs of the award, when available, contains most of the information within our model: whether the award has been granted, the rationale and allocation of costs. An exception is when costs have not been awarded, which is normally found in the summary by the Tribunal of the positions of the parties, within the claimant´s *petitio*.

Additionally, several economic variables have been considered: what was the amount requested by the claimant vis-à-vis the amount of principal obtained? In case of several figures, when and if available within the award, the lower figure has been input.

12	#AWARD REQUEST# <Read and focus on the ´award´ section or any reference to summary of the parties´ claims or claimant request for relief. How much is the claimant requesting? What is the amount that the claimant or claimants originally requested as compensation in their claim? If there are several figures, please state >
13	#Costs of proceedings AWARDED# <Read and focus on the ´award 'section. What are the costs to be reimbursed to the claimant when the outcome has been positive for the claimant? >
14	#AWARD - PRINCIPAL# <Read and focus on the ´award 'section. In case of positive outcome for the claimant, what is the principal amount of the damages awarded by their tribunal in the final decision? >
15	#AWARD-INTERESTS# <Read and focus on the ´award' section. In case of positive outcome for the claimant, what is the amount of interest that the Tribunal in their final decision orders the respondent to pay so as to compensate the claimant? >
16	#TOTAL ACTUAL AWARD# <Read and focus on the ´award´ section. In case of positive outcome, what is the award granted by the tribunal in their final decision in favour of the claimant. It can include principal and interest. >
17	#ESTIMATED TIME# <Please answer two questions: 1. What is the date of the tribunal decision (also referred to as the date of dispatch to the parties or the date at signature page). 2. In case of positive outcome for the claimant, what is the date where interest starts to accrue? it can also be the date where the tribunal states that the damages started >

Table 22. List of questions through prompting (version 9) - financial aspects (#12 through #17)

We have broken down the actual award received between the principal and the interest element. The latter could represent a significant amount in case of a lengthy process, or where two-digit compound interest rates apply. Some assumptions have taken place when considering an index (LIBOR, EURIBOR, US Treasury rates, etc). When a time series has not been detailed (e.g. short-term LIBOR without detailing whether deposits relate to 3, 6, 9 or 12 months), 12 months has been considered. If costs have been awarded, totally or partially at discretion of the tribunal - legal fees, other expenses including Tribunal administrative fees – these have been added to the actual amount received.

C- Categorisation

Legal Features

According to the UNCTAD data base, the number of alleged breaches amount to 2,594 whilst the breaches found amount to 443.

In average, for the 260 positive outcomes, 1.7 breaches are found. An overview of the breaches can be found in Table 23 below.

IIA provisions	Alleged breaches		Breaches found	
Direct expropriation	152	6%	**46**	**10%**
Indirect expropriation	534	21%	**78**	**18%**
Fair and equitable treatment/Minimum standard of treatment, including denial of justice claims	673	26%	**185**	**42%**
Full protection and security, or similar	320	12%	**25**	**6%**
Umbrella clause	192	7%	22	5%
National treatment	173	7%	10	2%
Most-favoured nation treatment	145	6%	5	1%
Performance requirements	15	1%	4	1%
Losses sustained due to insurrection, war, or similar events	6	0%	3	1%
Customary rules of international law	10	0%	1	0%
Transfer of funds	47	2%	8	2%
Arbitrary, unreasonable and/or discriminatory measures	278	11%	42	9%
Other	49	2%	14	3%
	2594		443	

Table 23. Alleged and found breaches of IIA provisions according to UNCTAD data base (Investment Policy Hub. (n.d.-d)).

We have focused on breaches found that account for 75% of the total or 84% if we consider ´arbitrary, unreasonable and/or discriminatory measures as failing to comply with fair and equitable treatment. Once Most-favoured Nation Treatment and Umbrella Clause are added, the breaches proportion amounts to 91%.

Time barring considerations (Aceris Law LLC. (n.d.))

The concept of extinctive prescription corresponds to the common law doctrine of laches: equity aids the vigilant, not those who sleep on their rights. International law does not lay down any general time limit concerning limitation periods, limitation of actions or statute of limitations, generally taken care of by national laws. Investment treaties can contain similar express provisions. A host State may rely upon the equitable notion of extinctive prescription in an attempt to defeat the claims.

Ratione personae and ratione materiae (Aceris Law LLC. (2022, June 26)).

As part of their objections to jurisdiction, States can invoke denial of benefits in investment arbitration based on the background of the investor (jurisdiction *ratione personae*) or the nature of the investment (jurisdiction *ratione materiae*). The former tries to avoid ´treaty shopping´ and can be found in the Energy Charter Treaty (article 17), whereby *"Each Contracting Party reserves the right to deny the advantages of this Part to: (1) a legal entity if citizens or nationals of a third state own or control such entity and if that entity has no substantial business activities in the Area of the Contracting Party in which it is organised".*

The latter intends to ensure a real economic relationship with the home state. Once more defined in the ETC (article 18(2)): *"Each Contracting Party reserves the right to deny the advantages of this Part to: (…) (2) an Investment, if the denying Contracting Party establishes that such Investment is an Investment of an Investor of a third state with or as to which the denying Contracting Party: (a) does not maintain a diplomatic relationship; or (b) adopts or maintains measures that:(i) prohibit transactions with Investors of that state; or (ii) would be violated or circumvented if the benefits of this Part were accorded to Investors of that state or to their Investments."*

Minimum Standard and Full Protection and Security (FPS) standard referred to investor or the investment (Müller, D., Fuchs, A., & Pika, M. (Eds.). (2024, March 4)).

The Full Protection and Security (FPS) standard creates an obligation for the host State (i) not to harm investors/investments through acts of State organs or acts otherwise attributable to the State and (ii) to protect investors and investments against actions of private parties, *e.g.* in the course of civil unrest. The standard is complementary to the State's monopoly over the use of physical force and the prohibition of vigilante justice.

Accordingly, the full protection and security standard can be violated through State action as well as inaction (Kinnear, M., & Fischer, G. R. (Eds.) (2015)).

The minimum protection standard provides a "floor" to treatment that States are obliged to respect vis-à-vis foreign nationals and their nationals (Alvarez, B. (Author). (2024, February 13)).

Fair and Equitable Treatment (FET) standard in Investment Arbitration

Fair and equitable treatment is a prominent standard of protection in investment arbitration disputes, present in most bilateral investment treaties ("BITs") (Newcombe, A. P., & Paradell, L. (2009), page 255). There are several variations in the drafting of fair and equitable treatment provisions, although arbitral tribunals have been keen on interpreting fair and equitable treatment as an autonomous and independent treaty standard (Newcombe, A. P., & Paradell, L. (2009), pages 264-265).

Three main approaches to interpreting fair and equitable treatment based on the BIT's language have been identified: FET subject to (i) the Minimum Standard of Treatment understood as a standing body of customary rules agreed by the host states of investment to protect an alien from another country (Islam, & Ghosh. (2018), Islam, R. (2019)); (ii) the Principles of International Law, as an obligation to be carried out "*in accordance with*" the sources of international law. Finally, (iii) FET as an Autonomous Standard is the preferred construction amongst arbitral tribunals as some BITs refer to fair and equitable treatment delinked from international law or the minimum standard of treatment (Islam, R., Islam, & Ghosh. (2018), page 68). Such clauses give significant discretion to arbitrators in interpreting fair and equitable treatment.

Fair and equitable treatment could be based on

- Arbitrary measures fail, by definition, to be fair and equitable.
- Damage to investors' legitimate expectations. Many tribunals have accepted legitimate expectations as a subcategory of fair and equitable treatment (International Arbitration Attorney (2022, January 23)), based on (i) changes of circumstances, (ii) the host state's conduct and representations (typically, in the form of oral or written statements) (Newcombe, A. P., & Paradell, L. (2009), pages 279-280) - when were these presentations made, by which level of

authority, any disclaimer regarding the state´s undertakings; and (iii) the investor´s conduct – what kind of expertise took place assessing or interpreting these representations.

- Transparency means that "*the legal framework for the investor's activities and operations is clearly laid out and that any decisions affecting the investor can be traced back to that legal framework.*" (Diehl, A. (2012), page 369)
- Lack of due process is often associated with the notion of denial of justice (see below), although with a broader sense, being applied to all forms of decision-making, including measures taken by the government on an administrative and legislative level ((Dumberry, P. (2013), page 231)).

Paparinskis, M. (2013), in a very limited sample, 24 awards in the period 2010-2011 highlighted that 10 awards had found breach of *"fair and equitable treatment or international minimum standard or any other obligation"*.

In the study at hand, within the sample and from a data classification perspective, and not entering to the discussion of the academic distinction, we have embedded the international minimum standard and discriminatory measures within the Fair and Equitable treatment. Unlike the UNICTAD website, we have separated Denial of International Justice from the FET, in search for further granularity.

Denial of Justice in International Arbitration (International Arbitration. (2021, August 11))

Denial of justice may be recognized in the following circumstances: "*refusal of access to court to defend legal rights, refusal to decide, unconscionable delay, manifest discrimination, corruption, or subservience to executive pressure"* (Paulsson, J. (2005), page 204).

Case law and doctrine reach the same conclusions as the Tribunal at Lion Mexico Consolidated LP v. United Mexican States (International Centre for Settlement of Investment Disputes. (2021)) where "*denial of justice is an international wrong which breaches the fair and equitable treatment standard"*.

Denial of justice in international arbitration concerns acts or omissions of a State's judiciary for which a State may be internationally liable. Although functionally independent body from a State's executive and government, the judiciary is a State's organ (Mourre, A., & Vagenheim, A. (2010)).

Expropriation in Investment Arbitration (International Arbitration Attorney. (2022, March 13))

Each State has the right to exercise sovereignty over its territory and an obligation to respect properties belonging to foreigners. A State may, in special circumstances and meeting certain criteria, expropriate a foreign investor's property (Newcombe, A. P., & Paradell, L. (2009), page 321). Exceptions to measures equivalent to expropriation or nationalisation must meet the following cumulative criteria (i) public purpose; (b) non-discriminatory manner; (c) prompt, adequate, and effective compensation; and (d) in accordance with due process of law and Minimum Standard of Treatment (article 6 of the 2012 U.S. Model Bilateral Investment Treaty).

Under customary international law, expropriation may be divided into direct and indirect expropriation.

In the former, the host State deliberately seizes property and transfers its rights to itself or to a State entity (Newcombe, A. P., & Paradell, L. (2009), page 322), traditionally found in the context of nationalisation of strategic sectors and industries, such as roads, parks, mines and oil fields. Tribunals have found indirect expropriation in a wide array of State measures, sometimes even if the investor retains the investment's formal ownership, including requisition of lands, forced sales, exorbitant taxation, deprivation of profits, interference in the management of a business, termination of rights, such as licences, contracts or debts, blocking and harassment of employees, blockage of plants, and prohibition on the repatriation of profits (Newcombe, A. P., & Paradell, L. (2009), page 328).

From a classification perspective, it has been sought whether Direct or Indirect expropriation has a relationship with the rest of the variables.

Two additional variables have been included, given the availability within the UNCTAD database:

In the second phase of our data gathering, two additional legal variables have been added, so as to explore and quantify their impact vis-à-vis the existing selected variables.

Most-Favoured Nation (MFN)

Most-Favoured Nation clauses figure in the vast majority of investment protection treaties (Aceris Law LLC. (2021, February 18)). They are intended to ensure *"that a host country extends to the covered foreign investor and its investments, as applicable, treatment that is no less favourable than that which it accords to foreign investors of any third country".*

Whilst there is little doubt that an MFN Clause can be and has been used to import substantive protection standards such as Fair and Equitable Treatment (FEt) (Dumberry, P. (2017)), Full Protection and Security Standard (FPS), or Umbrella Clause (UC), more controversy emerges regarding the use of an MFN Clause in order to import more favourable procedural and/or dispute resolution provisions from a third treaty.

The inclusion of such a variable within our study research could therefore be debatable, since on one hand we have not figured out within the sample cases whether a substantive or dispute resolution call has been made, and within the former, whether the FET/FPS or UC was applicable.

Umbrella Clause (UC)

In investment arbitration, contract claims, or even State unilateral commitments, might be considered a violation of an international obligation under a BIT "imposing a requirement on each Contracting State to observe all investment obligations entered into with investors from the other Contracting State" (Wong, J. (2006)). The umbrella clause must be incorporated within the corresponding BIT (Aceris Law LLC. (2022, May 1)), that is, this protection ("umbrella"), is not automatic.

Outcome

The Tribunal acting in the arbitration can reject the claim and accept a respondent´s objections to jurisdiction in favour of the respondent or, if accepted, continue to evaluate on merits. Should the Tribunal find the respondent State liable for any reason, whether from a formal or substantive law perspective, this is a win for the claimant. In this case, compensation is normally granted.

Occasionally, a settlement is reached, or the claim is discontinued. The former could be construed as if the claimant has obtained some kind of compensation, however it is normally not public. The latter invites consideration that the claimant has withdrawn his claim, although, once more, we have not found sufficient information in the database. As stated previously, this research focuses on the awards in favour of the claimant, where compensation has been granted.

Financial variables

An initial note on currency, exchange rate and inflation: the figures sampled have been expressed in USD at the exchange rate calculated to the date of dispatch of the awards when applicable. Inflation has not been taken into account whilst interest on the compensation amount intends to mitigate this circumstance.

Quantum requested

The expert in damages on behalf of the claimant justifies the amount, which is refuted by the respondent´s expert. The input figure has not considered the interest requested, so as to make a like-for-like comparison of the award granted by the Tribunal. In case of several figures or scenarios filed by the claimant, the lowest figure has been selected.

Costs

These include legal fees and experts´ fees as well as related expenses, from witnesses to advances to the Tribunal. Costs would normally include legal expenses plus Tribunal and administrative fees, which could represent, approximately, between 10% and 35% of the total costs of proceedings.

Given the discretion of the tribunal in cost allocation, (i) the costs could be awarded or not, in fairness, irrespective of winning or losing and (ii) in case of positive outcome for the claimant, the amount of costs awarded could be different from the costs that have been claimed. When available, we have included the costs actually incurred versus the costs actually awarded, which is a more precise representation of what have been the actual costs to reach the award.

Principal

This amount awarded by the tribunal, which should not be taxed by the respondent State. The award can also be granted as a lump sum, i.e. without breaking down between principal and interest, or because it is not applicable, for instance, when a discount factor has been applied.

Interest

Only on few occasions we have found that the tribunal undertakes the actual calculation of the interest. Interest can be reflected into the final award by detailing the methodology for calculation. The Tribunal normally refers (i) to an inception time where compensation needs to accrue interest; (ii) a rate or reference index (e.g. LIBOR, EURIBOR) or a fixed-income security (US Treasury bonds) within a particular time reference (annual, quarterly, semi-annual), (iii) indicating whether the interest is simple or compound , and (iv) occasionally adding a spread.

When not available for the time period, reasonable assumptions have taken place. For instance, the average rate for the years within the period and annual time reference if no other reference is expressed by the Tribunal in the award.

Late payment interest after the award was issued have not been considered in the model, since it is assumed that the award is cashed at the time of the issue of the award.

Total Award

Quantum actually received, including interests and the reimbursement of costs, if finally awarded in favour of the claimant.

Estimated time

Interest intends to further compensate for the loss of value of money over time. The estimated time refers to the time elapsed from the moment that interest starts to accrue until the date of the award. Of all the assumptions, the least realistic is that the date of payment coincides with the date of issue of the award: further claims in civil jurisdiction could delay the payment in case of a positive outcome for the claimant that has been challenged by the respondent. Even if late interest applies, the awards are contested in the civil jurisdiction. Very few progress, well below one out of ten, however finding these statistics is very challenging.

It is indeed unlikely that the payment takes place that very same date, despite some awards including late payment interest (or continuous interest until payment takes place). In order to mitigate – partially - this circumstance, when calculating the IRR for the funder, we have extended the time to recovery by rounding up the years to the upper digit (e.g. 4.2 years would lead to 5 years for IRR calculations for the funder). As explained above, a delay in 1, 2 or more years significantly impacts the IRR for any TPF.

Ratios applicable to the Claimant

Discount RATIO 1 (or ´**Hit Ratio**´): how much principal has been received vs how much was expected as shown in the ask to the Tribunal.

Hit Ratio = Principal / Claim Requested
Principal does not include interests or costs (if awarded)
This ´hit ratio´ percentage could be expressed as
Hit ratio = 1- Discount

Table 24. Hit Ratio - applicable to Claimant.

This ratio could be used by the parties or the analysts to assess the reasonableness or proportionality of the amount sought in relation to the damages claimed. It can also provide insights into the potential risks associated with the arbitration case, e.g. a small hit ratio could have been caused by a disproportionate ask, or a low award granted, or both.

As previously stated, in case of several amounts claimed, the lowest number has been selected.

Cash on cash RATIO 2 (´**CoC**´ for the claimant or ´**MOIC Claimant**´): multiple that expresses the times the actually-received quantum contains the investment (costs of the process).

MOIC (Claimant) = Total Award / Costs
where
Total Award = Principal + Interest + Costs
Costs in Total award are added if and only if they have been awarded

Table 25. Cash on cash ratio or Multiple on Invested Capital (MOIC) for the claimant

Ratios applicable to a Third-Party Funder

Cash on cash (´**CoC**´) or **MOIC**´ or **RATIO 3** for the TPF is defined as the cash-on-cash amount received by the funder assuming that disbursed the costs (budget or investment) and received 40% of the proceeds. To restate, this Cash-on-Cash ratio is equivalent to Multiple on Invested Capital.

MOIC (TPF) = MOIC (Claimant) x 40%

Table 26. Cash on cash ratio or Multiple on Invested Captial (MOIC) for the TPF

Internal Rate of Return (´**IRR**´) for the TPF or **RATIO 4.** Should the funder assume 100% of the costs and receive only 40% of the proceeds, what would be IRR, taking into account the timespan.

The period is comprised between two dates: (i) when the damages started – and interest began to accrue - and (ii) the date of the award, as a proxy of payment of the award.

For IRR calculations, we have rounded up the time period (e.g. when 4.7 years become a calculation at year 5.0). In the simulation exercise, however, we account using the precise figure, without rounding up.

The initial investment or outflow for the investor is the costs figure. The return or inflow represents 40% of the total award, calculated as total award, including interests and costs - if and when awarded by the Tribunal.

$$0 = \mathrm{NPV} = \sum_{t=1}^{T} \frac{C_t}{(1+IRR)^t} - C_0$$

where:

Ct=Net cash inflow during the period t

*C*0=Total initial investment (costs at the beginning of the interest period)

IRR=The internal rate of return

t=The number of time periods

Table 27. IRR formula (applicable to TPF)

Proxies and Caveats

Related to timespan.

(a) The date of the investment return is assumed to be the award date. However, the payment in practice occurs at a later stage, whether a few days or weeks or even after a longer period if the award is challenged requesting clarification, or even further in time if contested in the civil jurisdiction.

The award can be requested to be supplemented, rectified, interpreted, revised, or, exceptionally, annulled (International Centre for Settlement of Investment Disputes. (2022)). From the moment the award is on one hand rendered until it is paid by the losing party and on the other collected by the wining party can vary from a few days or months to several years (e.g. 3 years in the US for a foreign arbitral award under the Federal Arbitration Act (Salomon, C. T., & Yamamoto, H. (2020, June 9)).

The award can be paid promptly or require enforcement actions in domestic courts or through international mechanisms such as the New York convention on the Recognition and Enforcement of Foreign Arbitral Awards (Lu, M. (2005)). The time to challenge the awards in domestic jurisdictions varies from country to country. Refusal to pay by the losing party may imply initiating legal proceedings to enforce the award, which may lead to seizure of assets. If the Court finally confirms the decision of the arbitration tribunal, interest also accrues.

Needless to add, further delays would further dent the IRR, so the IRR output should carefully be caveated in this respect. Despite mitigating partially, for IRR calculations we have mitigated this later variable by rounding up the years.

This does not mean that the dates are not certain or that they do not exist, only that they are not publicly available, so a reasonable proxy is in order.

(b) The costs are normally incurred at different moments of time; although the bulk are normally disbursed before filing the claim - legal fees, experts, advance payments, etc. Where the date of commencement of procedures is available, it is more conservative to assume that the investment starts when the damage starts to occur, before filing the claim.

The bulk of the legal fees are incurred before filing the claim. Other costs must be projected over the expected time of the process.

Related to the costs or investment.

(c) Additional to timing for costs incurred, in terms of amount: all the costs incurred by the funder are the investments undertaken by the claimant, requested by the claimant to the tribunal. These costs are different from the relief sought, whether as a lost profit or actual damage.

The Tribunal has great discretion to allocate these costs, further to hearing each party´s justification. As previously mentioned, if costs allocated at the award are lower than the costs actually disbursed by the claimant or TPF, where available, the higher number has been input in the model.

Related to the award (total compensation)

(d) **Interest rates:** when available (e.g. USD Libor semi-annual), the average of the benchmark of the years corresponding to the lifespan.

(e) **Tax neutrality of the award:** The awards normally ensure the tax neutrality of the payment to the claimant. If a funder is located in a different jurisdiction, this might impact the risk-reward outcome of the transaction, becoming very relevant from an asset management portfolio.

4. Results

This section presents the details of the sample obtained and key statistics. In addition, the regression model obtained that allows estimating the main financial outputs of an award is explained.

Once the different relevant variables have been identified in section 3, a sample of 94 awards were randomly selected from the 260 awards positive to the claimant. For those 94 awards, the different value of the features have been extracted by processing the texts through the ChatPDF methodology previously described.

Representative sample

The 94 sample has been tested vis-à-vis the larger database of positive outcomes (n=260), to check whether it was a representative sample of the population. The numerical variable "amount claimed" was available in the UNCTAD data base for all the elements of the awards population.

Confronting the 94 sample values of "amount claimed" in log scale with the values for 240 awards, the probability histograms in Figure 6 are obtained. As can be seen, the sample distribution is similar to the population (All Claims) distribution.

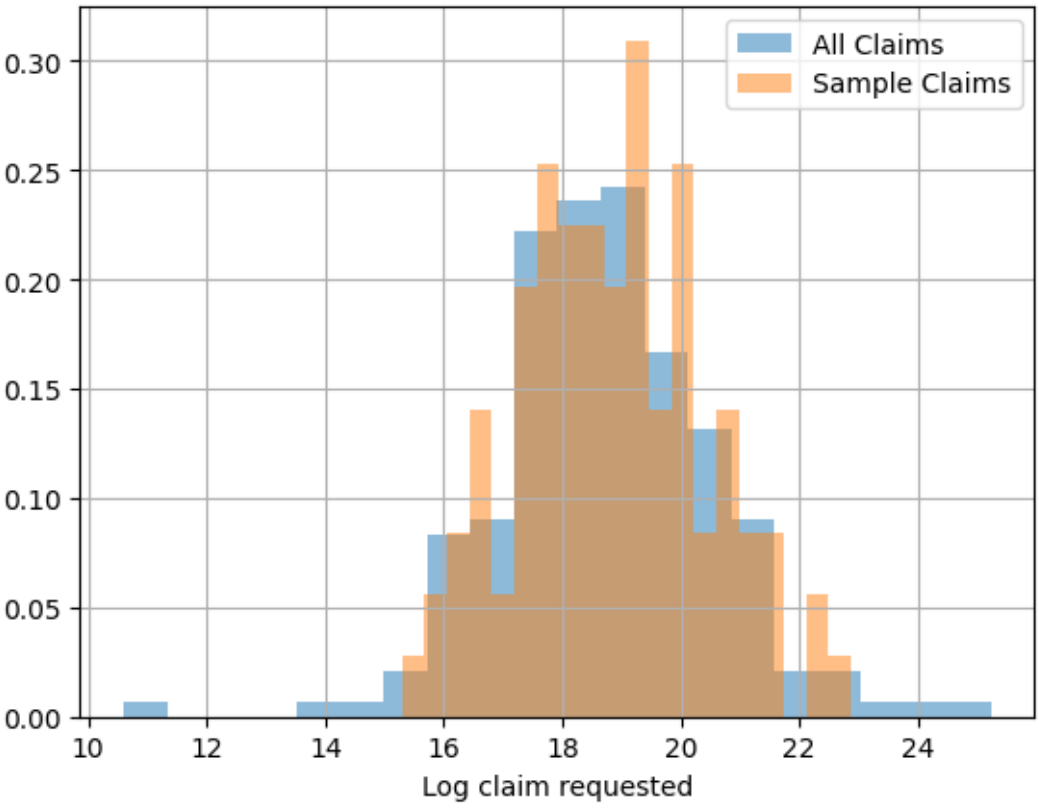


Figure 6. Graphic representation comparing the distribution of the variables in both sample and broader population.

Moreover, a wo-sample Kolmogorov-Smirnov test for goodness of fit is performed, obtaining a p-value of 0.99, thus, clearly not rejecting the null hypothesis of equal distribution. Hence, we can be confident that the sample is representative of the population.

Key Statistics for the Claimant

Table 28 presents a summary of all the variables obtained as well as the main statistics that characterize each of the variables.

FPSMinSt	Full Protection & Security and Minimum Standard
FET	Fair and Equitable Treatment)
DNJ	DNJ (Denial of Justice)
ExpYN	Expropriation (Yes/No)
MFN	Most-Favoured Nation (MFN)
UC	Umbrella Facility (UC)
CCY	Currency (all in USD)
MOICCLAIMANT	Multiple on Invested Capital for the claimant
CLAIM REQUESTED	Amount Requested
COSTS	Costs (legal, experts, administrative)
PRINCIPAL	Award received (without interests)
HIT RATIO	Principal / claim requested
INTEREST	Estimated within the timespan (period)
COSTS AWARDED (Yes/No)	Whether costs have been added to the award
TOTAL AWARD	Including interest and, when awarded, costs
YEARS	Period from start of damage to award date
BLENDED RATE	Applicable to the principal (estimated)

Variable	mean	std	min	Median	max	Type
FPS_MinSt	0.383	0.489	0	0	1	Dicotomic
FET	0.702	0.460	0	1	1	Dicotomic
DNJ	0.298	0.460	0	0	1	Dicotomic
Exp_Y_N	0.479	0.502	0	0	1	Dicotomic
MFN	0.021	0.145	0	0	1	Dicotomic
UC	0.074	0.264	0	0	1	Dicotomic
CLAIM_REQUESTED	5.29E+08	1.17E+09	4.40E+06	1.47E+08	8.50E+09	Numeric
COSTS	7.69E+06	8.26E+06	4.32E+05	5.45E+06	4.93E+07	Numeric
PRINCIPAL	2.06E+08	6.38E+08	7.86E+05	3.92E+07	4.22E+09	Numeric
HIT_RATIO	0.404	0.263	0.014	0.395	1.003	Numeric
Blended_rate	0.048	0.054	-0.001	0.033	0.279	Numeric
COSTS_AWARDED _Y_N	0.691	0.464	0	1	1	Dicotomic
TOTAL_AWARD	2.62E+08	8.02E+08	1.01E+06	5.12E+07	6.09E+09	Numeric
YEARS	6.503	3.158	1.395	6.158	19.855	Numeric
MOIC_Claimant	27.847	78.555	0.315	8.200	649.094	Numeric

Claimant´s perspective	
Total population (# of observations)	260 (n=94)
Amount requested: Mean (Median) Range	USD 529 M (USD 147 M) USD 4.4 M – 8.5 B
Hit ratio: Mean (Median)	40.4% (39.5%)
Award (principal only): Mean (Median) Range	USD 206 M (USD 39 M) USD 0.8 M – 4.2 B
Average costs: Mean (Median)	USD 7.6 M (USD 5.4 M)
Years – Mean (Median)	6.5 years (6.1 years)
Total Award: Mean (Median) Range	USD 262 M (51 M) 1.0 M – 6.1 B
MOIC: Mean (Median)	27.8 (8.2)

Table 28. Key Statistics (GRETL) for the Claimant (N=94). Source: Own data.

The emphasis on median terms instead of the average (or mean) is motivated by the goal of mitigating the skewness caused by significant outliers, such as multi-billion dollar requests.

Within the sample in median terms, USD 5.4 M of costs, a principal obtained of USD 39 M and a hit ratio amount received over amount requested of circa 40% - both mean and median – whilst the MOIC for the claimant is 8.2 times.

Our sample compares within the A&O report (2021) as follows.

Costs for claimant as median for central values (amounts in dispute between USD 50 M and 1 B) amount USD 4.7 and 7.1 M (Table 2) versus USD 5.4 M in our sample.

Regarding length, there is a two-year decalage: the median duration of approximately 3.9 years (Table 3) contrasts with the 6.1 years in our sample. This difference arises from the methodology difference: their calculation counts the years from the commencement of proceedings to the award date, whereas in our thesis, we calculate the years from the onset of accruing interest on damages.

Key Statistics for the Third-Party Funder

According to the definition of litigation finance agreement given by the European draft legislation, the TPF can subrogate in the position in the claimant in cases of purchase or assignment of the claim. For the sake of simplicity and illustrative purposes, we have considered the funder as a financer, rather than a purchaser throughout this research.

The relevant figures from a TPF that accounts for 100% of the costs and 40% of the proceeds show that the Internal Rate of Return (IRR) median is close to 30% (27.4%) whilst the Multiple On Invested Capital (MOIC) median is 3.3x.

Total population (# of observations)			260 (n=94)		
IRR **MOICTPF** **YEARSIRR**			Internal Rate of Return Multiple on invested Capital for the TPF Years (for IRR calculation purposes)		
Variable	**Mean**	**Median**	**S. D.**	**Mín**	**Max**
MOICTPF	11.14	3.280	31.42	0.1259	259.6
IRR	0.5240	0.2737	0.7748	-0.2257	4.891
YearsIRR	6.851	6.000	3.318	2.000	20.00
TPF					
MOIC – mean (median)			11.14x (3.3x)		
IRR – mean (median)			52.4% (27.4%)		
Years – mean (median)			6.9 years (6.0 years)		

Table 29. Key Statistics TPF (GRETL). Summary Statistical indices of the main variables used in the model. TPF´s perspective (N=94). Source: Own data.

Histograms and Correlation Matrix

Observing the distribution of some numerical variables, these are highly skewed. As a consequence, a logarithmic transformation has been applied to "claim requested", "principal", "total award" and "costs", in order to improve the linearity, homoscedasticity, and normality assumptions of the linear regression model.

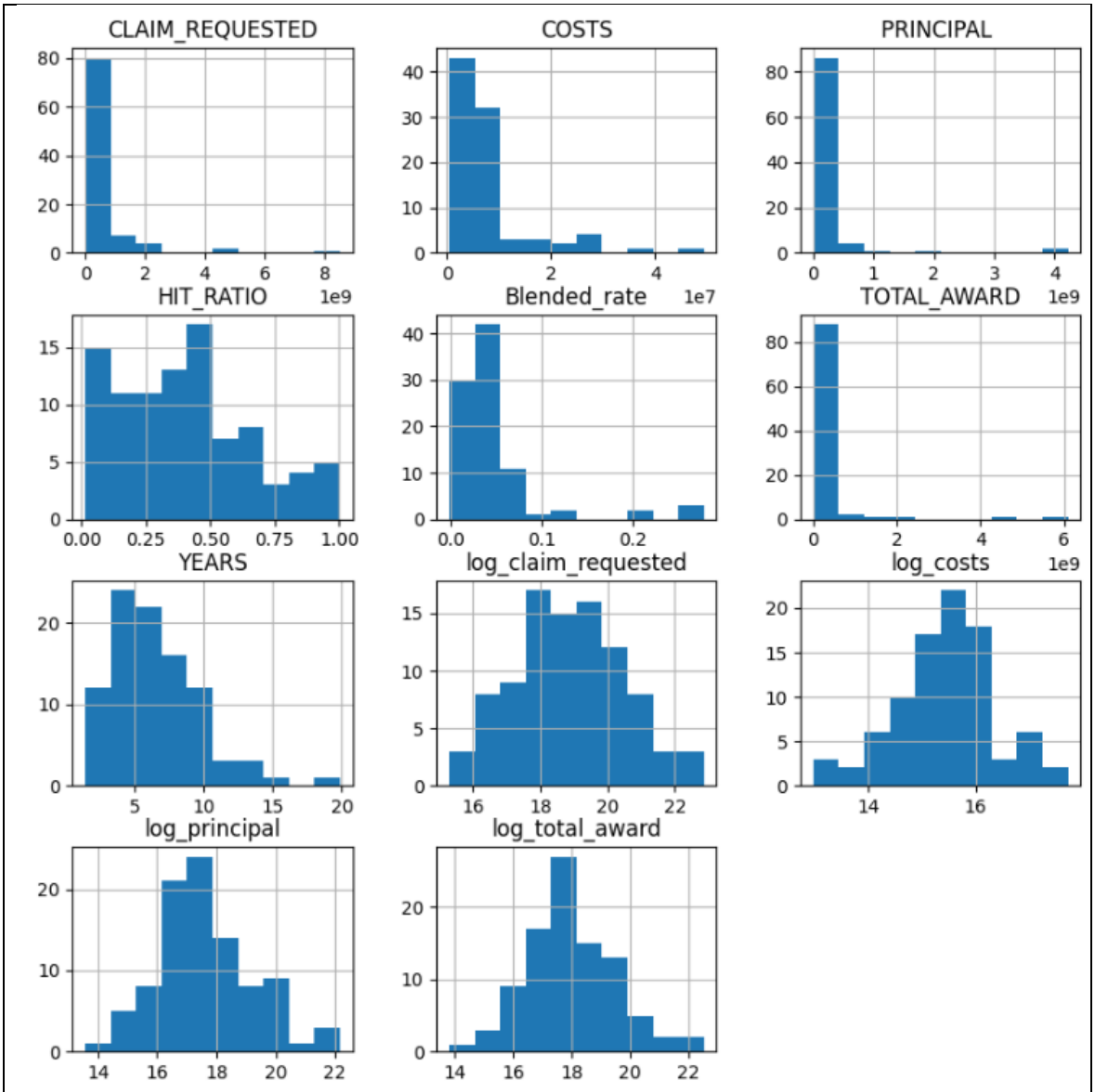


Figure 7. Histograms of selected variables. Source: Own data.

Moreover, outliers in "claim requested" and in "Principal" were found with extreme values far away from the sample. These abnormal situations are not of interest, and should be analysed separately. Therefore, the sample has been filtered to ensure "principal" values lower than USD 1.5e9 and "Claim requested" values lower than

USD 4e9. This effect has reduced the sample from 94 to 85 cases, hence, loosing a small amount of information. Additionally, we applied winsorization to the YEARS variable at the 0.9 percentile to address a few outliers (i.e., years greater than 10).

The following matrix of association measures quantify the degree of linear relationship between the selected variables. Note that no categorical variables are added to the mix, that is, those legal regressors based on discrimination, Full Equitable Treatment, Denial of Justice, MFN, or expropriation whether direct or indirect. Full matrix can be found in Appendix 4 (Figure 10.D).

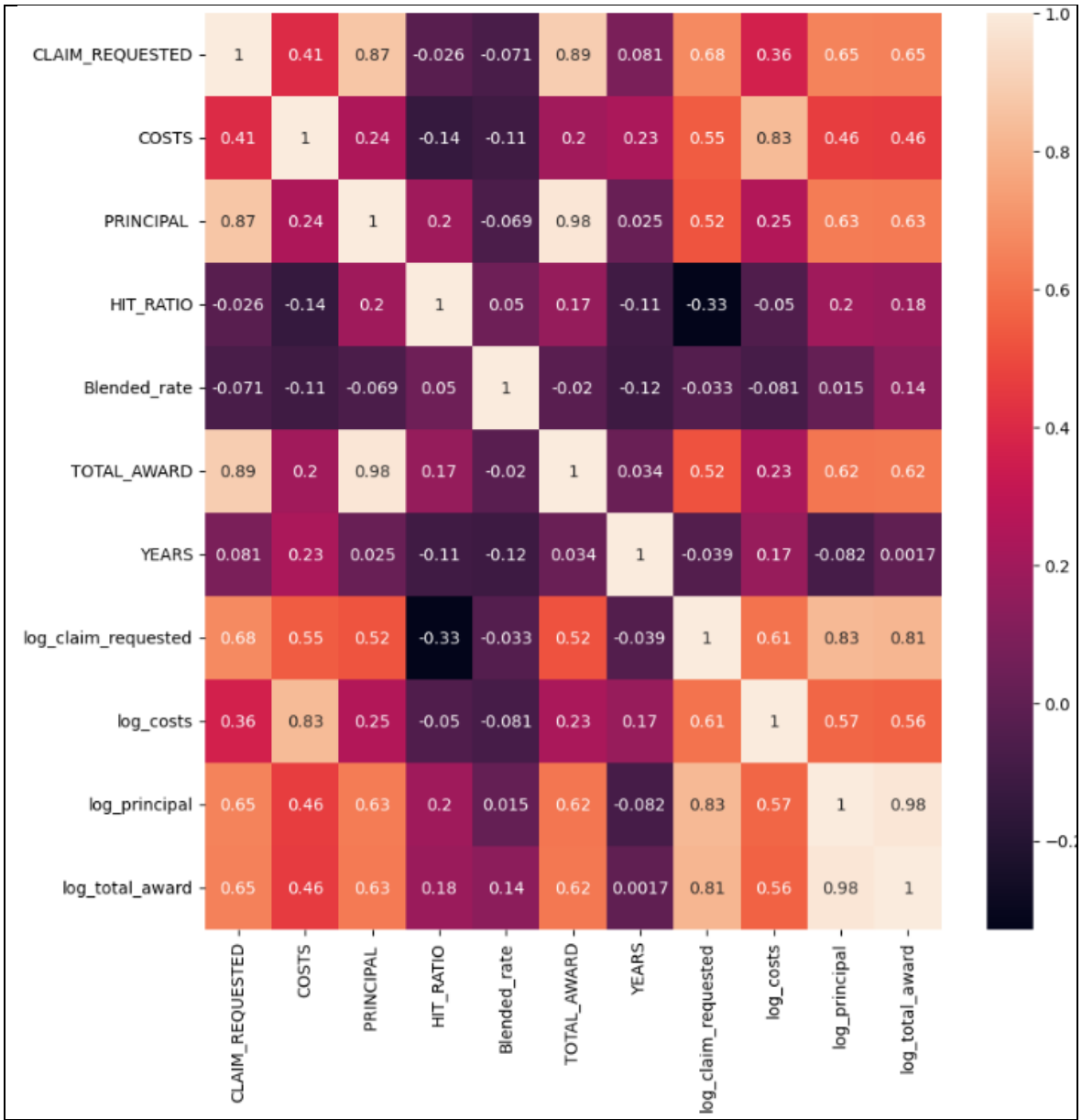


Figure 8. Correlation matrix. Source: Own data.

The variables that are strongly related correspond to those between the principal obtained and the claim requested (0.87, in Figure 8 above). Principal or total award and costs (log normalised) are correlated at close to 0.5. No correlation is observed between total award and years (<1%).

Models output

We aim, as guiding theory to find a calibration methodology that responds to the question under which circumstances a case could be subject to financing – whether prior to filing the claim or once it has been initiated, with a certain degree of confidence for reaching an ´investment grade´ status from the perspective of a potential claimant – or its third-party funder.

We have imported the data set in GRETL and python from Excel and specified the linear regression models.

Two different models are created in this thesis; hence, two dependent variables are defined from a claimant´s perspective: what would be the **principal** (damages award without interest) and **years.**

We have estimated the parameters of the linear regression model using Ordinary Least Squares (OLS) estimation. It computes the coefficients for each independent variable, and other relevant statistics such as R-squared - how well the independent variables in a regression model explain the variation in the dependent variable, R-squared adjusted to the number of predictors, standard errors, and t-statistics.

It is important to highlight that, from now onwards, the research will focus on the claimant´s perspective only. In other words, the model can be further elaborated in order to insert the 100% costs and 40% proceeds criteria. Using diagnostic tests and statistics provided by Gretl and python, the goodness-of-fit of the model has been undertaken and assessed the significance of the coefficients.

Interpretation and Critical analysis -claimant´s perspective

Estimation of the principal of the award

An initial model was estimated with all the input variables. As some variables appear as non-significant a process for reducing the number of input variables is performed.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:          log_principal   R-squared:                       0.655
Model:                            OLS   Adj. R-squared:                  0.613
Method:                 Least Squares   F-statistic:                     135.1
Date:                Wed, 24 Apr 2024   Prob (F-statistic):           1.26e-42
Time:                        07:24:37   Log-Likelihood:                -104.84
No. Observations:                  85   AIC:                             229.7
Df Residuals:                      75   BIC:                             254.1
Df Model:                           9
Covariance Type:                  HC1
=============================================================================================
                              coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------------------
Intercept                  -0.2297      2.047     -0.112      0.911      -4.241       3.782
log_claim_requested         0.6804      0.077      8.820      0.000 ***   0.529       0.832
log_costs                   0.2964      0.151      1.958      0.050 *    -0.000       0.593
FPS_MinSt                  -0.0101      0.300     -0.034      0.973      -0.599       0.579
FET                         0.3073      0.275      1.120      0.263      -0.231       0.845
DNJ                        -0.1159      0.325     -0.357      0.721      -0.752       0.521
Exp_Y_N                     0.1244      0.203      0.611      0.541      -0.274       0.523
MFN                         0.4819      0.466      1.035      0.301      -0.431       1.395
UC                          0.5524      0.214      2.581      0.010 **    0.133       0.972
COSTS_AWARDED_Y_N           0.1826      0.248      0.737      0.461      -0.303       0.668
==============================================================================
Omnibus:                       13.549   Durbin-Watson:                   1.556
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               14.676
Skew:                          -0.902   Prob(JB):                     0.000650
Kurtosis:                       3.944   Cond. No.                         491.
==============================================================================
```

```
OLS Regression Results
==============================================================================
Dep. Variable:          log_principal   R-squared:                       0.801
Model:                            OLS
Method:                 Least Squares   F-statistic:                 1.832e+04
Date:                Tue, 16 Apr 2024   Prob (F-statistic):          9.12e-116
Time:                        10:25:10   Log-Likelihood:                -106.02
No. Observations:                  85   AIC:                             218.0
Df Residuals:                      82   BIC:                             225.4
Df Model:                           3
Covariance Type:                  HC1
=============================================================================================
                              coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------------------
log_claim_requested         0.6957      0.075      9.319      0.000 ***   0.549       0.842
log_costs                   0.2874      0.089      3.245      0.001 ***   0.114       0.461
UC                          0.4677      0.164      2.857      0.004 **    0.147       0.789
==============================================================================
Omnibus:                       16.865   Durbin-Watson:                   1.578
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               19.930
Skew:                          -1.019   Prob(JB):                     4.70e-05
Kurtosis:                       4.215   Cond. No.                         88.5
==============================================================================
```

Table 30.A Principal amount estimation according to the research trained model (n=85). Source: own data.

After reducing the optimization of the structure of the model, the independent variables that demonstrate significance in the regression analysis for the principal as dependent variable are: claim requested, costs and Umbrella Clause. The R^2=0.801 value suggests that the independent variable accounts for a relevant amount of the variability observed. In the Ordinary Least Squares (OLS) regression model, the coefficient of the independent variable represents the change in the dependent variable associated with a one-unit change in the independent variable. However, as the dependent and some independent variables are in logarithms, the interpretation is in relative terms. The sign of the coefficient (positive or negative) in OLS regression determines the direction of the correlation between the independent and dependent variables.

As the coefficient is positive with the three variables, it indicates a positive correlation; in terms of order of magnitude of said increase, as the amount claimed or the costs increase. This coefficient indicates that a 1% increase in the claim requested (in log terms), keeping other factors constant, is associated with an approximately 0.6957% increase in the principal (in log terms). This coefficient implies that a 1% increase in costs (in log terms), holding all else constant, leads to an approximately 0.2874% increase in the principal (in log terms).

As an equation, the formula would be as follows:

*log_principal = 0.696*log_claim_requested + 0.287*log_costs + 0.468*UC*

or, once the logged form is converted to an original scale[24]:

Principal = (Claim Requested)^0.696 × (Costs)^0.287 × e^(0.468·UC)

Table 30.B Principal amount estimation – GRETL output and equation. Source: Own data.

A separate mention should be driven by the Umbrella Clause finding. Each unit increase in the UC, all else being equal, is associated with a 46.77% increase in the

[24] A simulation on the results of the equation, including the inverse logarithm of the equation, can be obtained through the website Wolframalpha https://www.wolframalpha.com (WolframAlpha. (n.d.))

principal (in log terms). This circumstance triggers the attention of the usefulness said clause for private investors in a contractual dispute with at State; reminding to find out whether a Treaty between their country and the host country exists and, if so, whether it incorporates an ´umbrella clause´ that renders and therefore enhances the protection to the investor, the investment, or both. Caution, however, is in order at this juncture, since it is the only categorical variable with statistical significance found, which could be derived by its relative frequency within the sample.

Estimation of years to the award

To remind, the concept years, as defined in the methodology, applies to the timespan between the moment when the interests of an award have started to accrue interests (damages begun), and the time to award.

After a selection of the most significant variables (once fitted and initial model for years), the final model obtained is as follows:

OLS Regression Results

Dep. Variable:	YEARS	R-squared:	0.213
Model:	OLS		
Method:	Least Squares	F-statistic:	185.8
Date:	Tue, 16 Apr 2024	Prob (F-statistic):	5.90e-40
Time:	10:25:13	Log-Likelihood:	-180.65
No. Observations:	85	AIC:	369.3
Df Residuals:	81	BIC:	379.1
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
log_claim_requested	-0.4447	0.190	-2.337	0.022 *	-0.823	-0.066
log_costs	0.9816	0.232	4.231	0.000 ***	0.520	1.443
FET	-1.7746	0.544	-3.265	0.002 **	-2.856	-0.693
DNJ	1.6104	0.545	2.953	0.004 **	0.526	2.695

Omnibus:	0.260	Durbin-Watson:	1.970
Prob(Omnibus):	0.878	Jarque-Bera (JB):	0.444
Skew:	0.010	Prob(JB):	0.801
Kurtosis:	2.646	Cond. No.	70.6

Table 31.A Years estimation according to the research trained model (n=85). Source: own data.

A statistically significant relationship between the following variables have been found: The p-values associated with the variables claim requested, costs, Fair and Equitable Treatment and Denial of International Justice are below the 5% (claim

requested) and 0.5% threshold (costs, FET and DNJ) for statistical significance at 95% and 99% degree of confidence, respectively.

In terms of explanation of the coefficients, the signal of amount claimed and FET is negative, meaning that, the larger the amount, the shorter the period to award. This is relatively simple to assume in the FET breach (reduction of 1.8 years), whose evidence would become a strong argument to win the case, whilst the amount of dispute could be driven by the need to close such reputational and potentially interest-bearing liability by the State - although we do not have a rationale for this particular outcome.

On the other hand, should the breaches alleged and found deal with Denegation of Justice, the tenor would extend in 1.6 years, whilst the increase of the process, according to the historical data gathered, is close to 1 year for every USD 1 M of costs. The R^2=0.213 indicating that there is still more variability of the dependent variable to be explained.

The prediction equation of the model is given by the following formula:

YEARS = - 0.445*log_claim_requested + 0.982*log_costs - 1.775*FET + 1.61*DNJ

Table 31.B Years estimation – GRETL output and equation. Source: Own data.

5. Discussion

Full model Simulation

Once the regression models are fitted, in order to have further discussion beyond average or mean values, we move onto prediction scenarios.

In Gretl, independent variables can be input into the fitted model to obtain predicted values for the dependent variable.

FET: 1	**Claim Requested : Budget**	**1:10**	**1:10**	**1:10**
DNJ: 0	**Claimed Requested ($)**	**20.000.000**	**50.000.000**	**100.000.000**
UC: 1	**Budget ($)**	**2.000.000**	**5.000.000**	**10.000.000**
Predicted Principal		$12,398,397.59	$30,519,546.11	$60,327,997.6
Predicted Years		5,0	5,5	5,9
Predicted Hit Ratio		61.99%	61.04%	60.33%
Predicted MOIC (claimant)		6,20	6,10	6,03
Predicted IRR (claimant)		121,67%	100.15%	98.92%
FET: 0	**Claim Requested : Budget**	**1:10**	**1:10**	**1:10**
DNJ: 1	**Claimed Requested ($)**	**20.000.000**	**50.000.000**	**100.000.000**
UC: 0	**Budget ($)**	**2.000.000**	**5.000.000**	**10.000.000**
Predicted Principal		$7,766,578.26	$19,117,990.17	$37,790,537.9
Predicted Years		8,4	8,9	9,2
Predicted Hit Ratio		38.83%	38.24%	37.79%
Predicted MOIC (claimant)		3,88	3,82	3,78
Predicted IRR (claimant)		41.22%	40.49%	36.05%

Table 32. Sensitivity of predicted output. Several scenarios comprising the same ratio budget (expected costs) to Claim request in a best-case scenario (FET and UC) and a worst-case scenario (DNJ). Source: own data.

For a similar expected claim to budget ratio (1:10), the principal, years to award, hit ratio and MOIC and IRR for the claimant have been calculated.

Principal is impacted by umbrella clause breach (positively), whilst tenor is reliant on the performance of Fair and Equitable Treatment (FET) – positively, i.e. reducing the years - and Denegation of International Justice (DNJ)- negatively for the claimant, i.e. increasing the yeas.

Consistent results have been obtained. Hit ratio as well as MOIC and IRR are considered from the perspective of the claimant, and are built from the above results. As expected, IRR and MOIC increase as tenor reduces. The hit ratio is larger in the best-case scenarios whilst the relationship of the budget vis-a-vis the hit ratio seems to be inelastic.

FET: 1	Claim Requested : Budget	1:05	1:10	1:50
DNJ: 0	Claimed requested ($)	50.000.000	100.000.000	500.000.000
UC: 1	Budget ($)	10.000.000	10.000.000	10.000.000
Predicted Principal		$37,246,476.96	$60,327,997.67	$184,845,519.3
Predicted Years		6,2	5,9	5,1
Predicted Hit Ratio		74.49%	60.33%	36.97%
Predicted MOIC (claimant)		3,72	6,03	18,48
Predicted IRR (claimant)		50.11%	98.92%	308.01%
FET: 0	Claim Requested : Budget	1:05	1:10	1:50
DNJ: 1	Claimed Requested ($)	50.000.000	100.000.000	500.000.000
UC: 0	Budget ($)	10.000.000	10.000.000	10.000.000
Predicted Principal		$23,331,860.11	$37,790,537.98	$115,790,543.2
Predicted Years		9,5	9,2	8,5
Predicted Hit Ratio		46.66%	37.79%	23.16%
Predicted MOIC (claimant)		2,33	3,78	11,6
Predicted IRR (claimant)		19.35%	36.05%	128.58%

Table 33. Predicted output sensitivity. Several scenarios comprising an increasing ratio budget (expected costs) to claim request in a best-case scenario (FET and UC) and a worst-case scenario (DNJ). Source: own data.

Sensitivity analysis as the budget to expected claim increases: Interestingly, tenors decrease slightly as the expected claim increases. It could be driven by the larger exposure of the case and the need to expedite the case. Once more, it is confirmed that budget is not a drive for hit ratio; but the alleged breaches; hit ratio decreases as the amount claimed increases, reducing the probability of winning, associated with a larger risk, the predicted IRR is larger.

Critical Analysis of the predictive validity and limitations

In terms of assessing the validity of the model vis-à-vis new cases based on historic performance, the expected principal – and total award - and the expected time to award are critical from the perspective of the claimant or its funder.

Facing a new case, a funder would have to undertake a due diligence to review the merits and analyse the risk-reward according to their own assessment criteria. It is sought, in a portfolio approach, that the actual returns from positive cases exceed losses from the investment in negative ones. Diversification is a valuable tool, as in any other asset management discipline.

Arbitration risk addresses ´quasi perfect´ market conditions, in the sense of (i) arbiters are top qualified professionals worldwide, independent and impartial, (ii) cases expose reputation of the parties involved therefore, with increased scrutiny derived from the transparency, and (iii) given the budgets involved, a high degree of professionalism is assumed by all the stake holders – i.a. damages experts, legal firms – whether appointed by the tribunal or the parties.

The claims have been filed by the right claimant, in due course and at the right place and time with the appropriate object (*ratione temporis, ratione personae and ratione materiae*). If the jurisdictional hurdles are overcome, a discussion on substance starts to take place: has there been discriminatory measures? ensured a min standard, provided the investor and investment with protection and security; has there been expropriation (direct or indirect). If so, with or without sufficient compensation. If so, what have been the damages - actual or lost profit -, since when and how is the claimant compensated over time, including the costs of bringing the case to arbitration.

The tool described in this research could be useful in two fronts: for new cases and for reviewing existing ones.

6. Conclusions and suggested next steps

This research introduces a portfolio-approach methodology utilizing historical data to assess litigation risk in international investment arbitration. Recognizing the limitations of our modelling in guaranteeing accuracy or predictive capability, the methodology presents a potential objective proxy, aiming to providing an additional tool for benchmarking Investor-State disputes, with the goal of reducing subjective interpretations.

In our view, the international arbitration field has the essential and minimal elements of any dispute resolution: claim, response, answer, rejoinder, award, a set of rules governing the process that are similar from jurisdiction to jurisdiction and a series of international treaties with a common ground of standards aiming to protect foreign investment and investors. Arbiters are professionals of outstanding academic or professional background, deemed to be impartial and independent in a highly transparent process.

Agnostic of, inter alia, the venue or sector of the claim and the composition of the Tribunal, a combination of technical–legal and econometric analysis has been undertaken with a particular focus of the interrelation between the qualitative and the quantitative variables. Whilst the limitations of the model are not minor, the selection of the regressors and access to data are critical to reach a practical data-driven model that addresses litigation risk based on observable input.

Should the data driven Asset Pricing litigation risk model be validated, the practical implications are three-fold.

On one hand, the model provides an additional tool to the asset manager to value and benchmark litigation cases in terms of amount expected, and risk-reward ratios, Internal Rate of Return (IRR) and Multiple On Invested Capital (MOIC). Addressing the merits, qualitative factors within the model, provides an additional tool to explain the rationale of the outcome, useful for decision making (a) in existing portfolios, to hedge risks – potentially leading to a divestment or a settlement decision - and, (b) for new cases, to set additional quantitative criteria to invest (anticipate returns), ultimately to optimise returns from the perspective of an investor or a funder.

Secondly, the exercise could be useful for claimants, funders - and their auditors - and any stake holder, from legal advisors to insurers, providing third parties an additional tool to validate expectations and fair valuation. Within reason, past information from external cases can add a triggering or valuation event beyond the sale of an asset, an actual award, exhaustion of appeal or any other "milestone factor".

Thirdly, from a records management perspective, the level of subjectivity could be reduced when addressing fair valuation.

___

We have also undertaken a theoretical framework over that limited number of positive cases: what investments would have met the hypothetical criteria of IRR above 20% in a period equal or below 6 years [can this be sorted out in the model]. The percentage above 20% is required in the industry given the inherent litigation risk: difficult to price, subjective by the counterparties, reliant on subjective judgment from a judge, arbiter or tribunal, dependant on changing or potentially changing rules (legislation and precedents). Here are the findings from our sample, based on 89 out of the 94 awards for which we were able to acquire the necessary underlying data, as outlined in the methodology of this research.

Third-Party Funder´s perspective

IRR < 20% and > 6 years	IRR>20% and > 6 years	IRR>20% and ≤ 6 years
34	14	41
38%	16%	46%

Table 34. Summary chart of a theoretical investment criteria >20% IRR and <6 years of time to award. Total # of cases (n=89). Source: own data.

62% of the awards would have yielded above 20% IRR, and below 50% would have met the two criteria. In 38% of cases neither criteria would have been met at all.

Claimant´s perspective		
Dependent variable	***Independent variables Statistical significance at 90% (*), 95% (**) or 90% (***)***	***R^2***
Principal	*Claim Requested (***)* *Costs (***)* *Umbrella Clause (**)*	*80.1%*
Principal = (Claim Requested)^0.696 × (Costs)^0.287 × e^(0.468·UC)		
Years	*Claim Requested (*)* *Costs (***)* *Fair and Equitable Treatment (FET) (**)* *Denial of Justice (DNJ) **	*21.3%*
*YEARS = - 0.445*log_claim_requested + 0.982*log_costs - 1.775*FET + 1.61*DNJ*		

Table 35. Summary Table. Source: Own data.

The conclusions of the research are summarised below:

(1) Regression OLS equations have been obtained to obtain both the expected principal and the time to award based on past information from a series of quantitative and qualitative information; the outcome is no guarantee of future performance or inform about the actual likelihood of new cases, however, provides a further risk management tool.

(2) We have found what we believe is an unprecedented liaison between the legal regressors and the financial variables of an award; this is based on the approach to ´law as data´, following a trend that has been fuelled by recent studies from the academia, and leveraged on a growing litigation finance industry; in particular, we have found that umbrella clause may lead to an increase of the principal to be awarded, whilst the time to award could be shortened when the respondent State has breached the Fair and Equitable Treatment standard, and the process is lengthier where the claimant has evidenced to suffer Denial of International Justice.

(3) We have run simulations on two models that intend to predict principal and time to award based on variables known ex ante (before addressing Court or Tribunal), which are (i) the amount claimed and (ii) the budget or expected investment, and (iii) the set of breaches alleged. Of interest to the claimant and its funder, what would be, based on past information, the Hit Ratio, Multiple On Invested Capital (MOIC) and the Internal Rate of Return (IRR) by the claimant, critical key performance indicators from an asset management perspective. After running sensitivities for the different variables, from budget to expected claim ratio and for different amounts claimed, we have found that the results are consistent.

Legal Finance Analytics – marrying legal, finance and technology – can represent an area of significant growth. Although the contribution to our research has been limited, it is undeniable that through Generative AI, vast amounts of data and alphanumerical records can be accessed and managed in a short period of time with increasing accuracy. New data mining technologies can lead to improved asset classification and subsequent risk management. Retrieval Augmented Generation are useful tools to lever Machine, Deep Learning and Pre-trained Transformers techniques within Large Language Models. Providing the model context through Prompt Engineering and iteration with statistical analysis can contribute as an additional management tool for the litigation finance and insurance industries.

Suggested Next Steps and future work

We would like to highlight several areas of interest within asset pricing litigation risk that are outside the scope of this research, albeit it could drive future lines of research and action.

1- *In connection with the existing model*

a) *Monte Carlo Simulations*

A Monte Carlo simulation – a multiple random sampling and statistical analysis (Raychaudhuri, S. (2008)) – could be further considered, in order to model scenarios and likelihoods of probabilistic forecasts.

Carter, R. C., & Long, R. J. (2009) have used this technique in an early stage of a claim: to assess a contractor´s claim recovery and as an estimation of costs to settle an insurance dispute.

The Monte Carlo simulation would be used in training the litigation risk management model to effectively characterize uncertainty by quantifying and understanding the range of possible outcomes and their associated probabilities, aimed to improve decision-making. Needless to add, the critical matter of accurately representing uncertainty is undertaken by incorporating actual data from legal and financial *corpa*.

b) *Introducing the analysis of negative outcomes analysis and regressors addressing jurisdiction.*

Focus on this research has taken place on positive outcomes, given the availability of the data. Having said this, the theoretical exercise could be extended to cases with negative outcome.

Probability of success = 1 – Probability of Default

The latter can be driven by not overcoming one or more jurisdictional hurdles - *ratione personae, ratione temporis and/or ratione materiae* – or because no breach has been found, despite allegations. For the time being, manual intervention is still needed; having said this, technology could overcome in the very near future the data mining limitations that we have encountered.

The formula of the Expected Value of a claim (Table 8) could be further expanded, adding the jurisdictional (procedural or formal elements) to the substantive law element (breaches). Aiming in the direction of a probability-weighted damages award function would be useful for investment decision process.

c) In terms of expected quantum and data validation

Further granularity on the elements of the award, in particular differentiating damages vs lost profit element, as well as the impact within sectors or geographies, provided amount and quality data available, could incorporate additional information into the modelling validation.

2- Comparison of litigation risk with other asset classes

In order to be transferrable to other asset classes - or subsegments - within litigation risk beyond international arbitration, three requirements seem in order: (i) abundant amount of data gathering, (ii) similar set of legal arguments (whether formal or substantive) within a common framework (rules and applicable law) and (iii) a critical analysis when selecting the regressors.

Steinitz, M., & Field, A. C. (2013) pointed out the similarities of litigation risk with venture capital: "*similar (mid-length) investment timelines; represent pools of investors´ capital, and profitability is measured across a portfolio of investments, not a single investment"*. Pricing litigation risk would lead to conclude that it can be considered formally as an asset class. The probability distribution function of the returns could therefore be benchmarked against other asset classes, for instance in the venture capital space.

3- Risk-management tool - rating, dynamic approach.

The proposed methodology could also be developed with a dynamic approach, given the ´common funnel´ of information within the process (with three instances) (i) Pleading-Defendant's Response; (ii) Plaintiff's Rejoinder-Defendant's Surrejoinder (iii) Judgment/Award. Valuation of litigation as a stochastic process is a recent and thriving trend (i.a. Ahmed, A. (2021), Grenadier, B., & Grenadier, S. R. (2024). Lera, S. C., Mahari, R., & Strub, M. S. (2022)).

7. References

Abramowicz, M., & Alper, O. (2013). Screening legal claims based on third-part litigation finance agreements and other signals of quality. *Vand. L. Rev.*, *66*, 1641.

Abrams, D. S., & Chen, D. L. (2012). A market for justice: A first empirical look at third party litigation funding. *University of Pennsylvania Journal of Business Law*, *15*, 1075.

Aceris Law LLC. (2022, June 26). Denial of Benefits in Investment Arbitration. Last accessed March, 2024, from https://www.acerislaw.com/denial-of-benefits-in-investment-arbitration/.

Aceris Law LLC. (n.d.). Time Limitations and Stale Claims in Investment Arbitration. Aceris Law LLC. Last accessed March, 2024, from https://www.acerislaw.com/time-limitations-and-stale-claims-in-investment-arbitration/.

Aceris Law LLC. (2021, February 18). MFN Clauses in Investment Arbitration. Last accessed March, 2024, from https://www.acerislaw.com/mfn-clauses-in-investment-arbitration/#_ftn3.

Aceris Law LLC. (2022, May 1). Umbrella Clauses in Investment Arbitration. Last accessed March, 2024, from https://www.acerislaw.com/umbrella-clauses-in-investment-arbitration/.

Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pp. 373-383. Spring.

Affaki, G. (2013). A Financing is a Financing is a Financing. *Dossier X: Third-party Funding in International Arbitration, ICC Publishing SA, Paris*.

Agulló, D. A. (2022). Los contratos de financiación de litigios por terceros (third-party funding) en España/Third-party litigation funding contracts in Spain. *Revista de Derecho Civil*, *9*(1), 183-231.

Ahmed, A. (2021). Valuation of Third Party Litigation Finance Contracts using a Real Option Methodology. *Columbia Law and Economics Working Paper, Forthcoming.* https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3836481

Aletras, N., Tsarapatsanis, D., Preoţiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective. *PeerJ Computer Science*, *2*, e93.

Alexander, C. S. (2023). Legal Tech, Litigation, and the Adversarial System. In D. F. Engstrom (Ed.), *Legal Tech and the Future of Civil Justice*. Cambridge University Press. Last accessed March, 2024, from https://www.cambridge.org/core/books/legal-tech-and-the-future-of-civil-justice/litigation-outcome-prediction-access-to-justice-and-legal-endogeneity/E489F1E681BC0EF05707242FDA4E22FD#A-sec-37.

Alvarez, B. (Author). (2024, February 13). *Tutorial video: Minimum Standard of Treatment (MST)*. Last accessed March, 2024, from https://jusmundi.com/en/document/publication/en-minimum-standard-of-treatment-mst.

Alschner, W., & Charlotin, D. (2018). The growing complexity of the International Court of Justice's self-citation network. *European Journal of International Law*, *29*(1), 83-112.

Ash, E., Chen, D. L., & Galletta, S. (2022). Measuring Judicial Sentiment: Methods and Application to US Circuit Courts. *Economica, 89*(354), 362-376.

Ashley, K. D. (2019). A brief history of the changing roles of case prediction in AI and law. *Law Context: A Socio-Legal J.*, *36*, 93.

Ashley, K. D., & Brüninghaus, S. (2009). Automatically classifying case texts and predicting outcomes. *Artificial Intelligence and Law*, *17*, 125-165.

Atkinson, K., Bench-Capon, T., & Bollegala, D. (2020). Explanation in AI and law: Past, present and future. *Artificial Intelligence, X*(Y), 103387.

Avraham, R., & Sebok, A. (2018). An empirical investigation of third party consumer litigant funding. *Cornell Law Review*, *104*, 1133.

Barros, R., Peres, A., Lorenzi, F., Krug Wives, L., & Hubert da Silva Jaccottet, E. (2018). Case law analysis with machine learning in Brazilian court. In *Recent Trends and Future Technology in Applied Intelligence: 31st International Conference on Industrial Engineering and Other Applications of Applied Intelligent*

Systems, IEA/AIE 2018, Montreal, QC, Canada, June 25-28, 2018, Proceedings 31, pp. 857-868. Springer International Publishing.

Bedi, S., & Marra, W. C. (2021). The Shadows of Litigation Finance. *Vand. L. Rev.*, *74*, 563.

Bench-Capon, T., Araszkiewicz, M., Ashley, K., Atkinson, K., Bex, F., Borges, F., ... & Wyner, A. Z. (2012). A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law. *Artificial Intelligence and Law*, *20*, 215-319.

Bertalan, V. G. F., & Ruiz, E. E. S. (2020, March). Predicting Judicial Outcomes in the Brazilian Legal System Using Textual Features. In *DHandNLP@PROPOR*, pp. 22-32.

Bloomberg News. (2020, June). Hedge Funds Have Turned Lawsuit Bets into a $39 Billion Industry. BNN Bloomberg. Last access March, 2024, from https://www.bnnbloomberg.ca/hedge-funds-have-turned-lawsuit-bets-into-a-39-billion-industry-1.1681919.

Blundell, M. (2022). What are the Chances?: The Predictive Analytics behind Third-Party Litigation Funding in Investment Arbitration.

Board, F. S. (2019). Global monitoring report on non-bank financial intermediation 2023. *Financial Stability Board*.

Boniol, P., Panagopoulos, G., Xypolopoulos, C., Hamdani, R. E., Amariles, D. R., & Vazirgiannis, M. (2020). Performance in the courtroom: Automated processing and visualization of appeal court decisions in France. *arXiv preprint* arXiv:2006.06251.

Branting, L. K., Pfeifer, C., Brown, B., Ferro, L., Aberdeen, J., Weiss, B., ... & Liao, B. (2021). Scalable and explainable legal prediction. *Artificial Intelligence and Law, 29*(2), 213-238.

Burford Capital Limited. (2023, May 16). Burford Capital reports full-year 2022 financial results; profitable growth and update on fair value accounting. Retrieved March, 2024, from https://investors.burfordcapital.com/news/news-details/2023/Burford-Capital-reports-full-year-2022-financial-results-profitable-growth-and-update-on-fair-value-accounting/default.aspx.

Burford Capital. (2023a). Annual Report for the year ended December 31, 2022 [*Form 10-K*]. Retrieved from Securities and Exchange Commission website. Last accessed March, 2024, from https://www.sec.gov/Archives/edgar/data/1714174/000171417423000112/tmb-20230913xex99d1.htm.

Burford Capital. (2023b). Interim Report for the six months ended June, 2023 [*Form 10-Q*]. Retrieved from Securities and Exchange Commission website. Last accessed March, 2024, from https://www.sec.gov/Archives/edgar/data/1714174/000171417423000112/tmb-20230913xex99d1.htm.

Burford Capital. (2023c). Burford Capital Reports Nine Month and Third Quarter 2023 Financial Results: Nine Month Consolidated Total Revenues Increased Fivefold to $794 Million. *Burford Capital Investors*. Retrieved from https://investors.burfordcapital.com/news/news-details/2023/BURFORD-CAPITAL-REPORTS-NINE-MONTH-AND-THIRD-QUARTER-2023-FINANCIAL-RESULTS-NINE-MONTH-CONSOLIDATED-TOTAL-REVENUES-INCREASED-FIVEFOLD-TO-794-MILLION/default.aspx. Last accessed March, 2024.

Burford Capital. (2024a, March 14). Burford Capital Reports Record 2023 Results. *Burford Capital Insights & News.* Retrieved from https://www.burfordcapital.com/insights-news-events/news-press-releases/burford-capital-reports-record-2023-results/. Last accessed March, 2024.

Burford Capital. (2024b, March 14). *Shareholder Letter* (p. xi). Retrieved from https://investors.burfordcapital.com/financials/Regulatory-News/regulatory-news-details/2024/Burford-Capital-Ltd---Annual-results-for-year-ended-December-31-2023/default.aspx. Last accessed April, 2024.

Burford Capital. (2024c, March 14). Full year 2023 and fourth quarter financial results presentation [PowerPoint slides]. Retrieved from https://investors.burfordcapital.com/events-and-presentations/default.aspx. Last accessed April, 2024.

Chalkidis, I., Androutsopoulos, I., & Aletras, N. (2019). Neural legal judgment prediction in English. *arXiv preprint arXiv:1906.02059*.

Carter, R. C., & Long, R. J. (2009). Applications of Monte Carlo simulations in dispute resolution and claims work. *AACE International Transactions*. Retrieved from Business Source Complete database.

Chauhan, N. K., & Singh, K. (2018, September). A review on conventional machine learning vs deep learning. In *2018 International Conference on Computing, Power and Communication Technologies (GUCON)*, pp. 347-352. IEEE.

Chen, S. X., & Hough, K. (2019, May). Researching third-party funding in investor-state dispute settlement. *New York University Law Global*. Retrieved from https://www.nyulawglobal.org/globalex/Third-Party_Funding_InvestorState_Dispute_Settlement.html. Last accessed in March, 2024.

Cohen, M. C., Dahan, S., Khern-Am-Nuai, W., Shimao, H., & Touboul, J. (2023). The use of AI in legal systems: determining independent contractor vs. employee status. *Artificial intelligence and law*, 1-30.

Cooter, R. D. (1991). Economic theories of legal liability. *Journal of Economic Perspectives, 5*(3), pp. 11-30.

Cooter, R., & Ulen, T. (2011). An economic theory of the legal process. In *Law and economics* (pp. 382-418). Addison-Wesley.

Cooter, R. D., & Rubinfeld, D. L. (1989). Economic analysis of legal disputes and their resolution. *Journal of economic literature*, 27(3), 1067-1097.

Cornell, B. (1990). The incentive to sue: An option-pricing approach. *The Journal of Legal Studies, 19*(1), 173-187.

Cremades, B. M. (2013). Third party funding in international arbitration. *Romanian Arbitration Journal, 7*(1), 1.

Dafe, F., & Williams, Z. (2021). Banking on courts: financialization and the rise of third-party funding in investment arbitration. *Review of international political economy*, *28*(5), 1362-1384.

Dahan, S. (2020). Analytics and EU Courts: The Case of Trademark Disputes. *Samuel Dahan and others,'Analytics and EU Courts: The Case of Trademark Disputes' in Tamara Capeta and Iris Goldner Lang (eds), The*

Changing European Union: A critical view on the role of the Courts (Hart Publishing 2022). This chapter is a drawn from a report commissioned by the EUIPO.

Davitti, D., & Vargiu, P. (2023). Litigation Finance and the Assetization of International Investment Arbitration. *The Journal of World Investment & Trade, 24*(3), 487-500.

De Morpurgo, M. (2011). A comparative legal and economic approach to third-party litigation funding. *Cardozo J. Int'l & Comp. L.*, *19*, 343.

Dhami, M. K., & Mandel, D. R. (2022). Communicating uncertainty using words and numbers. *Trends in Cognitive Sciences*.

Dias, J., Santos, P. A., Cordeiro, N., Antunes, A., Martins, B., Baptista, J., & Gonçalves, C. (2022). State-of-the-art in Artificial Intelligence applied to the Legal Domain. *arXiv preprint arXiv:2204.07047*.

Diehl, A. (2012). *The core standard of international investment protection: Fair and equitable treatment.* Kluwer Law International, p. 369.

Dumberry, P. (2013). *The Fair and Equitable Treatment Standard: A Guide to NAFTA Case Law on Article 1105.*

Dumberry, P. (2017). The Importation of the FET Standard through MFN Clauses: An Empirical Study of BITs. *ICSID Review-Foreign Investment Law Journal, 32*(1), 116-137.

Dunn, M., Sagun, L., Şirin, H., & Chen, D. (2017, June). Early predictability of asylum court decisions. In *Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law*, 233-236.

DW. (2023, April 17). EU exit from climate-killing energy treaty looms. *DW*. Retrieved from https://www.dw.com/en/eu-exit-from-climate-killing-energy-treaty-looms/a-65323709 (Accessed March, 2024).

Elnaggar, A., Otto, R., & Matthes, F. (2018). Named-entity linking using deep learning for legal documents: a transfer learning approach. *arXiv preprint arXiv:1810.06673*.

Fei, Z., Shen, X., Zhu, D., Zhou, F., Han, Z., Zhang, S., ... & Ge, J. (2023). LawBench: Benchmarking Legal Knowledge of Large Language Models. *arXiv preprint arXiv:2309.16289*.

Fore, J. (2019). A court would likely (60-75%) find-defining verbal probability expressions in predictive legal analysis. *Legal Communication & Rhetoric: JAWLD, 16*, 49.

Franck, S. D., & Wylie, L. E. (2015). Predicting outcomes in investment treaty arbitration. *Duke Law Journal*, *65*, 459.

Frankenreiter, J., & Livermore, M. A. (2020). Computational methods in legal analysis. *Annual Review of Law and Social Science*, *16*, 39-57.

Ethical implications of third-party funding in international arbitration. *Arbitration International*, *32*(3), 505-522.

Grenadier, B., & Grenadier, S. R. (2024). *The valuation of a dynamic litigation process: The lawsuit as a (real) option on an option*. Stanford University Graduate School of Business Research Paper No. 4566390. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4566390

Goeler, J. V. (2016). *Third-party funding in international arbitration and its impact on procedure.* Kluwer Law Internation

Goral, R. (2015). Justice dealers: The ecosystem of American litigation finance. *Stanford Journal of Law, Business & Finance, 21,* 98.

Guha, N., Nyarko, J., Ho, D. E., Ré, C., Chilton, A., Narayana, A., ... & Li, Z. (2023). Legalbench: A collaboratively built benchmark for measuring legal reasoning in large language models. *arXiv preprint arXiv:2308.11462*.

Guo, M. (2021). Internet court's challenges and future in China. *Computer Law & Security Review*, *40*, 105522.

Hartung, M., Bues, M. M., & Halbleib, G. (2017). *Legal tech.* Munich: CH Beck.

Hederer, C. (2016). *Third-Party Funding in International Investment Arbitration–Economic Analysis and Regulatory Options.* Available at SSRN 2952488.

Hodgson, M., Kryvoi, Y., & Hrcka, D. (2021). *Costs, Damages and Duration in Investor-State Arbitration.* British Irish Institute of International and Comparative Law and Allen & Overy, London.

Hodges, C., Vogenauer, S., & Tulibacka, M. (Eds.). (2010). *The costs and funding of civil litigation: a comparative perspective*. Bloomsbury Publishing.

Hodges OBE, C., Peysner, J., & Nurse, A. (2012). *Litigation funding: Status and issues.*

Hu, Z., Li, X., Tu, C., Liu, Z., & Sun, M. (2018, August). Few-shot charge prediction with discriminative legal attributes. *Proceedings of the 27th International Conference on Computational Linguistics, 487-498.*

Hussain, M. A., Labanieh, M. F., Mahdzir, N., Sulaiman, N., & Bawazir, O. S. A. (2023). The potential prospect of Artificial Intelligence (AI) in arbitration from the International, National and Islamic perspectives. *Journal of International Studies*, *19*(1), 95-122.

International Arbitration. (2021, August 11). ISDS. Retrieved October, 2023, from https://www.international-arbitration-attorney.com/arbitration/isds/page/3/. Last access in March, 2024.

International Arbitration Attorney. (2022, January 23). International Arbitration. International Arbitration Attorney. Retrieved from https://www.international-arbitration-attorney.com/arbitration/isds/page/3/. Last access in March, 2024.

International Arbitration Attorney. (2022, March 13). International Arbitration. International Arbitration Attorney. Retrieved from https://www.international-arbitration-attorney.com/arbitration/isds. Last access in March, 2024.

International Bar Association. (2024, February). IBA Guidelines on Conflicts of Interest in International Arbitration. Retrieved from https://www.ibanet.org/document?id=Guidelines-on-Conflicts-of-Interest-in-International-Arbitration-2024.

International Centre for Settlement of Investment Disputes. (2022). Post Award Remedies. Retrieved from https://icsid.worldbank.org/procedures/arbitration/convention/post-award-remedies/2022 Last access in March, 2024.

International Centre for Settlement of Investment Disputes. (2021). *Lion Mexico Consolidated LP v. United Mexican States, ICSID Case No. ARB(AF)/15/2, Award dated 20 September 2021, para. 205.*

International Legal Finance Association. (2023, June). Resourcing the Rule of Law in Europe [Full report]. Retrieved from https://uploads-ssl.webflow.com/5ef44d9ad0e366e4767c9f0c/6498d39ac7b645dcdc5b9a4d_Resourcing%20the%20Rule%20of%20Law%20in%20Europe_Full%20report.pdf. Last access in March, 2024.

Ishikawa, T. (2021, June 25). Funders' pricing and the real value of litigation risks [Blog post]. Practical Law Dispute Resolution Blog. Retrieved from http://disputeresolutionblog.practicallaw.com/funders-pricing-and-the-real-value-of-litigation-risks/. Last access in March, 2024.

Islam, R. (2019). Does Fair and Equitable Treatment (FET) Standard in International Investment Treaties Create a Customary Rule of International Law?. *Dhaka Univ. Stud. Part F*, *30*, 27.

Islam, R., Islam, & Ghosh. (2018). *Fair and Equitable Treatment (FET) Standard in International Investment Arbitration*. Singapore: Springer.

Jackson, R. M. (2010). Review of civil litigation costs (Vol. 2). The Stationery Office.

Jus Mundi. (n.d.). Minimum Standard of Treatment (MST). Jus Mundi. https://jusmundi.com/en/document/publication/en-minimum-standard-of-treatment-mst (Accessed March, 2024).

Kaddour, J., Harris, J., Mozes, M., Bradley, H., Raileanu, R., & McHardy, R. (2023). Challenges and applications of large language models. *arXiv preprint arXiv:2307.10169*.

Pearson, K. (1905). The problem of the random walk. *Nature*, *72*(1865), 294-294.

Kaplow, L., & Shavell, S. (2002). Economic analysis of law. In *Handbook of Public Economics* (Vol. 3, pp. 1661-1784). Elsevier.

Kashefi, A., & Mukerji, T. (2023). ChatGPT for programming numerical methods. *Journal of Machine Learning for Modeling and Computing, 4*(2).

Katz, D. M., Bommarito, M. J., & Blackman, J. (2017). A general approach for predicting the behavior of the Supreme Court of the United States. *PLoS ONE (Public Library of Science ONE), 12*(4), e0174698.

Khouri, S., & Hurford, K. (2012). Third party funding in international arbitration. *Practical Law for Companies*, (6), 35.

Kidd, J. (2011). To fund or not to fund: The need for second-best solutions to the litigation finance dilemma. *Journal of Law, Economics, and Policy, 8,* 613.

Kinnear, M., & Fischer, G. R. (Eds.). (2015). *Building International Investment Law: The First 50 Years of ICSID*. Kluwer Law International BV.

Kluwer Arbitration Blog. (2024, January 13). *2023 Year in Review: Latin America and Investment Arbitration*. Retrieved from https://arbitrationblog.kluwerarbitration.com/2024/01/13/2023-year-in-review-latin-america-and-investment-arbitration/ (Accessed March 25, 2024).

Kohlmeier, T. (2017). Litigation Funding and Predictive Analytics in Mania, K. (2023). Legal technology: Assessment of the legal tech industry's potential. *Journal of the Knowledge Economy*, *14*(2).

Kolanovic, M., & Krishnamachari, R. (2017). Big data and AI strategies: Machine learning and alternative data approach to investing. *JP Morgan Global Quantitative & Derivatives Strategy Report*, *25*.

Kowsrihawat, K., Vateekul, P., & Boonkwan, P. (2018, October). Predicting judicial decisions of criminal cases from Thai Supreme Court using bi-directional GRU with attention mechanism. In *2018 5th Asian Conference on Defense Technology (ACDT)*, 50-55. IEEE.

Lage-Freitas, A., Allende-Cid, H., Santana, O., & Oliveira-Lage, L. (2022). Predicting Brazilian court decisions. *PeerJ Computer Science*, *8*, e904.

Lai, J., Gan, W., Wu, J., Qi, Z., & Yu, P. S. (2023). Large Language Models in Law: A Survey. *arXiv preprint arXiv:2312.03718*.

Lawtomated. (2019, April 18). A.I. Technical Machine vs. Deep Learning. Retrieved from https://lawtomated.com/a-i-technical-machine-vs-deep-learning/, cit https://www.quora.com/What-is-the-difference-between-deep-learning-and-usual-machine-learning. Last accessed March, 2024.

Le Picard, H. (2022). MORAL HAZARD. A FINANCIAL, LEGAL, AND ECONOMIC PERSPECTIVE. *Politique Etrangere*, (2), 204-205.

Lebeuf, C. R. (2018). *A taxonomy of software bots: towards a deeper understanding of software bot characteristics* (Doctoral dissertation).

Langford, M., Behn, D., & Lie, R. (2020). Computational Stylometry: Predicting the Authorship of Investment Arbitration Awards. *Computational Legal Studies: The Promise and Challenge of Data-Driven Research*, 53.

Lera, S. C., Mahari, R., & Strub, M. S. (2022). *Litigation finance at trial: Model and data*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4091716

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.

Lex Anastasiana (506 AD). Law 22, XXXV, Book 4º of the Corpus Iuris Civilis Code, as cited in Spain Supreme Court decision 829/2004 (13th July); confirmed by Justinian, as cited in Spanish Supreme Court decision 976/2008 (31st October).

Li, S., Zhang, H., Ye, L., Guo, X., & Fang, B. (2019). Mann: A multichannel attentive neural network for legal judgment prediction. *IEEE Access*, *7*, 151144-151155.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, *55*(9), 1-35.

Litigation Finance Insider. (n.d.). [*Requires subscription to Litigation Finance Insider newsletter*]. Retrieved from https://litigationfinanceinsider.com/. Last accessed March, 2024.

Loevinger, L. (1963). Jurimetrics: The methodology of legal inquiry. *Law and contemporary problems*, *28*(1), 5-35.

Long, S., Tu, C., Liu, Z., & Sun, M. (2019). Automatic judgment prediction via legal reading comprehension. In *Chinese Computational Linguistics: 18th China National Conference, CCL 2019, Kunming, China, October 18–20, 2019, Proceedings 18* (pp. 558-572). Springer International Publishing.

Lu, M. (2005). The New York convention on the recognition and enforcement of foreign arbitral awards: Analysis of the seven defenses to oppose enforcement in the United States and England. *Arizona Journal of International and Comparative Law*, *23*, 747.

Luk, M. (2023, September). Generative AI: Overview, economic impact, and applications in asset management. *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.4574814.

Ma, L., Zhang, Y., Wang, T., Liu, X., Ye, W., Sun, C., & Zhang, S. (2021, July). Legal judgment prediction with multi-stage case representation learning in the real court setting. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 993-1002).

Mackaay, E., & Robillard, P. (1974). Predicting Judicial Decisions: The Nearest Neighbor Rule and Visual Representation of Case Patterns. Datenverarbeitung im Recht, 3, 302.

Marrow, P. B., Karol, M., & Kuyan, S. (2020). *Artificial Intelligence and Arbitration: The Computer as an Arbitrator—Are We There Yet?*

McConnell, D. J., Zhu, J., Pandya, S., & Aguiar, D. (2021, June). Case-level prediction of motion outcomes in civil litigation. In Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law, 99-108.

McShane, B. B., Watson, O. P., Baker, T., & Griffith, S. J. (2012). Predicting securities fraud settlements and amounts: a hierarchical Bayesian model of federal securities class action lawsuits. *Journal of Empirical Legal Studies*, *9*(3), 482-510.

Medvedeva, M., Vols, M., & Wieling, M. (2020). Using machine learning to predict decisions of the European Court of Human Rights. *Artificial Intelligence and Law*, *28*, 237-266.

Merlone, U., & Lupano, M. (2022). Third party funding: The minimum claim value. *European Journal of Operational Research*, *296*(2), 738-747.

Molot, J. T. (2009). A market in litigation risk. University of Chicago Law Review, 76, 367.

Molot, J. T. (2009). A Market Approach to Litigation Accuracy. *Searle Center, Northwestern University Law School, Third Party Financing of Litigation, September*, 24-25.

Molot, J. T. (2014). The Feasibility of Litigation Markets. *Indiana. Law Journal*, *89*, 171.

Molot, J. T. (2010). Litigation Finance: A Market Solution to a Procedural Problem. *Geo. Law Journal*, *99*, 65.

Morgan, C. (2020, February 28). Inside Arbitration: Data Analytics in International Arbitration: Balancing Technology with The Human Touch. *Herbert Smith Freehills*. Retrieved from https://www.herbertsmithfreehills.com/insights/2020-02/inside-arbitration-data-analytics-in-international-arbitration-balancing-technology-with-the.

Mourre, A., & Vagenheim, A. (2010). Some comments on denial of justice in public and private international law after Loewen and Saipem. *Liber Amicorum Bernardo Cremades. Wolters Kluwer, Alphen aan den Rijn*, 843-866.

Moya Fernández, A. J., Pérez-Pujazón, M. E., & Trigo Sierra, E. (2016). Cesión de créditos y cuestiones prácticas de interés: retracto de crédito litigioso y titulización. *Cesión de créditos y cuestiones prácticas de interés: retracto de crédito litigioso y titulización*, 48-64.

Müller, D., Fuchs, A., & Pika, M. (Eds.). (2024, March 4). *Full Protection and Security (FPS)*. Retrieved March, 2024, from https://jusmundi.com/en/document/publication/en-full-protection-and-security-fps.

Mumcuoğlu, E. (2022). *Prediction of Outcomes in Higher Courts of Turkey Using Natural Language Processing* (Doctoral dissertation, Bilkent Universitesi (Turkey)).

Newcombe, A. P., & Paradell, L. (2009). *Law and practice of investment treaties: standards of treatment*. Kluwer Law International BV.

Nieuwveld, L. B., & Sahani, V. S. (2016). *Third-party funding in international arbitration*. Kluwer Law International BV.

Niklaus, J., Chalkidis, I., & Stürmer, M. (2021). Swiss-judgment-prediction: A multilingual legal judgment prediction benchmark. *arXiv preprint arXiv:2110.00806*.

Ortolani, P., Janssen, A., & Wolters, P. (Eds.). (2022). *International Arbitration and Technology*. Kluwer Law International BV.

Osborne Clarke. (n.d.). Managing risk: Using third party funding to pay the costs of litigation and international arbitration. Retrieved March, 2024, from

https://www.osborneclarke.com/insights/managing-risk-using-third-party-funding-to-pay-the-costs-of-litigation-and-international-arbitration.

European Parliament resolution of 13 September 2022 with recommendations to the Commission on Responsible private funding of litigation (2020/2130(INL))

Paisley, K., & Sussman, E. (2018). Artificial intelligence challenges and opportunities for international arbitration. *New York Dispute Resolution Lawyer*, *11*(1), 35-40.

Panda, S. (2023). Enhancing PDF interaction for a more engaging user experience in library: Introducing ChatPDF. *IP Indian Journal of Library Science and Information Technology, 8*(1), 20-25. ttps://doi.org/10.18231/j.ijlsit.2023.004.

Paparinskis, M. (2013). *The international minimum standard and fair and equitable treatment*. OUP Oxford.

Paulsson, J. (2005). *Denial of justice in international law* (Vol. 17). Cambridge University Press.

Pereira, J., Assumpcao, A., Trecenti, J., Airosa, L., Lente, C., Cléto, J., ... & Lotufo, R. (2024). INACIA: Integrating Large Language Models in Brazilian Audit Courts: Opportunities and Challenges. *arXiv preprint arXiv:2401.05273*.

Pérez López, J. Á., Moya Fernández, A. J., & Trigo y Sierra, E. (2012). Cuestiones prácticas de las ventas de carteras de créditos. *Actualidad jurídica Uría Menéndez, 33*, 45-62.

Pinsolle, P. (2011). Le financement de l'arbitrage par les tiers. *Revue de l'Arbitrage*, *2011*(2).

Polinsky, A. M., & Shavell, S. (1997). Punitive damages: An economic analysis. *Harv. L. Rev.*, *111*, 869.

Popp, A. T. (2019). Federal Regulation of Third-Party Litigation Finance. *Vand. L. Rev.*, *72*, 727.

Posner, R. A. (2014). *Economic analysis of law*. Aspen Publishing.

Prabhu, A. (2023, February 15). Spain's Renewable Energy Disputes: Renewable Energy Needs Reliable Arbitration [Blog post]. *The Arbitration Brief*. Retrieved from https://thearbitrationbrief.com/2023/02/15/spains-renewable-energy-disputes-renewable-energy-needs-reliable-arbitration/.

R (on the application of PACCAR Inc and others) (Appellants) v Competition Appeal Tribunal and others (Respondents) [2023] UKSC 28

Rationalstat LLC. (2023, August 9). Global Litigation Funding Investment Market is poised to touch US$ 24.3 billion by the end 2028, driven by increasing awareness about litigation financing. *GlobeNewswire.* Retrieved from https://www.globenewswire.com/en/news-release/2023/08/09/2721731/0/en/Global-Litigation-Funding-Investment-Market-is-poised-to-touch-US-24-3-billion-by-the-end-2028-driven-by-increasing-awareness-about-litigation-financing.html.

Raychaudhuri, S. (2008, December). Introduction to monte carlo simulation. In *2008 Winter simulation conference* (pp. 91-100). IEEE.

Research Nester. (n.d.). Litigation Funding Investment Market. Retrieved from https://www.researchnester.com/reports/litigation-funding-investment-market/2800.

Rhee, R. J. (2006). A Price Theory of Legal Bargaining: An Inquiry into the Selection of Settlement and Litigation Under Uncertainty. *Emory Law Journal*, *56*, 619.

Rhim, Y. Y., & Park, K. (2023). The Artificial Intelligence in International Law. In *Revolutionary Approach to International Law: The Role of International Lawyer in Asia,* 215-237. Singapore: Springer Nature Singapore.

Rissland, E. L., Ashley, K. D., & Loui, R. P. (2003). AI and Law: A fruitful synergy. *Artificial Intelligence*, *150*(1-2), 1-15.

Rowles-Davies, N., & Cousins, J. (2014). *Third party litigation funding.*

Salomon, C. T., & Yamamoto, H. (2020, June 9). When the Statute of Limitations To Enforce a Foreign Arbitral Award in the United States Has Lapsed: Enforcing Foreign Arbitral Awards as Foreign Judgments in the United States. *New York Law Journal.* Retrieved from https://www.law.com/newyorklawjournal/2020/06/09/when-the-statute-of-limitations-to-enforce-a-foreign-arbitral-award-in-the-united-states-has-lapsed-enforcing-foreign-arbitral-awards-as-foreign-judgments-in-the-united-states.

Savelka, J., Ashley, K. D., Gray, M. A., Westermann, H., & Xu, H. (2023). Explaining Legal Concepts with Augmented Large Language Models (GPT-4). *arXiv preprint* arXiv:2306.09525.

Savelka, J., Ashley, K. D., Gray, M. A., Westermann, H., & Xu, H. (2023). Can GPT-4 Support Analysis of Textual Data in Tasks Requiring Highly Specialized Domain Expertise? *arXiv preprint arXiv:2306.13906.*

Schilder, F. (2023). *Legal Expertise Meets Artificial Intelligence: A Critical Analysis of Large Language Models as Intelligent Assistance Technology.*

Scherer, M., Goldsmith, A., & Flechet, C. (2012). Third party funding in international arbitration in Europe: Part 1-funders' perspectives. *International Business Law Journal, 207*.

Sebok, A. J. (2010). Betting on Tort Suits After the Event: From Champerty to Insurance. *DePaul Law Review*, *60*, 453.

Seidel, S. (2013). Third-party investing in international arbitration claims To invest or not to invest? A daunting question. In *Third-party Funding in International Arbitration.* Paris, France: ICC Publishing SA.

Sert, M. F., Yıldırım, E., & Haşlak, İ. (2022). Using artificial intelligence to predict decisions of the turkish constitutional court. *Social Science Computer Review*, *40*(6), 1416-1435.

Shabsigh, G., & Boukherouaa, E. B. (2023). Generative Artificial Intelligence in Finance. *FinTech Notes*, *2023*(006).

Shalaby, A. G., Abdelaziz, G. M., & Kandeel, M. E. (2022, November). Using Artificial Intelligence to Resolve Disputes through Online Arbitration. In *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)* (pp. 1-8). IEEE.

Shepherd, J. M., & Stone, J. E. (2015). Economic conundrums in search of a solution: The functions of third-party litigation fiance. Arizona State Law Journal, 47, 919.

Shulayeva, O., Siddharthan, A., & Wyner, A. (2017). Recognizing cited facts and principles in legal judgements. *Artificial Intelligence and Law*, *25*(1), 107-126.

Sokhansanj, B. A., & Rosen, G. L. (2022). Predicting Institution Outcomes for Inter Partes Review (IPR) Proceedings at the United States Patent Trial & Appeal

Board by Deep Learning of Patent Owner Preliminary Response Briefs. *Applied Sciences*, *12*(7), 3656.

Solas, G. M. (2019). *Third party funding*. Cambridge University Press.

Stadler, A. (2016). Funding of mass claims in Germany: Caught between a rock and a hard place?. In *Litigation, Costs, Funding and Behaviour,* 217-237. Routledge.

Steinitz, M. (2013). How Much Is That Lawsuit in the Window; Pricing Legal Claims. *Vand. L. Rev.*, *66*, 1889.

Steinitz, M., & Field, A. C. (2013). A model litigation finance contract. *Iowa L. Rev.*, *99*, 711.

Steinitz, M. (2014). Incorporating legal claims. *Notre Dame L. Rev.*, *90*, 1155.

Strom, R. (2020, June 11). Nobody knows litigation finance size, but it's not $85 billion. *Bloomberg Law.* Retrieved from https://news.bloomberglaw.com/business-and-practice/nobody-knows-litigation-finance-size-but-its-not-85-billion.

Strickson, B., & De La Iglesia, B. (2020, March). Legal judgement prediction for UK courts. In *Proceedings of the 3rd International Conference on Information Science and Systems,* 204-209.

Stroble, J. J., & Welikson, L. (2020). Third-Party Litigation Funding: A Review of Recent Industry Developments. *Defense Counsel Journal*, *87*, 1.

Sulea, O. M., Zampieri, M., Malmasi, S., Vela, M., Dinu, L. P., & Van Genabith, J. (2017). Exploring the use of text classification in the legal domain. *arXiv preprint arXiv:1710.09306*.

Sun, J., Huang, S., & Wei, C. (2024). Chinese legal judgment prediction via knowledgeable prompt learning. *Expert Systems with Applications*, *238*, 122177.

Surdeanu, M., Nallapati, R., Gregory, G., Walker, J., & Manning, C. D. (2011, June). Risk analysis for intellectual property litigation. In *Proceedings of the 13th International Conference on Artificial Intelligence and Law,* 116-120.

Sweify, M. F. (2023). *Third Party Funding in International Arbitration: A Critical Appraisal and Pragmatic Proposal*. Edward Elgar Publishing.

Tan, J., Westermann, H., & Benyekhlef, K. (2023). Chatgpt as an artificial lawyer. *Artificial Intelligence for Access to Justice (AI4AJ 2023)*.

Tarissan, F., & Nollez-Goldbach, R. (2016). Analysing the first case of the international criminal court from a network-science perspective. *Journal of Complex Networks*, *4*(4), 616-634.

Tippett, E. C., Alexander, C. S., Branting, K., Morawski, P., Balhana, C., Pfeifer, C., & Bayer, S. (2021). Does Lawyering Matter? Predicting Judicial Decisions from Legal Briefs, and What That Means for Access to Justice. *Texas Law Review*, *100*, 1157.

World Investment Report - United Nations Conference on Trade https://unctad.org/publication/world-investment-report-2023 accessed 25th October 2023

United Nations Conference on Trade and Development. (2023). World Investment Report 2023. UNCTAD. Retrieved from https://unctad.org/publication/world-investment-report-2023. Last access in March, 2024.

United Nations Conference on Trade and Development. (n.d.-a). Investment Dispute Settlement. UNCTAD Investment Policy Hub. Retrieved from https://investmentpolicy.unctad.org/investment-dispute-settlement. Last access in March, 2024.

United Nations Commission on International Trade Law. (n.d.-b). Investor-State Working Group. UNCITRAL. Retrieved from https://uncitral.un.org/en/working_groups/3/investor-state. Last access in March, 2024.

United Nations Conference on Trade and Development (UNCTAD). (n.d.-c). Enel Fortuna S.p.A. v. Republic of Panama. Investment Policy Hub. Retrieved March, 2024, from https://investmentpolicy.unctad.org/investment-dispute-settlement/cases/956/enel-fortuna-v-panama.

United Nations Conference on Trade and Development (UNCTAD). (n.d.-d). Investment dispute settlement. Investment Policy Hub. Retrieved March, 2024, from https://investmentpolicy.unctad.org/investment-dispute-settlement?status=1.

Van Boom, W. H. (2011). Third-Party Financing in International Investment Arbitration. Available at SSRN 2027114.

Van Boom, W. H. (Ed.). (2016). Litigation, costs, funding and behaviour: implications for the law. Taylor & Francis.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Veljanovski, C. (2011). *Third Party Litigation Funding in Europe.* Journal of Law, Economics, and Policy, 8, 405.

Virtucio, M. B. L., Aborot, J. A., Abonita, J. K. C., Avinante, R. S., Copino, R. J. B., Neverida, M. P., ... & Tan, G. B. A. (2018, July). Predicting decisions of the philippine supreme court using natural language processing and machine learning. In *2018 IEEE 42nd annual computer software and applications conference (COMPSAC),* Vol. 2, 130-135. IEEE.

Voss, J. (2022). Report with recommendations to the Commission on Responsible Private Funding of Litigation (VOSS REPORT).

Waihenya, J. (2021). Third Party Funding and Investment Arbitration–The Ménage À Trois Conundrum of International Arbitration? *Alternative Dispute Resolution*, 163.

Wang, C., Liu, X., Yue, Y., Tang, X., Zhang, T., Jiayang, C., ... & Zhang, Y. (2023). Survey on Factuality in Large Language Models: Knowledge, Retrieval and Domain-Specificity. *arXiv preprint arXiv:2310.07521*.

Waqar, M. (2022). *The Use of Artificial Intelligence in Arbitral Proceedings.* Ohio State Journal on Dispute Resolution, 37, 345.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, *35*, 24824-24837.

Westermann, H., Savelka, J., & Benyekhlef, K. (2023). Llmediator: Gpt-4 assisted online dispute resolution. *arXiv preprint arXiv:2307.16732*.

Westermann, H., Meeùs, S., Godet, M., Troussel, A., Tan, J., Savelka, J., & Benyekhlef, K. (2023). *Bridging the gap: Mapping layperson narratives to legal issues with language models*.

Westfleet Advisors. (2023). The Westfleet Insider 2022 Litigation Finance Market Report. Retrieved from https://www.westfleetadvisors.com/wp-content/uploads/2023/02/WestfleetInsider-2022-Litigation-Finance-Market-Report.pdf. Last access in March, 2024.

Wilinski, P., & Durbas, M. (2021). Data mining, text analytics and international commercial arbitration. *Forthcoming in: Pietro Ortolani et al.(eds.): International Arbitration and Technology, Wolters Kluwer*.

WolframAlpha. (n.d.). Wolfram|Alpha: Computational intelligence. Retrieved from https://www.wolframalpha.com. Last access in May, 2024.

Wolters Kluwer. (n.d.). Kluwer Arbitration. Retrieved from https://www.wolterskluwer.com/en/solutions/kluwerarbitration. Last access in October, 2023.

Wong, J. (2006). *Umbrella clauses in bilateral investment treaties: Of breaches of contract, treaty violations, and the divide between developing and developed countries in foreign investment disputes.* Georgetown Journal of International Law, 14, 135.

Yang, W., Jia, W., Zhou, X., & Luo, Y. (2019). Legal judgment prediction via multi-perspective bi-feedback network. *arXiv preprint arXiv:1905.03969*.

Zeleznikow, J. (2023). The benefits and dangers of using machine learning to support making legal predictions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, e1505.

Zhang, D., Schilder, F., Conrad, J. G., Makrehchi, M., Rickenbach, D. V., & Moulinier, I. (2023). Making a Computational Attorney. In *Proceedings of the 2023 SIAM International Conference on Data Mining (SDM),* 965-968. Society for Industrial and Applied Mathematics.

Zhong, H., Guo, Z., Tu, C., Xiao, C., Liu, Z., & Sun, M. (2018). Legal judgment prediction via topological learning. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, 3540-3549.

Zhong, H., Xiao, C., Tu, C., Zhang, T., Liu, Z., & Sun, M. (2020). How does NLP benefit legal system: A summary of legal artificial intelligence. *arXiv preprint arXiv:2004.12158*.

Appendix 1 - Legal Judgment Prediction Literature Review.

Year	Author	Sample size (#)		Topic or sentences source		Methodology	F1-score *
2024	Pereira, J. et al (2024)	122	Brazil	Brazilian Audit Courts	-	Ad hoc LLM prompting chatGPT-4 to extract text (F1 average)	59.2%
2024	Sun, J., Huang, S., & Wei, C. (2024)	241.434	China	Chinese Legal Judgement Prediction competition dataset	DL	Prompt Learning Frameworks	70%
2023	Schilder, F. (2023)	17.215	France	French Court of Appeal	ML	NLP methods to lawyers´ ranking	N.A.
2023	Shalaby, A. G. (2023)	7.482	China	Private lending	DL	Multi-task framework and life-cycle case representation	86.5%
2023	Sherry Xin Chen (2019)	538	Canada	Unemployment law (worker or contractor status)	ML	Decision Tree	89.5%
2022	Lage-Freitas, A. et al (2022)	4.403	Brazil	Brazilian Court decisions	ML&DL	ML algorithms and classifiers and DL models (web scraper)	80.2%
2022	Sert, M. F. et al (2022)	338	Turkey	Turkish Constitutional Court (public morality)	DL	Multi-Layer perceptron [BoW (Bag of Words)]	98.7%
2022	Sokhansanj, B. A., & Rosen, G. L. (2022)	10.462	US	PTAB (Patent Trial & Appeal Board)	DL	CNN-Attention model and XGBoost model	72%
2021	Niklaus, J. et al (2021)	85.000	Switzerland	Federal Supreme Court (trilingual corpus)	DL	BERT-based methods	70%
2021	Mumcuoğlu, E. (2022)	93.340	Turkey	Turkish Constitutional Court and Courts of Appeal	DL	Several (GRUs, LSTMs and BiLSTMs),	87%
2021	Rhim, Y. Y. et al (2023)	184.125	US	State of Connecticut Judicial branch civil cases	ML&DL	Methods to predict motions with info available to all parties (av.	64.4%
2021	Savelka, J. et al (2023)	16.024	Int	WIPO cases	DL	accuracy based on induced (predictive) tags	79.5%
2020	Bertalan, V. G. F., & Ruiz, E. E. S. (2020)	782	Brazil	Sao Paolo State Higher Court	ML	Supervised machine learning algorithm (text crawler)	98%
2020	Medvedeva, M., Vols, M., & Wieling, M. (2020)	1.942	Europe	European Convention oF Human Rights	ML	Automatic information extraction and Natural Language Proces	75%
2020	Strickson, B., & De La Iglesia, B. (2020)	4.959	UK	UK Court judgments	ML	Legal Judgement Prediction - models applied to data set	69%
2019	Chalkidis, I. et al (2019)	11.500	Europe	European Convention of Human Rights	DL	Language models based on artificial neural networks	82%

In terms of accuracy, the Table only highlights those results within the research studies with highest F1 score (*). We restate the formula: ***F1 Score****=(2*Precision*Recall)/Precision + Recall)* where *Precision = True Positive / (True Positive + False Positive);* and *Recall = True Positive / (True Positive + False Negative)* or percentage of total relevant results correctly classified under the used algorithm.

Year	Author	Sample size (#)		Topic or sentences source		Methodology	F1-score *
2019	Li, S. et al (2019)	1.367.654	China	Supreme People´s Court of China (criminal cases)	DL	Attention neuronal networks and word embeddings	40-91%
2019	Yang, W. et al (2019)	185.228	China	Supreme People´s Court of China	DL	Multi-Perspective Bi-Feedback Network	46-86%
2018	Barros, R. et al (2018)	10.000	Brazil	Brazilian labour court	ML	Bayesian networks to classify [accuracy]	90%
2018	Elnaggar, A. et al (2018)	20.000	Germany	Eur-Lex corpus	DL	Transfer Learning (Named Entity Linking)	98.9%
2018	Hu, Z. et al (2018)	383.697	China	China Judgments online (charge prediction criminal cases)	DL	Attribute-attentive charge prediction model	73.1%
2018	Kowsrihawat, K. et al (2018)	1.207	Thailand	Thai Supreme Court	DL	Bi-GRU and attention model	74.3%
2018	Long, S. et al (2019)	100.000	China	Divorce decisions China Judgements online	DL	Attention-based neuronal network models	85.1%
2018	Virtucio, M. B. L. et al (2018)	8.132	Philippines	Philippines Supreme Court decisions (criminal cases)	ML	Random forest classifier and n-grams for future extraction	59%
2018	Zhong, H. et al (2018)	113.536	China	Supreme People´s Court (charge prediction criminal cases)	DL	Topological multi-task learning framework	78.3%
2017	Katz, D. M. et al (2017)	28.084	US	US Supreme Court (case outcome)	ML	time-evolving random forest classifier	69%
2017	Popp, A. T. (2019).	254.021	US	Asylum court decisions	ML	Random forest model [accuracy]	81.5%
2017	Shulayeva, O. (2017)	50	UK&I	Principles & facts citations British & Irish Legal Institute	ML	Separate case facts and legal principles (Naive Bayes Multimodal Classifie	81%
2017	Sulea, O. M. et al (2017)	126.865	France	French Supreme Court (case ruling, law area and timing)	ML	Text classification methods [F1 relates to case ruling]	96%
2016	Aletras, N. et al (2016)	584	Europe	European Convention oF Human Rights	ML	Support-Vector Machine (violation of 3 articles)	79%
2012	McShane, B.B. et al (2012)	5.898	US	Securities fraud litigation	-	Hierarchical Bayesian models	
2012	Pérez López, J. Á.et al (2012)	864	US	US Federal employment cases	ML	Linguistic and machine learning	74.2%

**F1 refers, in case of several datasets, the one with the highest score [unless indicated otherwise]*

Table 36. Legal Judgment Prediction. This table was created by the author based on a review of the legal judgment prediction literature.

Appendix 2 - Case Study Data

2.1 Total population (n=260)

1	nachingwea-and-others-v-tanzania
2	united-agencies-v-algeria-i-
3	mercuria-v-poland-ii-
4	bbva-v-bolivia
5	elliott-v-korea
6	manolium-processing-v-belarus
7	olympic-entertainment-v-ukraine
8	pacc-v-mexico
9	rand-investments-and-others-v-serbia
10	rushydro-v-kyrgyzstan
11	sunlodges-v-tanzania
12	westwater-resources-v-turkey
13	zhongshan-fucheng-v-nigeria
14	air-canada-v-venezuela
15	bank-melli-and-bank-saderat-v-bahrain
16	de-sutter-and-others-v-madagascar-ii-
17	diag-and-va-v-czechia
18	ecodevelopment-and-ecoenergy-v-tanzania
19	gardabani-and-silk-road-v-georgia
20	magyar-farming-and-others-v-hungary

21 rockhopper-v-italy
22 slot-v-poland
23 agroinsumos-ibero-americanos-and-others-v-venezuela
24 alhambra-v-kazakhstan
25 biram-and-others-v-spain
26 cengiz-v-libya
27 cordoba-beheer-and-others-v-spain
28 dominion-minerals-v-panama
29 espf-and-others-v-italy
30 etrak-v-libya
31 eurus-energy-v-spain
32 glencore-international-and-c-i-prodeco-v-colombia-i-
33 goljev-ek-and-others-v-bosnia-and-herzegovina
34 gramercy-v-peru
35 grot-and-others-v-moldova
36 g-ri-and-yamant-rk-v-syria
37 infracapital-v-spain
38 jsc-tashkent-and-others-v-kyrgyzstan
39 kunsttrans-v-serbia
40 naftogaz-and-others-v-russia
41 oschadbank-v-russia

42 9ren-holding-v-spain
43 aboukhalil-v-senegal
44 aktau-petrol-v-kazakhstan
45 baywa-r-e-v-spain
46 cairn-v-india
47 cavalum-sgps-v-spain
48 cef-energia-v-italy
49 cube-infrastructure-and-others-v-spain
50 dayyani-v-korea
51 el-jaouni-v-lebanon
52 everest-and-others-v-russia
53 foresight-and-others-v-spain
54 greentech-and-novenergia-v-italy
55 hydro-and-others-v-albania
56 hydro-energy-1-and-hydroxana-v-spain
57 jgc-v-spain
58 jkx-oil-gas-and-poltava-v-ukraine
59 kci-v-gabon
60 lion-v-mexico
61 maessa-and-semi-v-ecuador
62 manchester-securities-v-poland

63 novenergia-v-spain
64 operafund-and-schwab-v-spain
65 soles-badajoz-v-spain
66 stabil-and-others-v-russia
67 steag-v-spain
68 strabag-v-libya
69 ukrnafta-v-russia
70 watkins-and-others-v-spain
71 bear-creek-mining-v-peru
72 casinos-austria-v-argentina
73 city-state-v-ukraine
74 cyprus-popular-bank-v-greece
75 flemingo-dutyfree-v-poland
76 Horthel vs Poland
77 infrared-and-others-v-spain
78 masdar-v-spain
79 nextera-v-spain
80 olin-v-libya
81 pl-holdings-v-poland
82 renergy-v-spain
83 rwe-innogy-v-spain

84 sodexo-pass-v-hungary
85 trinh-and-bin-chau-v-viet-nam-ii-
86 union-fenosa-v-egypt
87 vodafone-v-india-i-
88 zelena-v-serbia
89 de-sutter-and-others-v-madagascar-i-
90 deutsche-telekom-v-india
91 edenred-v-hungary
92 eiser-and-energ-a-solar-v-spain
93 g-ne-tekstil-and-others-v-uzbekistan
94 houben-v-burundi
95 infrastructure-services-and-energia-termosolar-formerly-antin-v-spain
96 karkey-karadeniz-v-pakistan
97 lee-v-kyrgyzstan
98 mol-v-croatia
99 mytilineos-v-serbia-ii-
100 okkv-v-kyrgyzstan
101 rreef-v-spain
102 sorelec-v-libya
103 south-american-silver-v-bolivia

104 stans-energy-v-kyrgyzstan
105 up-and-c-d-holding-v-hungary
106 valores-mundiales-and-consorcio-andino-v-venezuela
107 venezuela-us-v-venezuela
108 Windstream vs Canada
109 wwm-and-carroll-v-kazakhstan
110 yukos-capital-v-russia
111 dan-cake-v-hungary
112 devas-v-india-i-
113 garc-a-armas-and-garc-a-gruber-v-venezuela
114 gavazzi-v-romania
115 gavrilovic-v-croatia
116 lone-star-and-others-v-korea
117 rusoro-mining-v-venezuela
118 saint-gobain-v-venezuela
119 tenaris-and-talta-v-venezuela-ii-
120 tethyan-copper-v-pakistan
121 uab-v-latvia
122 al-kharafi-v-libya-and-others
123 arif-v-moldova
124 baggerwerken-v-philippines

125 bahgat-v-egypt-i-
126 belokon-v-kyrgyzstan
127 copper-mesa-v-ecuador
128 crystallex-v-venezuela
129 gamesa-v-syria
130 garanti-koza-v-turkmenistan
131 khan-resources-v-mongolia
132 koch-minerals-v-venezuela
133 longreef-v-venezuela
134 merck-v-ecuador
135 murphy-v-ecuador-ii-
136 oieg-v-venezuela
137 oxus-gold-v-uzbekistan
138 tenaris-and-talta-v-venezuela-i-
139 the-pv-investors-v-spain
140 awdi-v-romania
141 border-timbers-and-others-v-zimbabwe
142 british-caribbean-bank-v-belize-i-
143 energoalians-v-moldova
144 flughafen-z-rich-v-venezuela
145 guaracachi-v-bolivia

146 stati-and-others-v-kazakhstan
147 teco-v-guatemala
148 tidewater-v-venezuela
149 von-pezold-and-others-v-zimbabwe
150 white-industries-v-india
151 abengoa-v-mexico
152 bogdanov-v-moldova-iii-
153 deutsche-bank-v-sri-lanka
154 dogan-v-turkmenistan
155 edf-v-hungary
156 gold-reserve-v-venezuela
157 servier-v-poland
158 swisslion-v-macedonia
159 teinver-and-others-v-argentina
160 uab-v-serbia
161 valle-esina-v-russia
162 achmea-v-slovakia-i-
163 ata-construction-v-jordan
164 burlington-v-ecuador
165 clayton-bilcon-v-canada
166 inmaris-perestroika-v-ukraine

167 intersema-bau-v-libya
168 karmer-marble-v-georgia
169 marion-unglaube-v-costa-rica
170 perenco-v-ecuador
171 remington-v-ukraine
172 tatneft-v-ukraine
173 alpha-projektholding-v-ukraine
174 conocophillips-v-venezuela
175 fuchs-v-georgia
176 hochtief-v-argentina
177 impregilo-v-argentina-i-
178 mobil-and-murphy-v-canada-i-
179 mobil-and-others-v-venezuela
180 rdv-v-guatemala
181 renta-4-s-v-s-a-and-others-v-russia
182 sgs-v-paraguay
183 tza-yap-shum-v-peru
184 chevron-and-texpet-v-ecuador-i-
185 lemire-v-ukraine-ii-
186 occidental-v-ecuador-ii-
187 quiborax-v-bolivia

188 sistem-v-kyrgyzstan
189 vestey-v-venezuela
190 ares-and-metalgeo-v-georgia
191 cargill-v-mexico
192 desert-line-v-yemen
193 funnekotter-v-zimbabwe
194 hulley-enterprises-v-russia
195 kardassopoulos-v-georgia
196 micula-v-romania-i-
197 pren-nreka-v-czech-republic
198 rosinvest-v-russia
199 rumeli-v-kazakhstan
200 saipem-v-bangladesh
201 siag-v-egypt
202 veteran-petroleum-v-russia
203 walter-bau-v-thailand
204 yukos-universal-v-russia
205 adm-v-mexico
206 bogdanov-v-moldova-i-
207 cargill-v-poland
208 corn-products-v-mexico

209 duke-energy-v-ecuador
210 eastern-sugar-v-czech-republic
211 gemplus-v-mexico
212 mobil-v-argentina
213 oko-v-estonia
214 saur-v-argentina
215 talsud-v-mexico
216 total-v-argentina
217 adc-v-hungary
218 awg-v-argentina
219 bg-v-argentina
220 continental-casualty-v-argentina
221 edf-and-others-v-argentina
222 el-paso-v-argentina
223 national-grid-v-argentina
224 parienti-v-panama
225 petrobart-v-kyrgyz-republic
226 suez-and-interagua-v-argentina
227 suez-and-vivendi-v-argentina-ii-
228 france-telecom-v-lebanon
229 lg-e-v-argentina

230 occidental-v-ecuador-i-
231 pseg-v-turkey
232 sempra-v-argentina
233 siemens-v-argentina
234 aig-v-kazakhstan
235 azurix-v-argentina-i-
236 cms-v-argentina
237 enron-v-argentina
238 goetz-v-burundi-ii-
239 mtd-v-chile
240 nykomb-v-latvia
241 cme-v-czech-republic
242 tecmed-v-mexico
243 feldman-v-mexico
244 middle-east-cement-v-egypt
245 mitchell-v-democratic-republic-of-the-congo
246 pope-talbot-v-canada
247 swembalt-v-latvia
248 myers-v-canada
249 pey-casado-and-allende-foundation-v-chile
250 wena-hotels-v-egypt

251	csob-v-slovakia
252	maffezini-v-spain
253	metalclad-v-mexico
254	vivendi-v-argentina-i-
255	biedermann-v-kazakhstan
256	fedax-v-venezuela
257	sedelmayer-v-russia
258	saar-papier-v-poland-i-
259	amt-v-democratic-republic-of-the-congo
260	aapl-v-sri-lanka

Table 37. List of positive outcome awards as of 31.7.2023. United Nations Conference on Trade and Development. (n.d.-d)

2.2 Selected sample - positive outcome (n=94)

Claimant vs Respondent	Case reference
Nachingwea vs Tanzania	ARB/20/38
Bbva vs Bolivia	ARB(AF)/18/5
Elliot vs Korea	PCA Case No. 2018-51
Manolium vs Belarus	PCA Case No. 2018-06
Olympic vs Ukrania	PCA Case No. 2019-18
Sunlodges vs Tanzania	PCA Case No. 2018-09
Zhongshan vs Nigeria	UNCTAD (64)
Bank-melli vs Bahrain	PCA Case No. 2017-25
Magyar farming vs Hungary	ARB/17/27
Rockhopper vs Italy	ARB/17/14
Cengiz vs Libya	ICC 21537/ZF/AYZ
Cordoba vs Spain	ARB/16/27
ESPF B teiligungs vs Italy	ARB/16/5
Eurus Energy vs Spain	ARB/16/4
Glencore vs Colombia	ARB/16/6
Gramercy vs Peru	(UNCT/18/2)
Infracapital v. Spain	Case No. ARB/16/18
Naftogaz vs Russia	PCA Case No. 2017-16
9REN vs Spain	ARB/15/15

BayWa vs Spain	ARB/15/16
CEF Energia vs Italy	2015/158
Cube infrastructure vs Spain	ARB/15/20
Greentech vs Italy	SCC Case No. 2015/095
Foresight et al vs Spain	SCC Case No. 2015/150
Hydro Energy 1 and Hydroxana v Spain	ARB/15/42
Lion vs Mexico	ARB(AF)/15/2
Manchester vs Poland	UNCTAD (62)
Novenergia vs Spain	2015/063
SolEs Badajoz vs Spain	ARB/15/38
Stabil vs Russia	PCA Case No. 2015-35
Strabag vs Lybia	ARB(AF)/15/1
Bear Creek vs Peru	ARB/14/21
Casinos vs Argentina	ARB/14/32
flemingo-dutyfree-v-poland	PCA 2016-12-08
Infrared et al vs Spain	ARB/14/12
Masdar Solar et al vs Spain	ARB/14/1
NextEra et al vs Spain	ARB/14/11
PL Holdings vs Poland	V 2014/13
RENERGY v. Spain	ICSID Case No. ARB/14/18
Union Fenosa vs Egypt	ARB/14/4

Deutsche Bank vs Sri Lanka	ARB/09/02
Eiser et al vs Spain	ARB/13/26
Infra Ser L (formerly Antin) vs Spain	ARB/13/31
Karkey Karadeniz vs Pakistan	ARB/13/1
Mol vs Croatia	ARB/13/32
RREEF Infrastructure vs Spain	ARB/13/30
Silver vs Bolivia	PCA No. 2013-15
Windstream Energy v Canada	PCA Case No. 2013-22
Devas v India (I)	PCA Case No. 2013-09
lone-star-and-others-v-korea	ARB/12/37
Rusoro vs Venezuela	ARB(AF)/12/5
Tenaris vs Venezuela	ARB/12/23 ii
UP and CD vs Hungary	No. ARB/13/35
Valores mundiales vs Venezuela	ARB/13/11
Gavrilovic vs Croatia	ARB/12/39
Saint Gobain vs Venezuela	
Tethyan vs Pakistan	ARB/12/1
UAB E Energia vs Latvia	ARB/12/33
Bahgat vs Egypt	PCA Case No. 2012-07
Valeri vs Kyrgyz	UNCTAD (126)
Copper Mesa vs Ecuador	PCA 2012-2

Khan Resources v. Mongolia	PCA Case No. 2011-09
Koch Minerals et al vs Venezuela	ARB/11/19
OI European Group vs Venezuela	ARB/11/25
Oxus Gold vs Uzbekistan	UNCTAD (17)
Hassan Awdi et al vs Romania	ARB/10/13
British Caribbean vs Belize	PCA 2010-18
EnergoAlliance vs Moldova	UNCTAD (143)
Flughafen Zürich VS. Venezuela	ARB/10/19
Stati et al vs Kazakhstan	116/2010
Teco vs Guatemala	ARB/10/17
Tidewater vs Venezuela	ARB/10/5
White industries vs India	UNCTAD (89)
Abengoa vs Mexico	ARB(AF)/09/2
Gold Reserve vs Venezuela	ARB/(AF)/09/1
Burlington vs Ecuador	ARB/08/05
Marion Unglaube vs Costa Rica	ICSID Case No. ARB/08/1
Perenco vs Ecuador	ARB/08/6
OaO TatNeft vs Ukraine	N.A.
Hochtief vs Argentina	ARB/07/31
Impregilo vs Argentina	ARB/07/17
Tza Yap Shum v Peru	No. ARB/07/6

Lemire vs Ukraine	ARB/06/18 ii
Quiborax et al vs Bolivia	ARB/06/2
Funnekotter et al vs Zimbabwe	ARB/05/6
Ioannis et al vs Georgia	ARB/05/18
Ioan et al vs Romany	ARB/05/20
Walter Bau vs Thailand	UNCTAD (203)
Gemplus et al vs Mexico	ARB (AF) 04/3 ARB04/4
Saur vs Argentina	ARB/04/4
Talsud v. Mexico	(ICSID Case No. ARB(AF)/04/4)
EDF et al vs Argentina	ARB/03/23
MTD vs Chile	ARB/01/7
Vivendi vs Argentina	ARB/97/3

Table 38. List of awards within this research sample. United Nations Conference on Trade and Development. (n.d.-d) and Wolters Kluwer. (n.d.)

Appendix 3 – Histograms

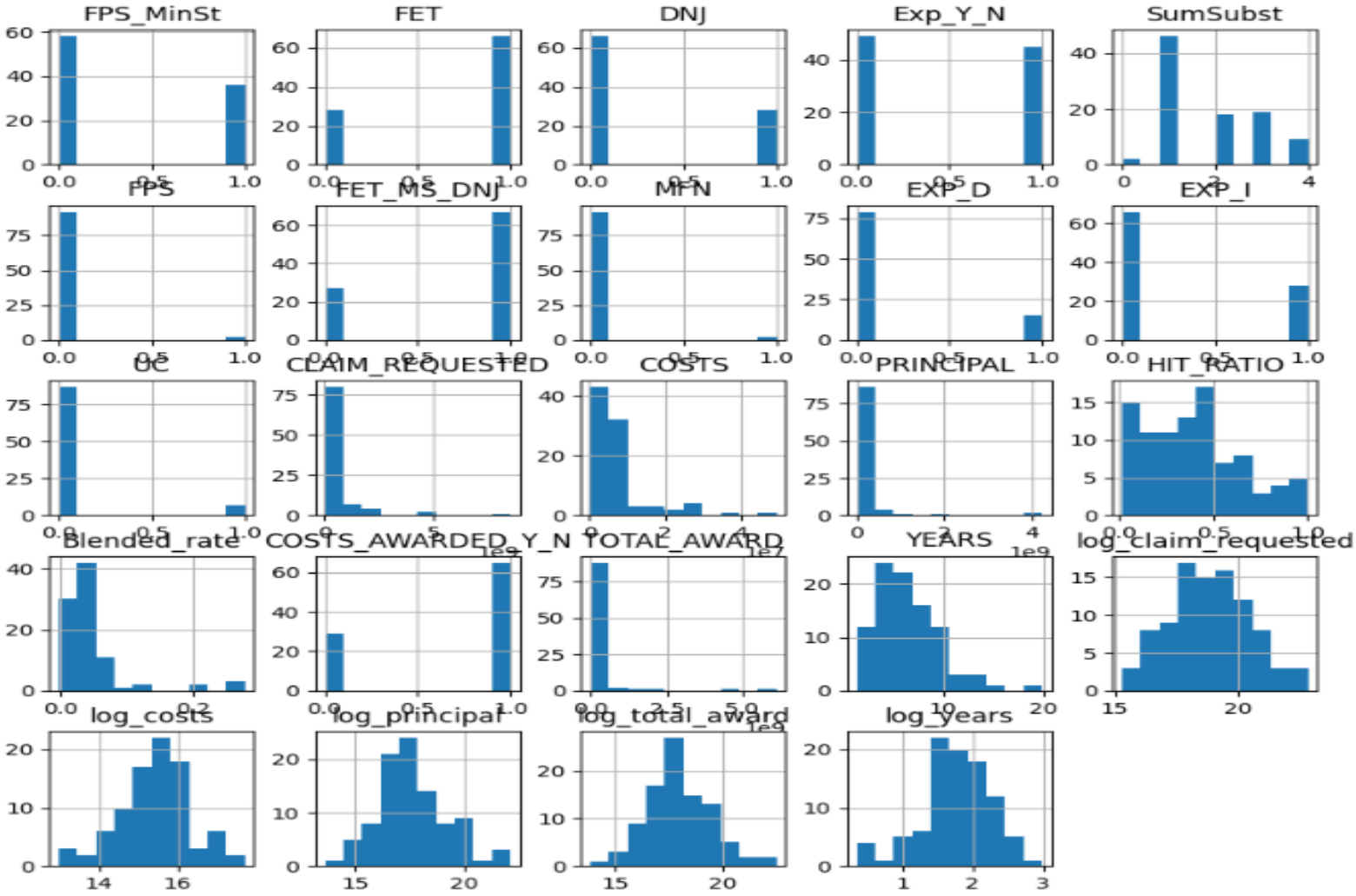


Figure 9. Histograms of all variables. Source: Own data.

Appendix 4 - Correlation graphics and matrix

Correlation graphics (i)

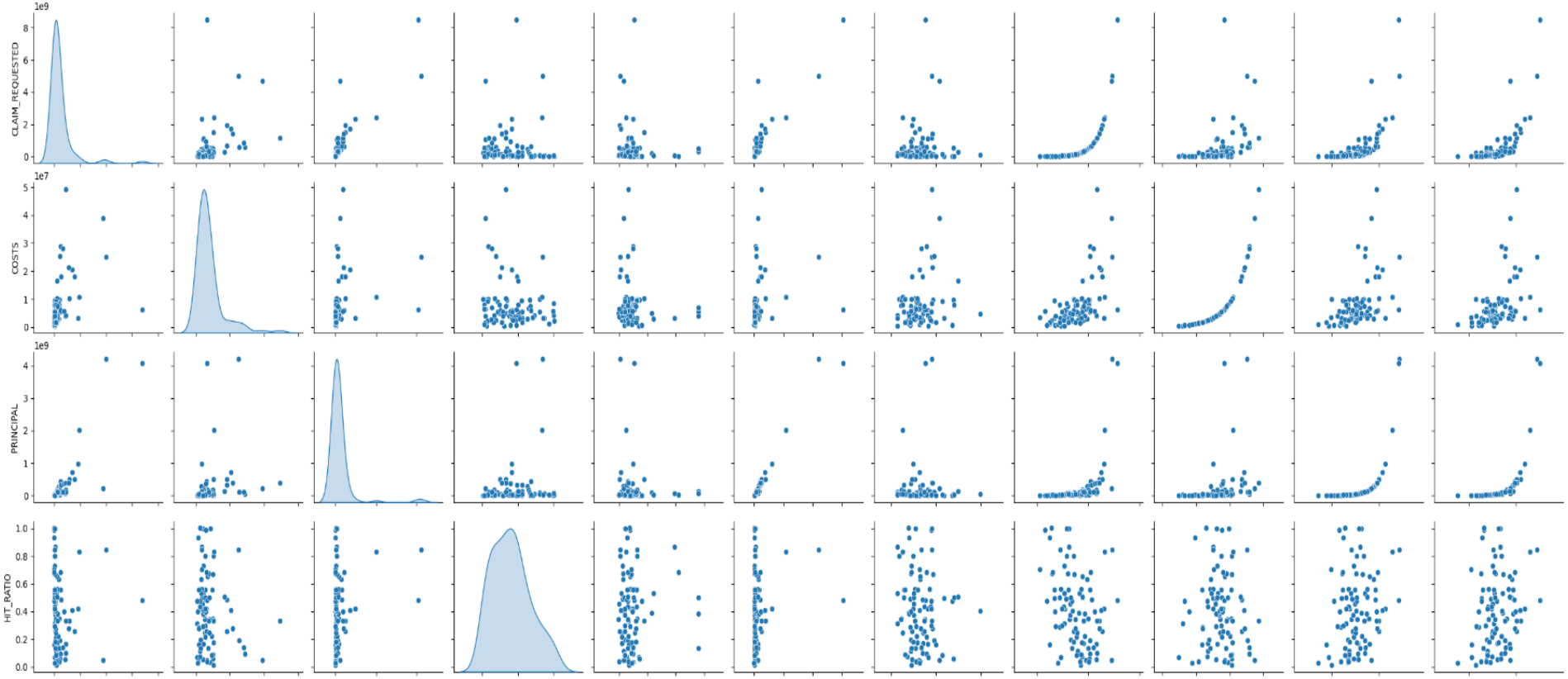


Figure 10.A. Correlation graphics related to claim requested, costs, principal, and hit ratio. Source: Own data.

Correlation graphics (ii)

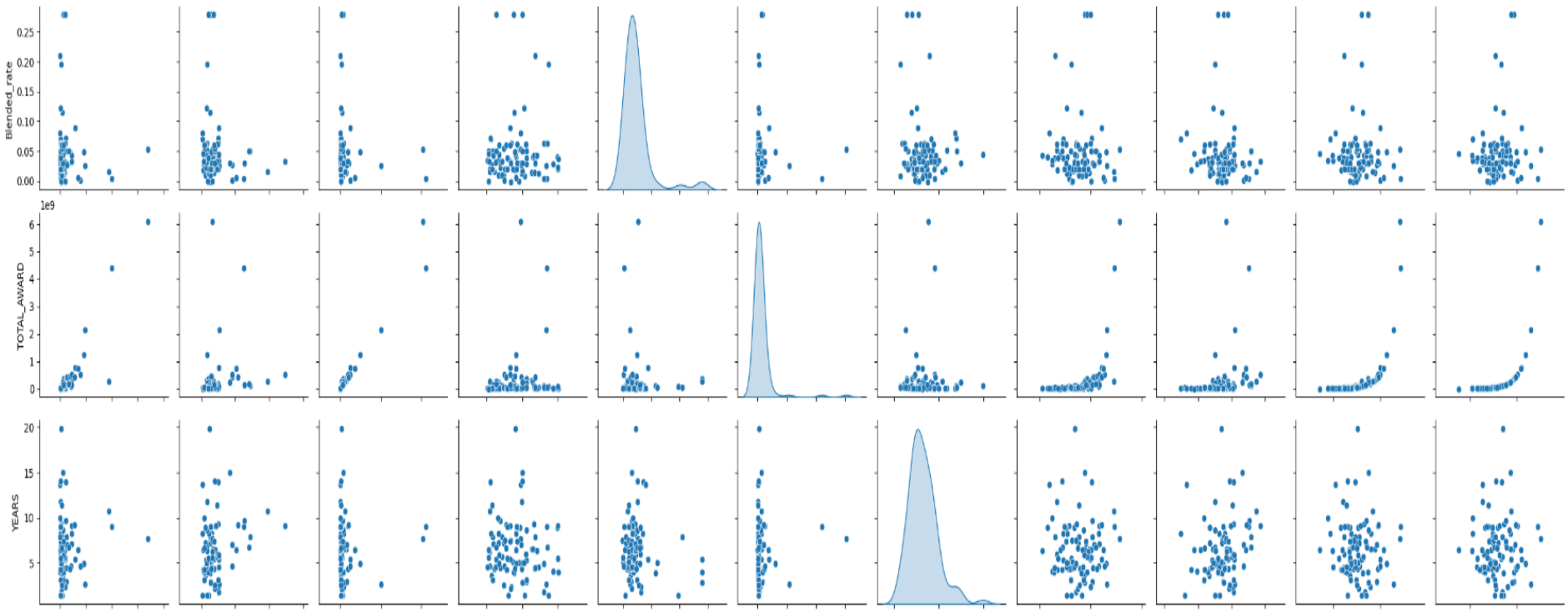


Figure 10.B. Correlation graphics related to blended rate, total award, and years. Source: Own data.

Correlation graphics (iii)

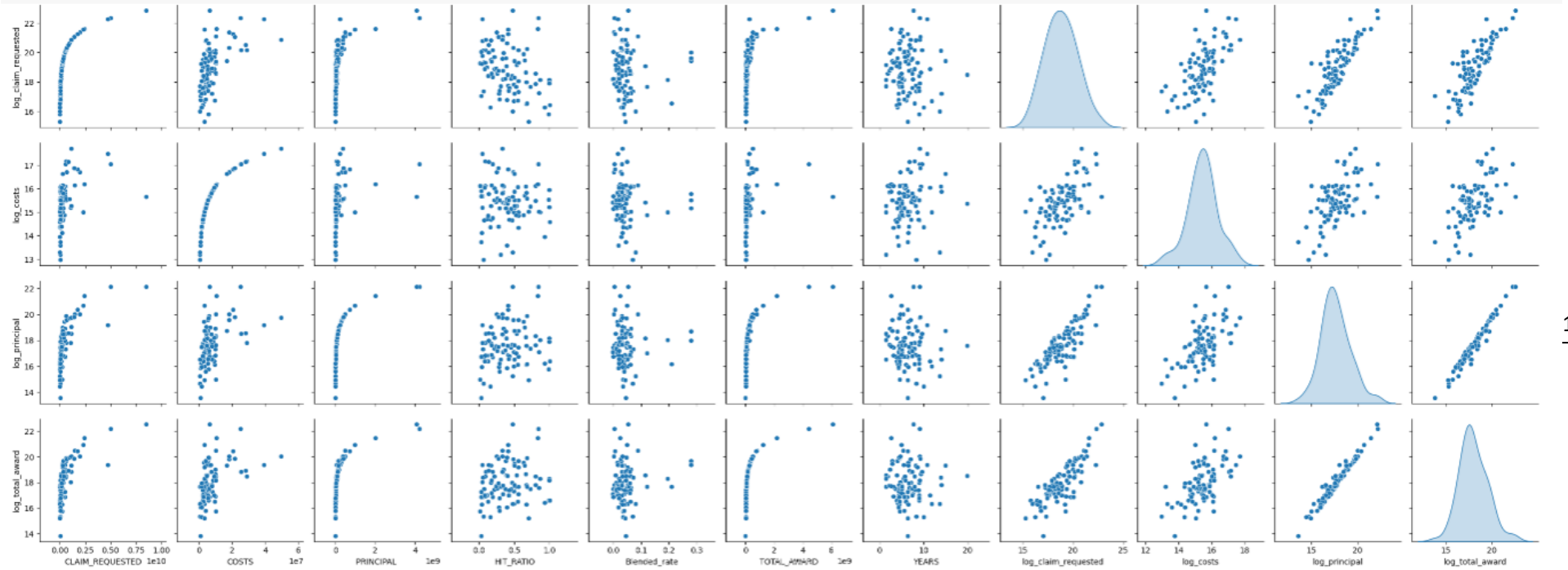


Figure 10.C. Correlation graphics related to claim requested (log), costs (log), principal (log) and total award (log). Source: Own data.

Correlation matrix (iv)

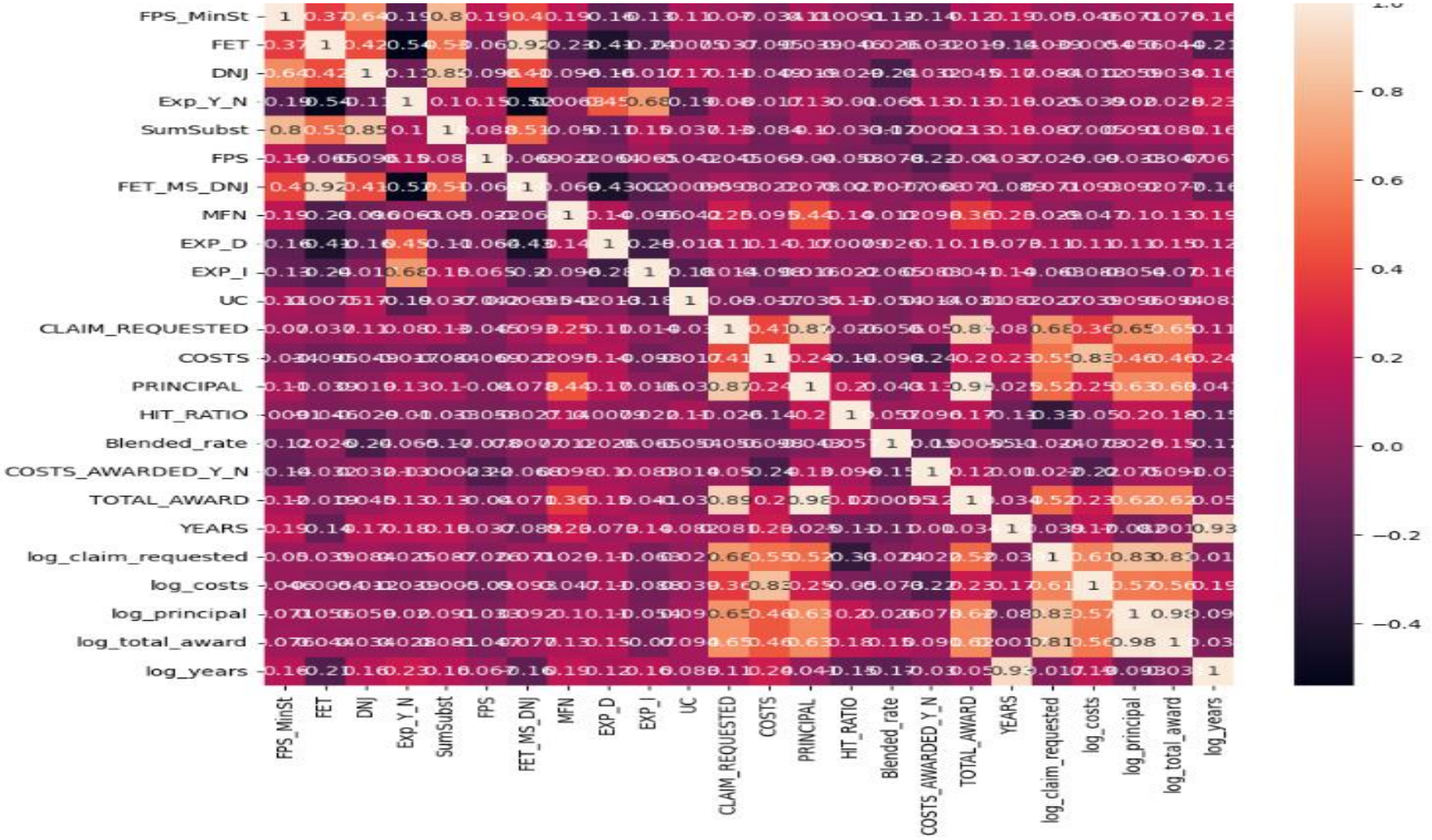


Figure 10.D Correlation matrix (all variables). Source: Own data.

Appendix 5 – OLS Models

5.1 – OLS regression for the Principal

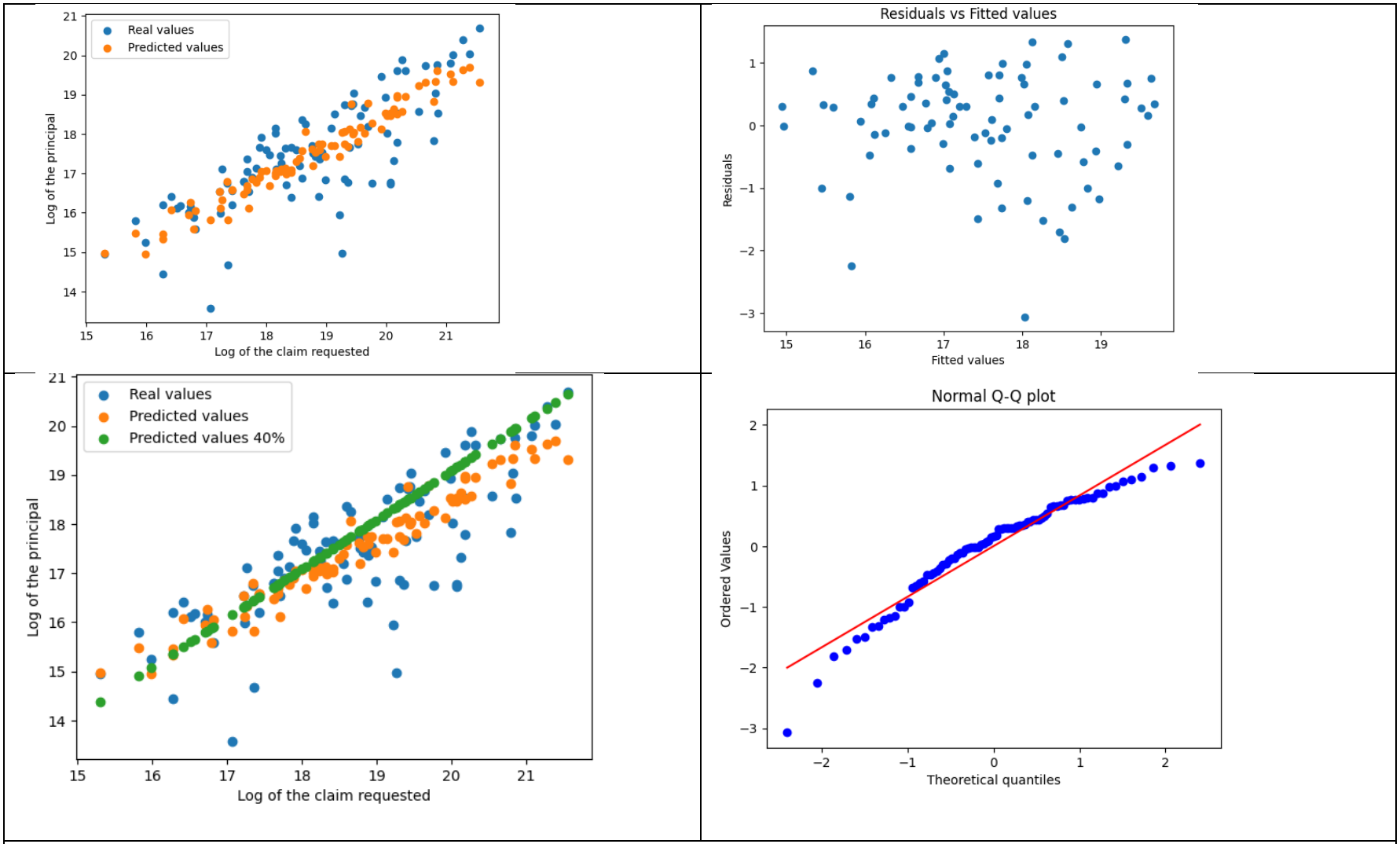
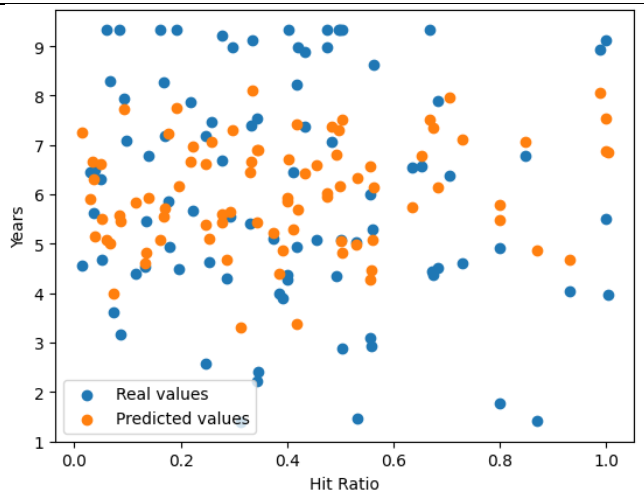
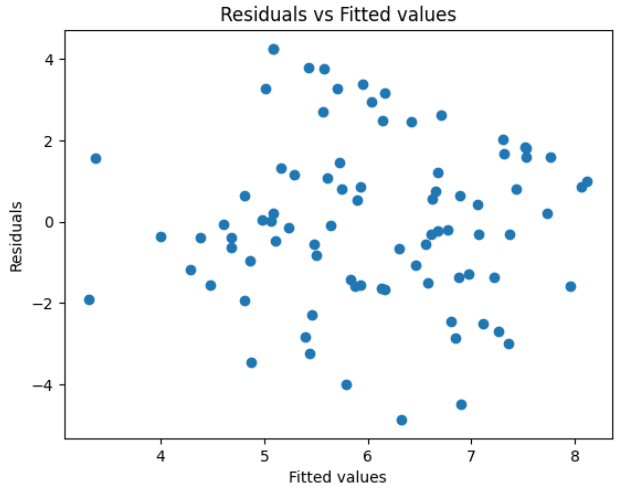
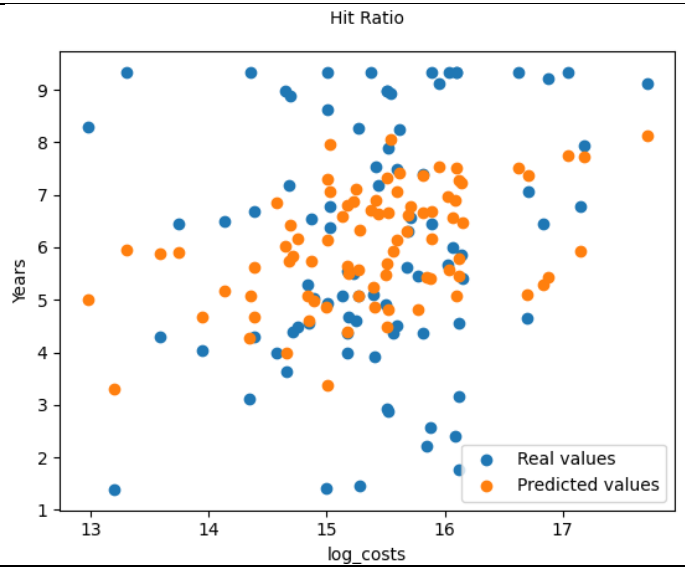


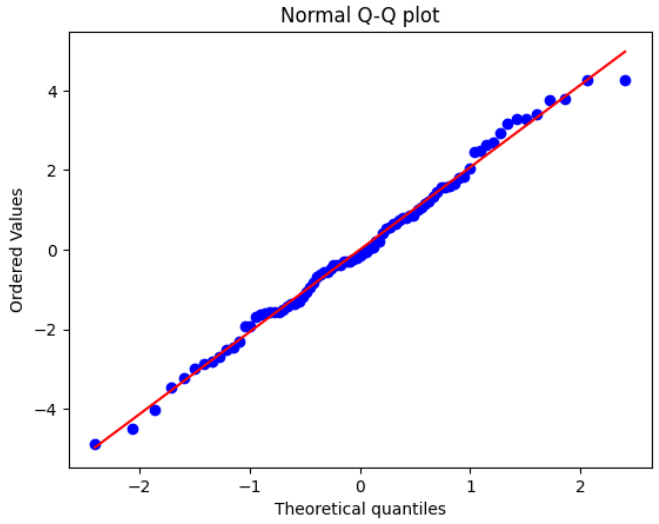
Figure 11. OLS Regression Model for the Principal. Left: Scatterplot of Logarithm of Claim Request vs. Logarithm of Principal (real vs predicted values), Right: Fitted Values vs. Residuals, and Quantile-Quantile plot. Source: Own data.

5.2 – OLS Regression for the Years


Years
Hit Ratio
Real values
Predicted values


Residuals vs Fitted values
Residuals
Fitted values


Hit Ratio
Years
log_costs
Real values
Predicted values


Normal Q-Q plot
Ordered Values
Theoretical quantiles