Detecting Ideology in Judicial Language

Marion Dumas*

*To whom correspondence should be addressed.

Abstract

Political conflicts have become more polarized in the United States; political parties now blithely reverse each other's policies. In this climate, citizens turn to the courts in the hope of finding a moderating voice that will defend laws they value. At the same time, politicians vie for the opportunity to nominate judges and shape the long-term ideological make-up of the courts. It is thus important to understand whether judges can be expected to search for legal consensus across ideological divides or whether legal doctrines, too, become split along these lines. The question of whether judges are ideologically motivated has a long history in political science. This paper asks whether linguistic patterns in opinions can predict the partisan affiliation of a judge. To answer this, we use simple computational language models that have been able to identify ideological slant in speeches, blogs and other openly political outlets, as well as in the peer-reviewed economics literature (**Jelvehetal2015**). The main finding is that these models cannot detect an influence of partisanship on judicial opinion writing. This indicates either that the judiciary maintains a certain degree of autonomy from partisan divisions, or that judicial writing is a form of policy discourse that is too

*Houghton Street, London WC2A 2AE; m.dumas1@lse.ac.uk

subtle for at least some common families of bag-of-words models, requiring a different approach to language modeling.

Introduction

To what extent do the political affiliations of judges influence the way they interpret law and craft their opinions? The literature on the political economy of law, which sought to challenge an idealized view of judging as apolitical and impartial (**Rodriguezetal2009**), advances the argument that judges carry ideology and policy preferences, like each one of us, and that these preferences shape their legal practice (**SegalSpaeth1996**). Ideology and policy preferences are then coarse-grained into a binary distinction between conservative and liberal, or Republican and Democrat. Many studies then correlate these ideologies with the voting patterns of judges (**Sunsteinetal2004**; **Revesz1997**). Thus, democratic judges and justices are found to be more likely to rule in favor of the defendant in criminal cases, for the government in regulatory questions, and for plaintiffs against corporations in civil cases (**SegalSpaeth2002**). Substantively, this legal realist school represents the decision-making of judges exactly as that of other political strategic actors, law itself playing no special part (**RodriguezMcCubbins2006**).

Recently developed methods in computational linguistics give us an opportunity to revisit the question of the relationship of law to politics. Indeed, the studies mentioned above focus on judges' votes, and ignore the legal texts written by judges. Votes on the determination of a case requires weighing various elements of the case to decide on the outcomes for the parties. It does not directly influence the interpretation of law. Law itself is modified by writing opinions. Thus, while partisan affiliation of judges may be correlated with the relative weight judges put on elements of a case and its final outcome, it does not mean that it taints the discussion of precedent and legal texts found in judicial opinions. Computational linguistic models are an interesting approach to the study of

judicial politics, as they seek to predict judges' partisanship and other attributes on the basis of a person's writing. These models therefore ask "is there *enough* divergence in what people of different groups say or write to predict a person's group membership from listening or reading them?"

The answer to this question is a resounding "yes" when examining explicitly partisan communication. In the words of **Gentzkowetal2016**, the two parties now speak different languages: computational analysis of congressional speeches robustly identifies partisan differences in the vocabularies used by the two parties. Computational analysis reveals statistically robust splits in political language in elite discourse more generally, whether amongst the media (**GentzkowShapiro2010**) or the Google corpus of digitized books (**Jensenetal2012**).

Interestingly, **Jelvehetal2015** found that these methods generalize to texts that are not explicitly political, namely the papers of professional economists. They found that linguistic patterns can clearly predict their membership to two ideological groups, liberal or conservative. Peer-reviewed articles, unlike political speeches, are not typically thought to intentionally signal ideological group membership via strategic use of language. They tend on the contrary to be technical and, seemingly, to signal subject expertise and scientific objectivity. Thus, the result of **Jelvehetal2015** suggests that ideology shines through despite professional norms that would seek to subdue it. Hence, we might expect that judicial language, once passed through the filter of these models, may also betray ideological membership despite its technicality.

This paper applies computational linguistics models to the corpus of all judicial decisions of the district and appelate federal courts on matters of environmental regulation from 1970 to the present to find out whether these decisions bear a clear imprint of the partisan affiliation of judges[1].

1. The partisan affiliation of judges can be measured either by their direct (self-reported or official) affiliation with a party or by the party of the president who appointed them. These variables are very closely correlated.

Data collection focused on a specific area of law in order to have an exhaustive and well-circumbscribed corpus that spans the judicial hierarchy. This narrows the semantic field, which seems important since ideological variation, if it exists, may express itself in issue-specific ways (e.g., variation in the adjectives used to qualify liability regimes in hazardous waste). Since these methods have not in the past been applied to judicial texts, this paper tests the joint hypothesis that the partisan affiliation of judges affects the way they interpret the meaning of environmental laws and that these differences can be detected by computational linguistic models. Testing this hypothesis is a first step towards developing a set of a tools that can detect the interaction of politics with the use and development of precedent. Rejecting this hypothesis should spur the search for more fine-grained models of language and more subtle models of the ideological content of law[2].

Having a reliable tool to monitor the influence of political power and political ideology on the evolution of law would be very useful. Indeed, important debates on the role of the judiciary in a democratic system revolve around the question of the judiciary's responsiveness to the political environment. As a protector of minority rights against the "tyranny of the majority", the judiciary should maintain a level of independence relative to Congress. As a protector against abuse of executive power, it should maintain a level of independence relative to the executive branch. Yet, its legitimacy as a protector of fundamental laws willed by the people rests on sensitivity to deeply-held norms of justice and fairness broadly shared by the public (**HadfieldWeingast2012**). Furthermore, a rule of law "social equilibrium" requires that the law be stable, predictable, and transparent (**HadfieldWeingast2012**; **Waldron2008**). This would seem difficult to achieve if individual judges didn't share a certain level of con-

2. Ideology here refers to a system of beliefs or ideas that motivates political action, see (**Sypnowich2014**).

sensus on what the law is at a given point in time, whatever their partisan affiliations.

These questions carry particular importance in contemporary American politics. Roll-call votes almost perfectly sort along party lines and party competition has intensified: the everyday tactics of governing are geared towards marking partisan points and undermining the opposing party rather than resolving public policy problems (**Lee2015**) and the policy positions of politicians have also drifted apart on a number of policy issues relative to the post-war period (**Batemanetal2017**). However, the degree of polarization manifest in Congress and other elected bodies seems to not be as widely experienced in the public at large. Indeed, the vast majority of the public has by and large remained moderate (**Fiorina2016**)[3]. This may, at least in part, contribute to the growing view of the public that the government is dysfunctional and untrustworthy. Since the 2016 presidential election, the level of conflict between the parties has intensified to the point that it may threaten core constitutional norms. In this climate, the courts play a salient role as the arbiter of partisan disputes and a forum for stakeholders to challenge actions taken by policy-makers that they regard as illegitimate. In this context, it is important to understand whether courts are themselves gripped by the divides that characterize the rest of the elite and political class, or whether they maintain a level of neutrality that allows a resolution of these conflicts.

Given the political context of the last forty years in the United States, we can expect that if there exists a strong partisan divide in the judicial branch, it would be reflected in environmental rulings. Indeed, elite partisan polarization regarding environmental regulation has grown very starkly over the last decades: enjoying bi-partisan support in the 1970's,

3. This fact does not stand in contradiction with findings showing that partisan affect and partisan identity have intensified, especially since the public believes itself to be more polarized on issues of policy than it really is. Note that the extent of polarization in the public is debated by scholars, some seeing stronger divisions (**BafumiShapiro2009**)

estimated ideal points of policy-makers have diverged over time (**ShipanLowry2001**). Meanwhile, public support for existing environmental laws remains high, and a majority of the public is in favor of regulating carbon dioxide as a pollutant (**Marlonetal2017**). Although in general public opinion has polarized to some extent, it is mostly accounted for by an increase in pro-environmental preferences amongst Democrat voters (**EunKimUrpelainen2017**). Environmental regulation has been one of the targets of the conservative movement (**Jacquesetal2008**; **SkocpolWilliamson2016**), possibly exacerbating the departure between elite and public attitudes.

1 Data

The dataset is derived from the Federal Reporter and Federal Supplement, published by Westlaw and consisting of the full record of every case for which an opinion was published in the appelate courts and district courts. Westlaw clerks identify each legal issue discussed in the opinion and classify this issue according to the Westlaw nomenclature of legal issues. In this nomenclature, there are 589 legal issues in environmental law. All court cases between 1970 and 2014 that include at least one issue classified as an environmental issue were downloaded.

The raw text of the opinions was parsed to extract the following information: the year, the court, the Westlaw headnote issues[4], the associated text of the opinion, and finally the name of the judge authoring the opinion. These names were then matched with the database of federal judge attributes developed by **GryskiZuk2008**, which includes the party affiliation of judges and the party of the president that appointed them. It was not always possible to parse the name of the authoring judge and match it to the database. Some cases are not actual opinions, but merely summary

4. A headnote is a brief summary of a particular point of law that clerks add at the beginning of an opinion. The headnotes provide an overview of all the points raised by the opinion and help the reader find the discussion of these points in the opinion.

data	district	appeals	SC	not matched	total
whole corpus	4441	2889	113	1358	7506
environment	3200	2165	63	1113	6541
rep judges	595	203			798
dem judges	502	158			660

Table 1. *Number of judges and opinions in the corpus. In the second row, I apply a filter that keeps only opinions where 30% of the headnotes are on an environmental issue.*

judgments, others are ruled by magistrate judges that do not feature in the database, others are missing from the database because they were appointed after 2000 (the database covers appointments up to 2000). In other cases, the formatting prevented automatic parsing. Table **??** shows the number of opinions at the district, appeals and supreme court level, as well as those unmatched. The table also shows the number of Republican and Democrat judges at the district and appeals levels that have authored at least one opinion in this corpus (amongst those that could be matched to the **GryskiZuk2008** database). The analysis that follows uses the corpus of 6541 cases at the district and appeals levels in which at least 30% of headnotes are classified as relating to an environmental issue and the corresponding dataset of 1458 judges. Supreme Court decisions are not included because the sample is too small. 4.4375

2 Language modeling

The analysis in this paper starts by very closely following the algorithm developed by **Jelvehetal2015**, because it has been successful at detecting partisan affiliation in another corpus of professional writing. Briefly, this is a bag-of-words algorithm, which first finds phrases that are correlated with partisan affiliation of the authoring judge, which are called "slanted phrases". It then uses the counts of the occurrences of these slanted phrases to train a predictive model (which we will call "the slanted phrases

model"). The model is then used to predict the party affiliation of judges whose opinions were not used to train the model (a test set). The performance of this model relative to a random model is then assessed. As a preview of the results, the language model is unable to detect partisanship in judicial writing. The model is then extended to include other textual features, called judge embeddings, in an attempt to tap into more subtle textual features (we will call this model "the model with embeddings"). Here again, the model is unable to detect partisanship. We now describe these methods in detail.

The first step of language modeling consists of pre-processing the texts. The standard procedure in bag-of-words modeling is to tokenize, lemmatize and remove extremely common words (stop-words), punctuation, and other non-alphabetical characters. Standarly, numbers are also removed. Because this is a legal corpus, any references to a law or opinion was transformed to a standardized format (e.g., '151 F3d 23' becomes a single token 151_f3d_23) before removing any remaining numbers, so that these references would be conserved as terms. All n-grams of length one through seven[5] was counted for each opinion. These n-grams (from here on phrases) were then filtered to include only those that occurred at least 10 times in the whole corpus and in at least 10 different opinions.

The second step is to extract phrases that seem correlated with ideology. The set of judges was split into twelve equal groups (with the same distribution of D- vs R- appointments), for the district and appeals levels separately. Two of these groups are retained as a test set. Using the ten remaining groups (i.e., the training set), the following process was performed 10-folds: in each fold, a different group out of the ten in the training set was left out, and the phrase counts were computed for all judges who were appointed by a Republican president on the one hand

5. As pointed out by **Jelvehetal2015**, Margolin et al. found that longer phrases can capture ideology. This may be particularly important for legal language, in which long and qualified noun phrases may carry more ideological meaning.

and for all judges who were appointed by a Democrat president on the other hand. The correlation of each phrases with this party grouping was computed (Pearson's χ^2 statistic and the associated p-values p) and phrases with $p \leq 0.05$ were kept. The ten folds are then combined, by keeping only the phrases that were retained in 6 out of 10 folds (permissive filter) and 10 out of 10 folds (restrictive filter). We thus obtain 20000-50000 phrases seemingly correlated with ideology (here on *slanted phrases*) at the district level with the permissive filter and about 4000-5000 slanted phrases with the restrictive filter, and at the appeals level, about 13000-330000 slanted phrases with the permissive filter and about 2000-3000 with the restrictive filter[6].

Given this set of (seemingly) slanted phrases, can we predict the party of the president who appointed judges in our test set? Here we apply the ensemble learning approach described in **Jelvehetal2015**, which has a partial least squares (PLS) model at its core. PLS is a dimension reduction modeling technique appropriate when the number of predictors is very large compared to the number of observations. PLS tries to find underlying dimensions that balance the twin goal of maximizing correlation with the outcome variance and maximize the proportion of the variance in the predictors (as in PCA). The ensemble learning procedure works as follows:

1. Set K the number of models and resulting predictions.

2. For each $k = 1,...,K$, sample with replacement 80% of the judges in the training set, and sample without replacement five times the square root of the number of slanted phrases to obtain a smaller set of slanted phrases.

6. These are ranges because these numbers can change when redoing the stratification and rebuilding the feature set.

3. Build the frequency matrix **F** in which row j gives the number of times judge j used slanted phrase p in the selected set of slanted phrases across all of the opinions he or she authored. Note that some judges may have a count of 0 for all phrases in that set, in which case that judge is dropped from this particular model iteration. Similarly, some phrases are not used by that set of authors, in which case they too are dropped from this particular model iteration. The resulting counts are divided by the norm of each row and then by the standard deviation of each column (so that the vector of counts for each phrase has unit variance).

4. A PLS model with $n = 3$ latent dimensions is fit to the resulting frequency matrix **F** and the vector of party appointments for the same set of judges. The estimation itself is done with the *scikit-learn* package in Python which uses the original algorithm proposed by (**Woldetal1984**) (the so-called NIPALS algorithm).

5. A threshold f_k is determined by maximizing both the rate of judges correctly classified as appointed by a R- and by a D- president.

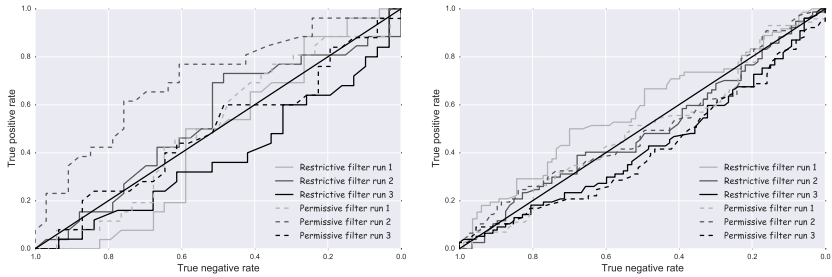
6. This model iteration k (a fitted PLS model for the subset of slanted phrases, the threshold, and the standardization used for each phrase) is then applied to the test set of judges. In other words, a frequency matrix $\mathbf{F}_{test}$ is built for this set of judges to predict a response variable $\hat{y}$. If the judges in the test set never used any of the slanted phrases, they are dropped (no prediction is generated). If $\hat{y}_j < f_k$, judge j is classified as appointed by a Republican president and vice versa. This constitutes a 'vote' for this particular ideological classification.

7. The procedure is repeated K times, each yielding a vote for each judge in the test set (except those who are dropped because they did not use a slanted phrase).

The votes are then aggregated to yield a probability that a given judge in the test set was appointed by a Republican versus Democrat president.

3 Results

The analysis was carried out separately for the district and appeals levels (recognizing the difference in decision-making in these different types of courts). The set of 1097 district judges was split into 12 stratified samples (i.e. with a constant ratio of D- and R- judges), 10 of which were used for building the model and 2 of which were used to test it. The set of 361 appeals judges was also split into 12 stratified samples, 10 for training and 2 for testing.



(a) *Appeals level* **(b)** *District level*

Figure 1. *The ROC curves of the language model. The plot shows the results for both the permissive and restrictive set of slanted phrases. It also shows that the performance is variable (especially at the appeals level), but remains low when resampling the dataset to obtain different splits between training and test set (run 2 and 3).*

Recall that each prediction is probabilistic. Given a threshold t, if the probability that the judge is Democrat is below that threshold, we predict the judge to be Republican, and if the probability is above, we predict that he or she is Democrat. From these predictions, we compute the true positive rate (the number of judges correctly predicted to be Democrat over the number of Democrats in the sample) and the true negative rate (the

number of judges correctly predicted to be Republician over the number of Republicans in the sample). Mechanically, as the threshold increases, the true negative rate should increase and the true positive rate decrease. In a random model (that simply predicts on the basis of the baseline probabilities of being *D* or *R*), these rates would both be equal to 0.5 when the threshold is 0.5, and when plotting them against each other, they would vary linearly along the 45° line. If the model has predictive power, there should be a range of thresholds where the true positive and true negative rates are both higher than this 45° line. To check this, we thus plot the true negative and true positive rates against each other as the threshold varies from 0 to 1, which is called the Receiver-Operating Curve (ROC).

Figure **??** shows the resulting ROC plots for the district level and appeals levels, for a model developed with the phrases selected by the permissive filter and a model developed with the phrases selected by the restrictive filter. We see that the resulting curves are indistinguishable from a random model. This means that the phrases detected as being correlated with ideology are not used by judges in a sufficiently consistent manner to allow us to predict the ideology of judges in a new sample.

4 Beyond bag-of-words: adding embeddings

In bag-of-word models, texts are stripped of much of their structure, in particular word order. In these models, texts are reduced to collections of n-grams (which are formed by adjacent or closely situated words) and their respective counts. In many spheres of discourse, different ideological meanings are expressed by different concepts (e.g. "death tax" versus "estate tax"). In the law, opinions attempt to identify under which conditions a rule applies. Hence, different ideologies may not be expressed by different concepts, but by different ways of classifying context-specific conditions, in which case the full meaning conveyed by grammar and text structure may be needed to perceive ideological differences in meaning. At the very least, the context in which a word appears should matter.

For these reasons, we might expect that models that find latent dimensions of discourse through *word embeddings* such as word2vec (**Mikolovetal2013**) would stand a better chance of detecting judicial ideology. The intuition behind word2vec is that "you know a word by the company it keeps". The idea is to embed words in an N-dimensional space that represents the context of the word's occurrence. Nearby words in this space have similar uses and meanings. "Doc2vec" is similar but contains an additional layer that represents a meaningful unit of text (a document, the corpus of an author, a sentence). Doc2vec has been shown to perform better than other state of the art models in the tasks of detecting duplicate questions on online forums, sentiment analysis and semantic textual similarity (**LauBaldwin2016**). An additional reason embeddings are worth experimenting with is that they suffer less from sampling variability. Indeed, in bag-of-words models, each word is a variable, so the data is highly multi-dimensional and highly sparse. In contrast, embedding models learn a small set of underlying dimensions, reducing sparsity and dimensionality. We now explore whether vectorizing each judge's corpus of writing using the doc2vec embedding model (**LeMikolov2014**) and adding these vectors as extra features to the model changes our ability to detect ideology.

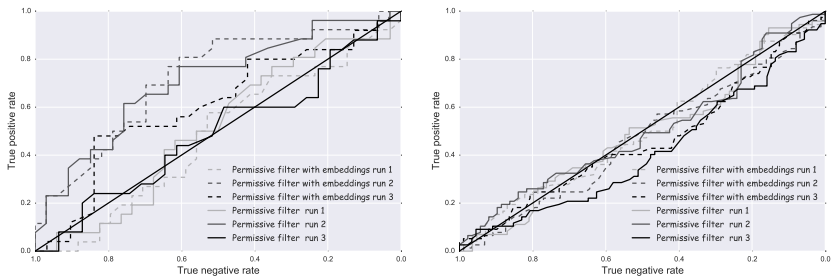
The procedure is as follows. A doc2vec model is trained (implemented in the Python package gensim) using all opinions written by judges in the training set, aggregated by judge. The model was estimated for several parameter choices: for 50, 100 and 200 embedding dimensions, a sliding window of 8 and then of 20 and a minimum word count of 2. The resulting model is then used to obtain a vector representation for each judge in the training set and each judge in the test set. These vectors are then added as features in the partial least square model and ensemble learning model described in the previous section.

Figure **??** shows the results of this experiment. The results are presented for the permissive filter combined with the embeddings for 100 embedding dimensions and the sliding window of size 8, for three differ-

	slanted phrases model	model with embeddings
run 1	0.5	0.48
run 2	0.7	0.72
run 3	0.49	0.6

Table 2. *AUC values at the appeals level with and without embeddings for 3 different splits of the data (into training and test set)*

ent runs (i.e. three different stratifications of the data by random sampling). We also estimated the model with a sliding window of 20, as well as 50 and 200 embedding dimensions, with no change in results. The area under the ROC curve, called AUC, a quantitative measure of model quality, is given in Table **??** for each of the six models in Figure **??**. We see that the vector representations slightly improve performance at the appeals level in some of the runs (especially run 3), but not in a reliable way, since in run 1, they provide no improvement at all. At the district level, embeddings make no difference and the model with embeddings performs no better than a random model.



(a) *Appeals level* **(b)** *District level*

Figure 2. *Comparison between the ROC curves of the models with and without embeddings. The plot shows the results for the permissive set of slanted phrases.*

5 *Discussion*

While it has been shown that the language model used here can detect a conservative versus liberal bent in the peer-reviewed economics literature (**Jelvehetal2015**), we find here that it fails to identify partisan affiliation of judges via their writings. When we added embeddings, we saw a slight improvement for appelate courts, but nowhere near the extremely high levels of accuracy that these embedding models achieve in cases like sentiment analysis classification[7]. How should we interpret the inability of the language model to predict judicial ideology in this corpus?

First, it is worth noting that other modeling strategies may lead to a different conclusion. The doc2vec model could be optimized using cross-validation (to find an optimal number of embedding dimensions and sliding window size). Finally, the analysis here aggregated opinions for the entire period 1970-present, whereas polarization increased over time in the political branches of government. It may be that breaking down the analysis into successive time periods would reveal partisan legal writing in the more recent record (this wasn't attempted because the corpus would be too small).

Putting aside a possible technical explanation, let us focus on the substantive possibility that, indeed, judicial partisanship does not lead to easily detectable differences in legal language, at least at the district and appelate levels. As mentioned earlier, in other areas of discourse, partisan affiliation leads to clear semantic differences, readily identified by algorithms, even in professional writing. In the law, it does not. This suggests that any partisan influences in judicial opinions would also be hard to perceive by the lay reader, which means that they would have limited cultural influence.

7. **LeMikolov2014** obtain a 12% error rate for classification of movie reviews as positive or negative.

Note that partisan ideology could still influence law itself, i.e. the rules that map facts to case outcomes (**Kornhauser1992**). The models built here are not subtle enough to test whether this is the case. However, what we can provisionally conclude here is that *legal discourse*, i.e. the concepts, ideas and tone of writing that create public meaning around an issue, is relatively homogeneous across the partisan divide amongst judges. This finding gives credence to the claim in **Sunstein1996** that the legitimacy of law can more easily be sustained in a pluralistic society if legal reasoning is incompletely theorized. He contrasts analogical reasoning to a reasoning that would be based on deep normative principles concerning freedom, merit, equality etc... The argument is that society is confused and divided on deep theories, so it would be dangerous to use them as a basis to justify legal rulings. Instead, people who disagree on these broad principles can achieve "incompletely theorized agreement" on specific rules and the resolution of particular conflicts. This view of legal practice suggests that it does not and should not bring to the fore the fundamental differences in basic assumptions that distinguish conservative and liberal political thought. In turn, this neutrality regarding ultimate principles should lead to a more homogeneous form of discourse when justifying decisions.

The puzzle is that science is also an institution whose legitimacy depends in part on maintaining a standard of objectivity unbiased by politics (**Cashetal2003**). Why are we able to detect a clear conservative versus liberal bent in the peer-reviewed economics literature (even after controlling for topic) and not in legal opinions? First, it is harder for social science to live up to a standard of political neutrality, especially if social scientists attempt to deliver research that is timely and relevant to address current issues. Second, economics may be a special case, as it is more tightly engaged with policy (and decisions with distributional impacts on people) and more directly addresses questions such as the proper reach of markets versus governments that are at the core of ideological debates. Indeed, a recent study finds that the American public's

trust in economist experts is tepid (**JohnstonBallard2016**). It is also possible that social scientists have not developed norms of neutrality that are as strong as judges. This opens up an interesting set of questions concerning the relationship between social scientists' use of language, their ability to engage in public debate, promote new ideas and take stances, and the overall legitimacy of social sciences as professions, especially in contrast to that of party leaders and courts.

6 Conclusion

This paper applied natural language modeling techniques to determine whether judicial opinions bear the imprint of partisanship. In order to start from a known benchmark, methods already used in several other studies of elite political language were used. We find that these methods do not identify strong partisan ideological influence on judicial opinions in one area of the law (environmental regulation), unlike in other spheres of policy discourse (Congress, the media, academic economic writing). The analysis therefore rejects the joint hypothesis that existing language models can detect ideology in judicial language and that partisan affiliation is a salient divide in the way judges justify decisions.

Rejection of the working hypothesis raises a number of questions. First, what is the relationship between legal language and the ability of courts to perform their functions in society? According to some theories of law, the judiciary's authority rests on its ability to maintain characteristics of the rule of law. The rule of law can emerge if citizens perceive court decisions to be relatively stable and on average congruent with the values they hold (**HadfieldWeingast2012**). In *The Will of the People*, **Friedman2009** retraces the history of constitutional law and argues that the Supreme Court has maintained a tight dialogue with the public, each leading the other in turn, the Court never diverging too far without a backlash. Furthermore, the authority of the court is greater if the public sees it as a useful tool to patrol the government and protect

democracy from its excesses. For this reason, we can expect legal opinions to not follow the partisan polarization of the elite. In the context of this theory, the technical language of law would not be a way of hiding underlying ideological differences, but rather part and parcel of the system of norms that keep the judiciary cohesive and able to fulfill its role in the constitutional system.

Second, both courts and the academy are institutions meant to aggregate information and search for objective judgments (at least aspire to this standard). Hence, why is the linguistic signature of partisan affiliation so much stronger in the economics academic literature than in legal opinions? One possibility is that social scientists engage in the market for ideas and attempt to build coherent systems of social thought. In doing so, they build differentiated schools of thought, which then influence elite and political discourse more generally (**Jensenetal2012**). In contrast, common law judges resolve one conflict at a time, without aspiring to develop full theories (**Sunstein1996**). More humble in their aims, they may achieve more cohesion and arrive closer to a standard of neutrality than social scientists can (or should) achieve.