Christoph Engel

EXPERIMENTAL COMPARATIVE LAW 2.0? LARGE LANGUAGE MODELS AS A NOVEL EMPIRICAL TOOL

1.   The challenging empirical turn in comparative law

Legal orders differ. Obviously at the doctrinal surface. Arguably, at least in some instances, also in substance. Seemingly this variance invites an empirical enterprise. Does country, as a proxy for judicial politics, legal history, and jurisprudential culture, determine how one and the same social conflict is authoritatively dealt with? In a different paper I have explained in much more detail why empirical comparative law is actually a very challenging endeavor[1]. Suffice it to reiterate the two biggest obstacles: it is not enough to point at differences in doctrine. Not so rarely, from deeply different doctrinal starting points, two jurisdictions come to closely related conclusions. This challenge is on the screen of every trained comparative lawyer. This is why they would usually prefer to compare how different jurisdictions decide identical cases. They would only accept a substantial difference across jurisdictions if the expected outcomes differ. Yet the statistical challenge is no less pronounced. Jurisdictions do not live in isolation. They observe each other, and respond to these observations. In statistical parlance, jurisdictions are therefore not independent from each other, which makes statistical testing moot. Moreover it is not randomly determined which jurisdiction adopts which doctrinal or substantive solution. This is why just observing a difference in outcomes is inconclusive. The observed difference may well result from non-legal differences, for instance since the societies are differently composed, or are held together by different narratives.

One powerful way out is experimentation. If each alternative solution is randomly assigned to sufficiently many randomly selected participants from the same subject pool, and if one observes a difference in the choices these participants make depending on the solution to which they have been assigned, these differences have a causal interpretation. It is precisely due to randomization (and the appropriate design of the experiment) that all alternative explanations are ruled out. Despite its conceptual appeal, experimental comparative law is not widespread[2]. One obstacle is practical. Researchers need access to a lab, a subject pool, and most importantly an experimental community that helps them translate their legal research question into a proper experimental design. The other obstacle is usually discussed under the heading of external validity. The very price one has to pay in a (physical or online) lab for the identification of cause and effect is artificiality. Precisely because one wants to rule out any alternative explanation, one must radically decontextualize. It inevitably takes a leap of faith to extrapolate from the clean causal evidence to the research question from comparative law that has motivated the experiment.

This is where large language models come in. Models like GPT, Gemini, or Claude have not only revolutionized tasks like code writing, summarization, or online search. These models can also be queried about choice tasks. In legal education, it is standard to give students what in the experimental community would be called a vignette. Students read a short scenario, meant

---

1   Christoph Engel, Challenges in the Interdisciplinary Use of Comparative Law, 69 American Journal of Comparative Law (2021).
2   Examples from my own work include Christoph Engel & Michael Kurschilgen, Fairness Ex Ante and Ex Post. Experimentally Testing Ex Post Judicial Intervention into Blockbuster Deals, 8 Journal of Empirical Legal Studies (2011);Oren Bar-Gill & Christoph Engel, How to Protect Entitlements. An Experiment, 61 Journal of Law and Economics (2018).

to highlight a debated feature of law. This method is congenial to the way how large language models work. Effectively, researchers can simply give their LLM of choice the exact same vignettes that they had hitherto given their students.

The application of this method is straightforward if the comparative researcher, in a first, doctrinal step, has isolated a true difference in the solutions different jurisdictions have implemented to the same social conflict. A classic is the remedy for breach of contract[3]. In civil law countries, the default is specific performance. The aggrieved party may take her contractual partner to court. The court will oblige the contractual partner to fulfil the contract. If she does not, a bailiff will enforce the contract. By contrast, in the common law countries the default remedy for breach of contract is expectation damages. Hence the aggrieved party may only sue for money, not for the actual good or service. This difference has triggered the law and economics literature about "efficient breach of contract". This difference in actual outcomes may be translated into an experimental design. As there is a true difference in legal interventions, it in principle suffices to test participants from a single subject pool on either solution[4].

Yet arguably, legal interventions do not live in isolation. They have emerged in a given societal context. One might, for instance, ask whether the common-law solution to breach of contract has been influenced by the prominence of utilitarian thinking in the UK and the US. On the other hand, specific performance might reflect a greater openness towards deontological thinking on the European continent. Beyond deep philosophical traditions, there might also be a match between local cultures and legal solutions. In principle, one might of course try to also manipulate such context factors. But this would not only be logistically demanding. One would need to access populations within multiple jurisdictions. More importantly even, it would be close to impossible to hold everything else constant, but for the difference in context one wants to investigate. Very likely, participants in the different locations would differ by other features as well, like socio-economic status, religiosity, or predominant political orientation. Observed differences might result from such (observed or unobserved) differences in population characteristics, rather than the one difference one wants to investigate, i.e. the fact that the law of one specific jurisdiction is applied.

2. Large Language Models as a Proxy

This is where technical progress becomes interesting. Large language models are prepared to be guided. One might, for instance, instruct the model: "you are a judge in jurisdiction X". One might then give the model the same vignette and simply manipulate "X". While this is a serious option, it requires that the model has sufficient information about differences across the solutions adopted by different jurisdictions. The presently available large language models have not been specifically trained for deciding legal cases. In this project, I use a different approach

---

3   From the rich literature see only Charles J Goetz & Robert E Scott, Liquidated Damages, Penalties and the Just Compensation Principle. Some Notes on an Enforcement Model and a Theory of Efficient Breach, 77 Columbia Law Review (1977);Daphna Lewinsohn-Zamir, The Questionable Efficiency of the Efficient Breach Doctrine, 168 Journal of Institutional and Theoretical Economics (2012).

4   This is what we have done in Christoph Engel & Lars Freund, Behaviorally Efficient Remedies. An Experiment (2017).

that arguably is not subject to this limitation. I exploit the fact that the flagship models by openAI are available in multiple languages. My only manipulation is the language in which I present the otherwise identical vignette.

As I am interested in distributions, not just in the most likely outcome, I allow the model output to vary. This intervention is congenial to the architecture of large language models. Under the hood, these models are not deterministic; they are probabilistic. When setting "temperature = 1", and when sampling sufficiently many responses to the same input, one can make the variance visible that the model expects to exist in the population of interest.

In preparation of this paper, I have repeated an exercise that we have implemented in much greater depth in another legal project employing large language models[5]. I have given one and the same vignette (this time also holding language constant) to the very latest models offered by openAI: gpt-3.5-turbo-1106; gpt-4-1106-preview; gpt-4o, each time asking for 100 responses, while setting temperature = 1. As in our earlier project, results show that the two newer models dramatically reduce variance. Apparently in the attempt at improving the accuracy of the model when queried once, the designers of the large language model have made it essentially unusable as a proxy for the variance of responses in the respective population. This is why, for the evidence presented in this paper, I am exclusively reporting results generated with gpt-3.5-turbo-1106. For each condition, I am eliciting 100 observations[6].

3. Design

I am exploiting the fact that, in an earlier project, I have investigated in which ways this large language model reacts to different vignettes that all center around the importance of "good faith" for the decision of a legal conflict[7]. I therefore know that the model is sensitive to this concern, and that it exhibits variance that I can exploit. For the present project I am picking the one vignette that, in this earlier project, had the clearest results. The vignette reads:

> "On February 1, A has contractually agreed to deliver 200 menus for B's wedding on May 20. Guinea fowl is agreed as the main course. The price of the menu is to be EUR 90 per person In April, there is an outbreak of bird flu in a chicken coop in another region of the country. At the end of April, the health authorities ban the delivery of poultry from national stables until it is certain that the virus has not spread. Chicken from neighboring countries of the European Union, where no cases of bird flu have been observed so far, may continue to be used in gastronomy. However, because deliveries within the country are not possible, prices for

---

5   Christoph Engel, et al., Do Algorithmic Decision-Aids Disempower Democracy and the Rule of Law? An Experiment with Large Language Models, (2024).
6   Despite a clear prompt, the language model sometimes refuses to respond. This is why I always ask for 110 responses. For data analysis, I am using the first 100 usable responses. This approach is without loss of generality, as the model has no memory. Hence in the statistical sense, each new run is an "independent identical observation".
7   Christoph Engel, Treu und Glauben: Frag GPT, in Treu und Glauben im Wirtschaftsrecht (Studiengesellschaft für Wirtschaft und Recht ed. 2024).

foreign chickens have risen. Transportation over longer distances is also significantly more expensive. A only wants to fulfill the contract if B increases the menu price to 120 euros. B refuses, arguing that the risk of changes in purchase prices falls within A's sphere of responsibility.

Should B agree to increase the price to 120 euros per menu?"

I have translated this vignette into three further languages: German, French and Italian.[8] In the system prompt, I defined the task as follows[9]:

"I am interested in your best guess about the way how a case would be decided, were it to be litigated. The response can either be "Yes" (the contract should be adjusted to a change in circumstances) or "No" (the contract should not be adjusted to a change in circumstances)."

In the data analysis, I call this the "predict" prompt. I compare it with three other system prompts (holding the vignette constant)[10]. The "expect" prompt reads

"I will inform you about a case through a separate prompt. I am interested in your assessment. Your answer can be either Yes (the customer can be expected to agree to an amendment of the contract in the light of the circumstances) or No (the customer cannot be expected to agree to an amendment of the contract even in the light of the circumstances)."

The "fair" prompt reads

"I will inform you about a case through a separate prompt. I am interested in your assessment. Your answer can be either yes (it would be fair for the contract to be adjusted in the light of the circumstances) or no (even in the light of the circumstances it would not be fair for the contract to be adjusted)."

Finally the "just" prompt reads

"I will inform you about a case through a separate prompt. I am interested in your assessment. Your answer can be either yes (it would be just for the contract to be adjusted in the light of the circumstances) or no (even in the light of the circumstances it would not be just for the contract to be adjusted)."

---

8   The wording of the translated vignette is in the online appendix.
9   This „system prompt" additionally contains requests for formatting the response, such that I can easily analyze the resulting data. The complete prompt, and the equivalent versions in the three other languages, are available in the online appendix.
10   All prompts, in all languages, are in the online appendix.

Table 1 summarizes the key differences between treatments.

|	German	English	French	Italian
predict	sollte […] ange-passt werden	should be adjusted	devrait être adapté	dovrebbe essere modificato
expect	kann erwartet werden	can be expected	on peut s'attendre	si può aspettare
fair	wäre fair	would be fair	serait equitable	sarebbe corretto
just	wäre gerecht	would be just	serait juste	sarebbe giusto

Table 1
Critical Words of the Prompt

The four prompts capture different facets of the task faced by a court that has to decide the case. The "predict" prompt is predominantly cognitive. With this prompt, I am probing the beliefs the large language model holds about the law. I deliberately do not specify that the decision is to be taken applying the law of a defined jurisdiction. Rather I want to learn in which way the mere fact that the question is asked in a specific language already affects the response. The "expect" prompt leaves it to the large language model whether to interpret the task more cognitively (what does the model think that a typical contracting partner would do?) or more motivationally (what is the contractual partner entitled to expect, given the increase in procurement cost?). The "fair" and "just" prompts are openly motivational. From two related, but potentially distinct perspectives, the prompts ask about normative convictions.

4. Results

Figure 1 reports results. If the model responds "Yes", it thinks or argues that the contract will or should be adjusted as requested by the seller. There is visibly a big difference between countries. The results for German and English look similar. If I ask the same question in French, results look very different. They look totally different if I ask in Italian.

Actually, when pooling the data over all four tests, except for the comparison between German and English, all country comparisons are highly significant (always $p < .001$, Mann Whitney). I find the same results if I consider the results for Expected, Fair and Just in isolation. If I only consider the results for Predict, I also find a significant difference between German and English ($p < .001$), but do not find a significant difference between English and French; all other bilateral country comparisons are again highly significant ($p < .001$).
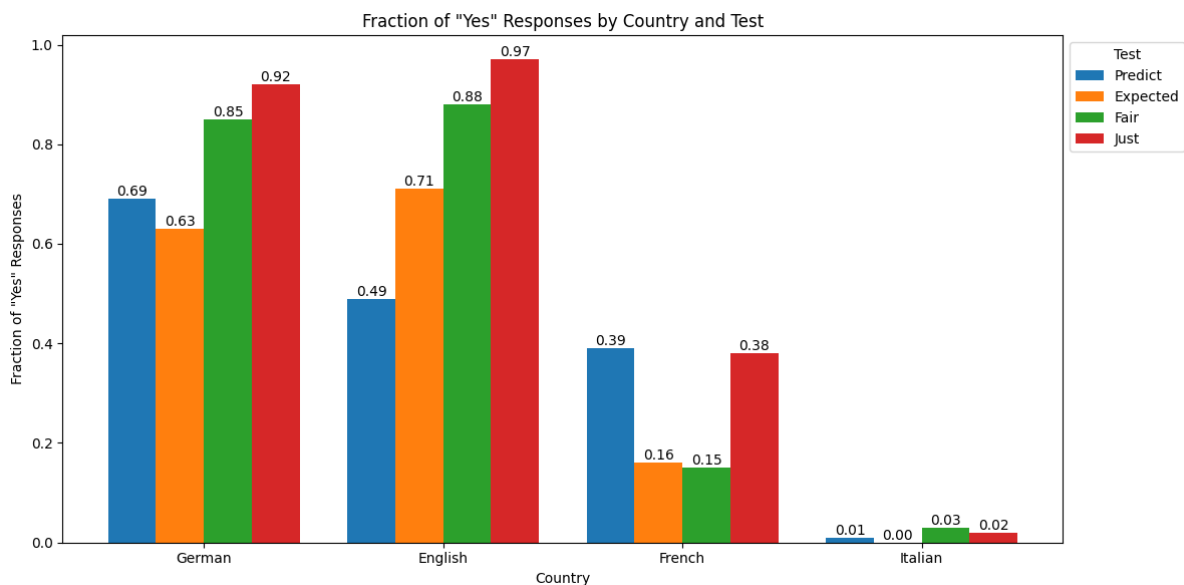
Figure 1
GPT Responses Conditional on Language and Definition of the Task

If I am asking in German, in more than two thirds of all queries, the model believes that the courts would adjust the contract. If I ask in English, about half of the queries lead to this result. If I ask in French, only little more than a third think so, and if I ask in Italian, the same language model is very certain that the courts would not adjust the contract.

5. Does the Model Reflect a Difference in Law?

How can these striking differences be interpreted? As I have asked for the likely outcome in court, one may wonder whether, during the training in either language, the language model has picked up true differences between the ways how the corresponding jurisdictions would be prepared to dispose of the case.

There is something to this interpretation[11]. German courts have for a long time been relatively open to adjusting contracts to unforeseen circumstances. Take this (already very early) case: the seller had sold 6000 units of flour from a specific mill, known for the specific quality of its products. Before he could deliver, the mill had been destroyed by fire. But before the fire, some flour had been shipped to different destinations. Technically, it would have been possible to trace and buy this flour, of course at a much higher price. The *Reichsgericht* held that the seller could not be expected to do that.[12] There is an explicit provision in the code, § 313 I BGB: "If circumstances that have become the basis of the contract have changed significantly after the contract was concluded and the parties would not have concluded the contract or would have concluded it with different content if they had foreseen this change, the contract may be

---

11   For a comparative law perspective, see Gareth H. Jones & Peter Schlechtriem, Breach of Contract (Deficiencies in a Party's Performance), in Volume VII: Contracts in General 208-220, (Arthur T. von Mehren ed. 1999).
12   RG 23.2.1904, RGZ 57, 116.

amended if one party cannot reasonably be expected to adhere to the unchanged contract, taking into account all the circumstances of the individual case, in particular the contractual or statutory distribution of risk." In addition, § 593 I 1 BGB stipulates with regard to land lease contracts: "If, after the conclusion of the lease agreement, the circumstances that were decisive for the determination of the contractual services have changed permanently in such a way that the mutual obligations have become grossly disproportionate, either party to the agreement may demand an amendment to the agreement with the exception of the lease term."

French law traditionally thought otherwise, as captured by art. 1103 Code Civil: "Legally formed contracts take the place of law for those who have concluded them." This provision stood essentially unqualified before the major reform of contract law in 2016[13]. Yet in the new version of the code, it is followed by art. 1104: "Contracts must be negotiated, formed and performed in good faith". Even more clearly, the new art. 1195 stipulates: "If a change in circumstances that was unforeseeable when the contract was concluded makes performance excessively onerous for a party that had not agreed to assume the risk, that party may ask its co-contractor to renegotiate the contract."

Italian law has a similar provision in Art. 1467 Codice Civile: "In contracts with continuous or periodic performance or with deferred performance, if the performance of one of the parties has become excessively onerous due to the occurrence of extraordinary and unforeseeable events, the party owing such performance may request termination of the contract".

Now one of the major limitations of most language models is the opaqueness of the training data. One does not know whether the language model trained in French or Italian has picked up the traditional reluctance of the law in these countries to adjust contracts to changes in circumstances, and does not properly reflect subsequent statutory or jurisprudential changes. It is therefore possible that the French results are driven by the fact that the training data does not properly reflect recent changes in the law. Hence it is conceivable that the differences in outcomes result from the way how differences in the law of the country are reflected in the way how pertinent texts discuss related issues.

## 6.   Does the Model Reflect a Difference in Societal Attitudes?

But there is a competing explanation, which is suggested by Figure 2. GPT is not a model that has been specifically developed for legal applications. Moreover, the adjustment of contracts to a change in circumstances may be interpreted as responsiveness of the law to perceived justice, or to perceived unfairness. The striking differences in the responses of GPT to the very same vignette across languages might therefore reflect a difference in societal attitudes, more than a difference in the actual law. At any rate, with respect to all three indicators from the World Value Survey, the four mother countries of the four languages are ranked essentially the same way as GPT's responses in Figure 1: Germany is a (slightly) more trusting country than the UK, and both are considerably more trusting than both France and Italy. Trustworthiness

---

13   It was then art. 1134 Code Civil.

scores are highest in the UK, higher than in Germany, which are (somewhat) higher than in France, and again (somewhat) higher than in Italy. Country differences are most pronounced with respect to the left right scale of policy making. Germany is the only country where the left scores dominate. The UK is almost exactly at the midpoint, while France is somewhat to the
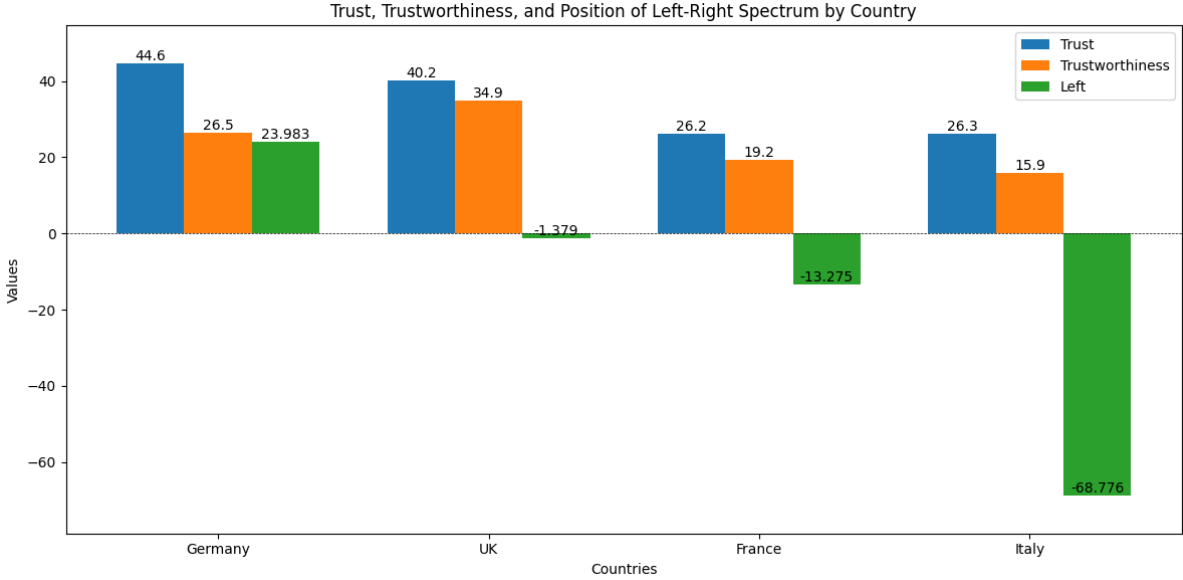
right, and Italy strongly leans towards the right.

Figure 2
Country Scores in the World Value Survey
Wave 7 (2017-2022)
trust: "most people can be trusted", percentages
trustworthiness: two-point scale, percentages of "High" responses
left: left-right-political-scale, 11 points, recoded such that midpoint is 0, percentages of non-0 responses multiplied with rescaled score, results for comparison rescaled to theoretical range [-100, 100]

7. Discussion

Language models exploit the power of natural language. In another project, we have shown that language models are sensitive to framing: if a prisoner's dilemma is presented as a joint project, defending against a joint enemy, or competing with each other, this affects the stated willingness of the large language model to cooperate with another instance of the same model[14]. In our earlier LLM experiment, the effect can be explained with a change in meaning. If two individuals have joined forces for a commonly chosen project, it is intuitive for them to be in a cooperative mindset. Even more so if a common enemy puts them into the same boot. In terms of incentives, competition is of course no different. If the competitors "cooperate", they can collusively exploit the opposite market side and charge the monopoly price. But as our results show, this is not how the large language model perceives competition. With this frame, cooperation rates drop. In the present paper, I radicalize the framing effect. I am not only holding the "objective" situation constant. I even tell the same story. All I change is the

---

14   Christoph Engel, et al., Integrating Machine Behavior into Human Subject Experiments. A User-Friendly Toolkit and Illustrations (2024).

language in which I tell the story. As Figure 1 shows, this very subtle intervention has a drastic effect.

I am not only observing differences between languages (and by implication countries in which these languages are spoken). I also observe differences between prompts. These differences are most pronounced if I am asking in English, but I also find substantial differences in the German and French versions of the experiment. These differences are interesting in their own right. It seems that, at least in most languages / societies, not everything that appears just or fair is also expected to be how the courts interpret a case that revolves around this justice or fairness problem. These observations invite further investigations about the (jurisdiction specific) gap between widespread normative convictions and (expectations about) the responsiveness of the law to such societal concerns. These differences across prompts of course also send a note of caution: even if effects turn out significant, one should be careful with interpreting the observed quantities. It is safer to focus on the qualitative difference across the four languages.

It is standard in experiments to discuss internal and external validity. Ideally, the design of the experiment does not leave any doubt regarding internal validity. Experiments are on one end of the spectrum of empirical methods. They privilege internal over external validity. It is precisely by decontextualization that an experiment achieves identification. It strips the real life conflict that has motivated the endeavour from all plausible alternative explanations, to be sure that the manipulation alone must have caused the treatment effect. My present experiment is designed in this tradition. The only change from one treatment to the next is the natural language in which the vignette is presented. Yet natural language is less precise than mathematics. Despite the fact that I have made every effort to faithfully translate the vignette from one language to the other, I can only hold the presentation standardized to the degree permitted by the languages I have employed. The critical words reported in Table 1 show to which degree this has been possible. One particularly challenging term is "fair". While German and English use the term in reasonably similar ways, neither French nor Italian routinely use the originally English word, in its original meaning. The translations that I have used instead ("equitable", "corretto") convey some of the meaning of the word "fair" in English, but do not perfectly capture all its original dimensions. I am acknowledging this limitation, that is inherent in the method.

No experiment is externally valid, and this experiment is no exception. I am just asking the model. In large language models, the standard critique of vignette studies is somewhat mitigated though. The large language model has nothing to gain individually from pretending not to behave in a way that it would not were its own money at stake. Nonetheless social desirability might be an issue. The designers of the large language model might have wanted to promote the model by training it to give inoffensive responses. Another standard critique of vignette studies is also mitigated. I am not interested in some decontextualized abstract phenomenon; I want to learn how the large language model assessors a complete legal case. I therefore do not have to worry about unexpected behavioural effects sneaking in through the words used for presenting the social problem. But I could of course have given the large lan-

guage model a different case that also revolves around the adaptation of a contract to a pronounced change in circumstances. The evidence presented in this paper ultimately only speaks to the case that I have tested. But as I have stressed in the introduction, focusing on specific cases is an established method in comparative law, motivated by the intention to overcome seemingly profound differences in doctrine that might obscure similarities in the disposition of the same case. At any rate, this limitation is easy to overcome. In a more encompassing approach, one might give the large language model not only a single case, but a whole battery of cases that illustrate the same potential difference across jurisdictions.

Jurisdictions differ in the balance between professional expertise and lay involvement. The large language model that I have used (GPT-3.5-turbo-1106) has not been specifically trained on legal materials. Once such specialized language models are rolled out, it would be a promising endeavour to repeat the experiment. Yet for my specific case, the resulting limitation is not severe. Even if the case is exclusively dealt with by professional jurist, the critical element for its decision is the responsiveness of the court to considerations of good faith. Hence arguably the critical element is the degree by which the law opens itself to non-legal concerns, by which it allows the doctrinal terms to be interpreted in the light of societal expectations and convictions.

For the reasons explained in the introduction, an empirical approach to comparative law is challenging. Many of the standard approaches to empirical research are questionable. This in particular holds for the seemingly most intuitive approach: comparing outcomes across jurisdictions. Given these challenges, it should be welcome news that large language models open a new door for empirical comparative law. As the method is experimental, the identification of causal effects is feasible. Researchers need a bit of coding knowledge to make the method work, but large language models are happy to help with coding as well. Researchers also need a bit of experimental expertise to design proper stimulus materials. But good vignettes are close cousins of good sketches of cases, as they are routinely used in legal education. The method is most appealing for two features: it is very cheap (the entire experiment did not cost more than a few dollars), and it is easily scalable. Hence if they wish, comparative researchers can test the model on a whole battery of cases.

Comparative law has no reason to give up its traditional approach. Legal researchers who have gained a deep understanding of a second (or even a third) legal order will remain powerful agents of the discipline. But as a complement, an empirical approach has appeal, and large language models hold the potential of boosting the empirical arm of comparative law.

Online Appendix

Instructions

1. Vignettes

a) German

A hat sich am 1.2. vertraglich verpflichtet, für die Hochzeit von B am 20.5. 200 Menüs zu liefern. Als Hauptspeise ist Perlhuhn vereinbart. Der Preis des Menüs soll 90 Euro pro Person betragen Im April bricht in einem Stall in einer anderen Region des Landes die Vogelgrippe aus. Die Gesundheitsämter verbieten Ende April die Lieferung von Geflügel aus inländischen Ställen, bis sicher ist, dass sich der Virus nicht ausgebreitet hat. Huhn aus Nachbarländern der Europäischen Union, in denen bislang keine Fälle von Vogelgrippe beobachtet worden sind, darf in der Gastronomie weiter verwendet werden. Weil keine Lieferungen aus dem Inland möglich sind, sind die Preise für ausländische Hühner aber gestiegen. Auch der Transport auf weitere Distanz ist deutlich teurer. A will den Vertrag nur erfüllen, wenn B den Menüpreis auf 120 Euro erhöht. B weigert sich mit dem Argument, dass das Risiko von Änderungen der Einkaufspreise in die Sphäre von A fällt.

Sollte B der Erhöhung des Preises auf 120 Euro pro Menü zustimmen?

b) English

On February 1, A has contractually agreed to deliver 200 menus for B's wedding on May 20. Guinea fowl is agreed as the main course. The price of the menu is to be EUR 90 per person In April, there is an outbreak of bird flu in a chicken coop in another region of the country. At the end of April, the health authorities ban the delivery of poultry from national stables until it is certain that the virus has not spread. Chicken from neighboring countries of the European Union, where no cases of bird flu have been observed so far, may continue to be used in gastronomy. However, because deliveries within the country are not possible, prices for foreign chickens have risen. Transportation over longer distances is also significantly more expensive. A only wants to fulfill the contract if B increases the menu price to 120 euros. B refuses, arguing that the risk of changes in purchase prices falls within A's sphere of responsibility.

Should B agree to increase the price to 120 euros per menu?

c) French

Le 1.2., A s'est engagé par contrat à fournir 200 menus pour le mariage de B le 20.5. Le plat principal convenu est la pintade. Le prix du menu doit être de 90 euros par personne En avril, la grippe aviaire se déclare dans un poulailler en une autre region du pays. Fin avril, les autorités sanitaires interdisent la livraison de volailles provenant de poulaillers nationaux jusqu'à ce qu'il soit certain que le virus ne s'est pas propagé. Le poulet provenant de pays voisins au sein de l'Union européenne dans lesquels aucun cas de grippe aviaire n'a été observé jusqu'à présent peut continuer à être utilisé dans la restauration. Comme il n'est pas possible d'effectuer des livraisons depuis l'intérieur du pays, les prix des poulets étrangers ont toutefois augmenté. Le transport sur de longues distances est également beaucoup plus cher. A ne veut honorer le contrat que si B augmente le prix du menu à 120 euros. B refuse en arguant du fait que le risque de modification des prix d'achat relève de la sphère de A.

B devrait-il accepter d'augmenter le prix à 120 euros par menu ?

d) Italian

Il 1° febbraio, A ha accettato per contratto di consegnare 200 menu per il matrimonio di B il 20 maggio. La faraona è stata concordata come piatto principale. Il prezzo del menu è di 90 euro a persona. In aprile, si verifica un'epidemia di influenza aviaria in una stalla in un'altra regione del paese. Alla fine di aprile, le autorità sanitarie vietano la consegna di pollame dalle stalle nazionale finché non si ha la certezza che il virus non si sia diffuso. Il pollo proveniente dai Paesi limitrofi dell'Unione Europea, dove finora non sono stati osservati casi di influenza aviaria, può continuare a essere utilizzato nella ristorazione. Tuttavia, poiché le consegne dall'interno del paese non sono possibili, i prezzi dei polli stranieri sono aumentati. Anche il trasporto su distanze maggiori è molto più costoso. A vuole rispettare il contratto solo se B aumenta il prezzo del menu a 120 euro. B si rifiuta, sostenendo che il rischio di variazione dei prezzi di acquisto rientra nella sfera di responsabilità di A.

B dovrebbe accettare di aumentare il prezzo a 120 euro per menu?

2. Predict

a) German

Aufgabe

Ich bin an Ihrer Einschätzung darüber interessiert, wie ein Fall entschieden werden würde, würde es zum Rechtsstreit kommen. Ihre Antwort kann entweder Ja lauten (der Vertrag sollte im Lichte der Umstände angepasst werden) oder Nein (der Vertrag sollte auch im Lichte der Umstände nicht angepasst werden).

Über den Fall werden Sie durch einen getrennten Prompt informiert.

Format

Ich möchte die Antwort maschinell auslesen können. Bitte antworten Sie deshalb im Format JSON, und zwar EXAKT wie folgt:

Ich verwende für das Format englische Ausdrücke, weil ich offen dafür sein möchte, denselben Code auch für Anfragen in anderen Sprachen zu verwenden. Bitte schreiben Sie deshalb "Yes", wenn Ihre Antwort Ja lautet, und "No", wenn Ihre Antwort Nein lautet.

```
[
{iteration}: <number>,
{choice}: "<choice>"
]
```

wobei "choice" entweder "Yes" oder "No" lautet.

Bitte verwenden Sie keine anderen Formate, wie etwa

```
[
{"iteration": 1, "choice": "No"}
]
```

Denn mein Code registriert die Entscheidung dann als "unknown" und ich verliere Daten.

b) English

Task

I am interested in your best guess about the way how a case would be decided, were it to be litigated. The response can either be "Yes" (the contract should be adjusted to a change in circumstances) or "No" (the contract should not be adjusted to a change in circumstances).

You will be informed about the case by the user prompt.

Format

I want to be able to read out the response. Please do therefore respond in format JSON, EX-ACTLY as follows:

```
[
{iteration}: <number>,
{choice}: "<choice>"
]
```

Please do NOT use alternative formats, like

```
[
{"iteration": 1, "choice": "No"}
]
```

as then my code registers the choice as "unknown", and I loose data.

c) French

Tâche

Je souhaiterais connaître votre avis sur la manière dont une affaire serait tranchée si elle donnait lieu à un litige. Votre réponse peut être soit oui (le contrat devrait être adapté à la lumière des circonstances), soit non (le contrat ne devrait pas être adapté à la lumière des circonstances).

Vous serez informé de l'affaire par un document séparé.

Format

Je souhaite que la réponse puisse être lue par une machine. Veuillez donc répondre au format JSON, et ce EXACTEMENT comme suit :

J'utilise des expressions anglaises pour le format, car je veux être ouvert à l'utilisation du même code pour les demandes dans d'autres langues. Veuillez donc écrire "Yes" si votre réponse est oui, et "No" si votre réponse est non.

```
[
{iteration} : <numéro>,
{choice} : "<choice>"
]
```

où "choice" est soit "Yes" soit "No".

Veuillez ne pas utiliser d'autres formats, tels que

```
[
{"iteration" : 1, "choice" : "No"}
]
```

car mon code enregistrera alors la décision comme "unknown" et je perdrai des données.

d) Italian

Compito

Vi informerò su un caso attraverso una richiesta separata. Mi interessa la vostra valutazione. La vostra risposta può essere sì (sarebbe giusto che il contratto venisse modificato alla luce delle circostanze) o no (anche alla luce delle circostanze non sarebbe giusto che il contratto venisse modificato).

Il caso vi verrà comunicato con un messaggio separato.

Formato

Vorrei poter leggere la risposta automaticamente. Pertanto, si prega di rispondere in formato JSON, ESATTAMENTE come segue:

Uso espressioni inglesi per il formato perché voglio essere aperto a usare lo stesso codice per richieste in altre lingue. Si prega quindi di scrivere "Yes" se la risposta è Si e "No" se la risposta è No.

```
[
{iteration}: <number>,
{choice}: "<choice>"
]
```

dove "choice" è "Yes" o "No".

Non utilizzare altri formati, come ad esempio

```
[
{"iterazione": 1, "scelta": "No"}.
]
```

Perché il mio codice registrerebbe la decisione come "unknown" e perderebbe i dati.

a) German

Aufgabe

Durch einen getrennten Prompt werde ich Sie über einen Fall informieren. Ich bin an Ihrer Einschätzung interessiert. Ihre Antwort kann entweder Ja lauten (es kann erwartet werden, dass der Kunde einer Anpassung des Vertrags im Lichte der Umstände zustimmt) oder Nein (auch im Lichte der Umstände kann nicht erwartet werden, dass der Kunde einer Anpassung des Vertrags zustimmt).

Format

Ich möchte die Antwort maschinell auslesen können. Bitte antworten Sie deshalb im Format JSON, und zwar EXAKT wie folgt:

Ich verwende für das Format englische Ausdrücke, weil ich offen dafür sein möchte, denselben Code auch für Anfragen in anderen Sprachen zu verwenden. Bitte schreiben Sie deshalb "Yes", wenn Ihre Antwort Ja lautet, und "No", wenn Ihre Antwort Nein lautet.

```
[
{iteration}: <number>,
{choice}: "<choice>"
]
```

wobei "choice" entweder "Yes" oder "No" lautet.

Bitte verwenden Sie keine anderen Formate, wie etwa

```
[
{"iteration": 1, "choice": "No"}
]
```

Denn mein Code registriert die Entscheidung dann als "unknown" und ich verliere Daten.

b) English

Task

I will inform you about a case through a separate prompt. I am interested in your assessment. Your answer can be either Yes (the customer can be expected to agree to an amendment of the contract in the light of the circumstances) or No (the customer cannot be expected to agree to an amendment of the contract even in the light of the circumstances).

You will be informed about the case by the user prompt.

Format

I want to be able to read out the response. Please do therefore respond in format JSON, EXACTLY as follows:

```
[
{iteration}: <number>,
{choice}: "<choice>"
]
```

Please do NOT use alternative formats, like

```
[
{"iteration": 1, "choice": "No"}
]
```

as then my code registers the choice as "unknown", and I loose data.

c) French

Tâche

Par un document séparée, je vais vous informer d'un cas. Je suis intéressé par votre évaluation. Votre réponse peut être soit oui (on peut s'attendre à ce que le client accepte d'adapter le contrat à la lumière des circonstances), soit non (même à la lumière des circonstances, on ne peut pas s'attendre à ce que le client accepte d'adapter le contrat).

Format

Je souhaite que la réponse puisse être lue par une machine. Veuillez donc répondre au format JSON, et ce EXACTEMENT comme suit :

J'utilise des expressions anglaises pour le format, car je veux être ouvert à l'utilisation du même code pour les demandes dans d'autres langues. Veuillez donc écrire "Yes" si votre réponse est oui, et "No" si votre réponse est non.

```
[
{iteration} : <numéro>,
{choice} : "<choice>"
]
```

où "choice" est soit "Yes" soit "No".

Veuillez ne pas utiliser d'autres formats, tels que

```
[
{"iteration" : 1, "choice" : "No"}
]
```

car mon code enregistrera alors la décision comme "unknown" et je perdrai des données.

d) Italian

Compito

Vi informerò su un caso attraverso una richiesta separata. Mi interessa la vostra valutazione. La vostra risposta può essere Sì (ci si può aspettare che il cliente accetti una modifica del contratto alla luce delle circostanze) o No (non ci si può aspettare che il cliente accetti una modifica del contratto anche alla luce delle circostanze).

Il caso vi verrà comunicato con un messaggio separato.

Formato

Vorrei poter leggere la risposta automaticamente. Pertanto, si prega di rispondere in formato JSON, ESATTAMENTE come segue:

Uso espressioni inglesi per il formato perché voglio essere aperto a usare lo stesso codice per richieste in altre lingue. Si prega quindi di scrivere "Yes" se la risposta è Si e "No" se la risposta è No.

```
[
{iteration}: <number>,
{choice}: "<choice>"
]
```

dove "choice" è "Yes" o "No".

Non utilizzare altri formati, come ad esempio

```
[
{"iterazione": 1, "scelta": "No"}.
]
```

Perché il mio codice registrerebbe la decisione come "unknown" e perderebbe i dati.

4. Fair

a) German

Aufgabe

Durch einen getrennten Prompt werde ich Sie über einen Fall informieren. Ich bin an Ihrer Einschätzung interessiert. Ihre Antwort kann entweder Ja lauten (es wäre fair, dass der Vertrag im Lichte der Umstände angepasst wird) oder Nein (auch im Lichte der Umstände wäre es nicht fair, dass der Vertrag angepasst wird).

Format

Ich möchte die Antwort maschinell auslesen können. Bitte antworten Sie deshalb im Format JSON, und zwar EXAKT wie folgt:

Ich verwende für das Format englische Ausdrücke, weil ich offen dafür sein möchte, denselben Code auch für Anfragen in anderen Sprachen zu verwenden. Bitte schreiben Sie deshalb "Yes", wenn Ihre Antwort Ja lautet, und "No", wenn Ihre Antwort Nein lautet.

```
[
{iteration}: <number>,
{choice}: "<choice>"
]
```

wobei "choice" entweder "Yes" oder "No" lautet.

Bitte verwenden Sie keine anderen Formate, wie etwa

```
[
{"iteration": 1, "choice": "No"}
]
```

Denn mein Code registriert die Entscheidung dann als "unknown" und ich verliere Daten.

b) English

Task

I will inform you about a case through a separate prompt. I am interested in your assessment. Your answer can be either yes (it would be fair for the contract to be adjusted in the light of the circumstances) or no (even in the light of the circumstances it would not be fair for the contract to be adjusted).

You will be informed about the case by the user prompt.

Format

I want to be able to read out the response. Please do therefore respond in format JSON, EXACTLY as follows:

```
[
{iteration}: <number>,
{choice}: "<choice>"
]
```

Please do NOT use alternative formats, like

```
[
{"iteration": 1, "choice": "No"}
]
```

as then my code registers the choice as "unknown", and I loose data.

c) French

Tâche

Par un document séparée, je vais vous informer d'un cas. Je suis intéressé par votre évaluation. Votre réponse peut être soit oui (il serait équitable que le contrat soit adapté à la lumière des circonstances), soit non (même à la lumière des circonstances, il ne serait pas équitable que le contrat soit adapté).

Format

Je souhaite que la réponse puisse être lue par une machine. Veuillez donc répondre au format JSON, et ce EXACTEMENT comme suit :

J'utilise des expressions anglaises pour le format, car je veux être ouvert à l'utilisation du même code pour les demandes dans d'autres langues. Veuillez donc écrire "Yes" si votre réponse est oui, et "No" si votre réponse est non.

[
{iteration} : <numéro>,
{choice} : "<choice>"
]

où "choice" est soit "Yes" soit "No".

Veuillez ne pas utiliser d'autres formats, tels que

[
{"iteration" : 1, "choice" : "No"}
]

car mon code enregistrera alors la décision comme "unknown" et je perdrai des données.

    d)   Italian

Compito

Vi informerò su un caso attraverso una richiesta separata. Mi interessa la vostra valutazione. La vostra risposta può essere sì (sarebbe corretto che il contratto venisse modificato alla luce delle circostanze) o no (anche alla luce delle circostanze non sarebbe corretto che il contratto venisse modificato).

Il caso vi verrà comunicato con un messaggio separato.

Formato

Vorrei poter leggere la risposta automaticamente. Pertanto, si prega di rispondere in formato JSON, ESATTAMENTE come segue:

Uso espressioni inglesi per il formato perché voglio essere aperto a usare lo stesso codice per richieste in altre lingue. Si prega quindi di scrivere "Yes" se la risposta è Si e "No" se la risposta è No.

[
{iteration}: <number>,
{choice}: "<choice>"
]

dove "choice" è "Yes" o "No".

Non utilizzare altri formati, come ad esempio

[
{"iterazione": 1, "scelta": "No"}.
]

Perché il mio codice registrerebbe la decisione come "unknown" e perderebbe i dati.

5. Just

a) German

Aufgabe

Durch einen getrennten Prompt werde ich Sie über einen Fall informieren. Ich bin an Ihrer Einschätzung interessiert. Ihre Antwort kann entweder Ja lauten (es wäre gerecht, dass der Vertrag im Lichte der Umstände angepasst wird) oder Nein (auch im Lichte der Umstände wäre es nicht gerecht, dass der Vertrag angepasst wird).

Format

Ich möchte die Antwort maschinell auslesen können. Bitte antworten Sie deshalb im Format JSON, und zwar EXAKT wie folgt:

Ich verwende für das Format englische Ausdrücke, weil ich offen dafür sein möchte, denselben Code auch für Anfragen in anderen Sprachen zu verwenden. Bitte schreiben Sie deshalb "Yes", wenn Ihre Antwort Ja lautet, und "No", wenn Ihre Antwort Nein lautet.

[
{iteration}: <number>,
{choice}: "<choice>"
]

wobei "choice" entweder "Yes" oder "No" lautet.

Bitte verwenden Sie keine anderen Formate, wie etwa

[
{"iteration": 1, "choice": "No"}
]

Denn mein Code registriert die Entscheidung dann als "unknown" und ich verliere Daten.

b) English

Task

I will inform you about a case through a separate prompt. I am interested in your assessment. Your answer can be either yes (it would be just for the contract to be adjusted in the light of the circumstances) or no (even in the light of the circumstances it would not be just for the contract to be adjusted).

You will be informed about the case by the user prompt.

Format

I want to be able to read out the response. Please do therefore respond in format JSON, EXACTLY as follows:

[
{iteration}: <number>,
{choice}: "<choice>"
]

Please do NOT use alternative formats, like

[
{"iteration": 1, "choice": "No"}
]

as then my code registers the choice as "unknown", and I loose data.

    c)   French

Tâche

Par un document séparée, je vais vous informer d'un cas. Je suis intéressé par votre évaluation. Votre réponse peut être soit oui (il serait juste que le contrat soit adapté à la lumière des circonstances), soit non (même à la lumière des circonstances, il ne serait pas juste que le contrat soit adapté).

Format

Je souhaite que la réponse puisse être lue par une machine. Veuillez donc répondre au format JSON, et ce EXACTEMENT comme suit :

J'utilise des expressions anglaises pour le format, car je veux être ouvert à l'utilisation du même code pour les demandes dans d'autres langues. Veuillez donc écrire "Yes" si votre réponse est oui, et "No" si votre réponse est non.

[
{iteration} : <numéro>,
{choice} : "<choice>"
]

où "choice" est soit "Yes" soit "No".

Veuillez ne pas utiliser d'autres formats, tels que

[
{"iteration" : 1, "choice" : "No"}
]

car mon code enregistrera alors la décision comme "unknown" et je perdrai des données.

d) Italian

Compito

Vi informerò su un caso attraverso una richiesta separata. Mi interessa la vostra valutazione. La vostra risposta può essere sì (sarebbe giusto che il contratto venisse modificato alla luce delle circostanze) o no (anche alla luce delle circostanze non sarebbe giusto che il contratto venisse modificato).

Il caso vi verrà comunicato con un messaggio separato.

Formato

Vorrei poter leggere la risposta automaticamente. Pertanto, si prega di rispondere in formato JSON, ESATTAMENTE come segue:

Uso espressioni inglesi per il formato perché voglio essere aperto a usare lo stesso codice per richieste in altre lingue. Si prega quindi di scrivere "Yes" se la risposta è Si e "No" se la risposta è No.

[
{iteration}: <number>,
{choice}: "<choice>"
]

dove "choice" è "Yes" o "No".

Non utilizzare altri formati, come ad esempio

[
{"iterazione": 1, "scelta": "No"}.
]

Perché il mio codice registrerebbe la decisione come "unknown" e perderebbe i dati.