American Political Science Review (2020) 114, 1, 291–301
doi:10.1017/S0003055419000522 © American Political Science Association 2019

Letter

Understanding Delegation Through Machine Learning: A Method and Application to the European Union

L. JASON ANASTASOPOULOS *University of Georgia*
ANTHONY M. BERTELLI *Bocconi University and Pennsylvania State University*

Delegation of powers represents a grant of authority by politicians to one or more agents whose powers are determined by the conditions in enabling statutes. Extant empirical studies of this problem have relied on labor-intensive content analysis that ultimately restricts our knowledge of how delegation has responded to politics and institutional change in recent years. We present a machine learning approach to the empirical estimation of authority and constraint in European Union (EU) legislation, and demonstrate its ability to accurately generate the same discretionary measures used in an original study directly using all EU directives and regulations enacted between 1958–2017. We assess validity by training our classifier on a random sample of only 10% of hand-coded provisions and replicating an important substantive finding. While our principal interest lies in delegation, our method is extensible to any context in which human coding has been profitably produced.

INTRODUCTION

Delegation of powers represents a grant of authority by politicians, who hold policy-making power as a constitutional matter, to an agent or set of agents, whose powers are determined by the conditions in enabling laws. Beyond the policy substance of the powers to be exercised, delegation so defined has two essential elements that any quantitative strategy to measure it must capture: (1) the identity of the agent or agents who hold policy-making *authority* as a result of the delegation and (2) any *constraint* or conditions on the exercise of that authority. For instance, a delegating law might state that an environmental agency (authority) shall make rules and regulations to protect wildlife, but only after consultation with the public (constraint). As important as it is, delegating legislation is not the only kind of law enacted by governments, and researchers have applied restrictive selection criteria to yield a small sample on which to apply a labor-intensive coding framework. We offer a machine learning approach for estimating authority and constraint directly from legislative texts that leverages an existing coding framework, and then apply it to over 59,000 European Union (EU) laws from 1958–2017[1].

Delegation in the EU requires an important choice of agents: the European Commission (EC), the principal executive body, or the national administrations of the EU member states. Franchino (2004, 2007) creates a human coding framework to test his theory of this choice in 158 *major* laws from 1958–93. His framework produces delegation (authority) and constraint ratios, combining them into a composite index of discretion as introduced in a study of delegation by the US Congress (Epstein and O'Halloran 1999).

We reproduce and extend this human coding effort in four stages. First, we train a series of gradient-boosted tree (GBT) text classifiers to reproduce Franchino's coding on his full original sample of major legislation. Second, we machine-code *all articles* in EU legislation between 1958–2017 using these classifiers. Third, we estimate delegation and constraint ratios and the discretion index from the larger machine-coded dataset. Fourth, we apply a discontinuity test to assess the consistency of these measures after the end of the human-coding period in 1993. Fifth, we assess the validity of our measures by replicating two empirical results from Franchino (2007) regarding the decision rule under which laws are enacted using out-of-sample predictions for discretion indices.

We contribute an alternative to the growing variety of methods for capturing delegation through proxy measures on a large scale. Various studies use the number of words in a statute to capture constraint, arguing that as word count increases, agent discretion is more restricted (Clinton et al. 2012; Huber and Shipan 2002; Huber, Shipan, and Pfahler 2001). Vakilifathi (2017) improves on this measure by distinguishing between optional and

L. Jason Anastasopoulos, Assistant Professor of Political Science and Assistant Professor of Public Administration and Policy, University of Georgia, ljanastas@uga.edu.
Anthony M. Bertelli, Professor of Political Science, Bocconi University; Sherwin-Whitmore Professor of Public Policy and Political Science, Pennsylvania State University, bertelli@psu.edu.

We thank Fabio Franchino for providing the data and for helpful comments. Nicola Palma, Maulik Shah and Giulia Leila Travaglini provided excellent assistance in data collection and preparation. We thank Moritz Osnabrugge, Matia Vannoni, Arthur Spirling, Erik Voeten and Michael Bailey for helpful comments. Replication files are available at the American Political Science Review Dataverse: https://doi.org/10.7910/DVN/FF3DQM.

Received: June 25, 2018; revised: July 8, 2019; accepted: August 19, 2019; First published online: November 11, 2019.

[1] To the best of our knowledge, this collection encompasses all of the active English language EU legislation currently available on EUR-Lex, a comprehensive database of European Union legislation maintained by the EU: https://eur-lex.europa.eu/homepage.html.

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

mandatory provisions by identifying contextual triggers such as the use of "shall" versus "may." O'Halloran et al. (2018) employ a similar methodology, exploring changes over time in key words suggesting delegation. Ash, Morelli, and Vannoni (2018) contribute to a more formal natural language processing (NLP) approach, first defining a set of rules or legal ontologies that are associated with important delegation terms, parts of speech, and the sentence structure of legislative documents, identifying the presence or absence of those terms, parts of speech, and structures as a means of measuring delegation. These methods make assumptions about how authority and constraint become manifest as key terms or phrases in legal texts, and measure these concepts as a function of the frequency with which key terms are identified. Our approach avoids such assumptions, replicating human codings that were based on full-text, expert readings of a wide variety of legal documents. Recent work by O'Halloran et al. (2015) on financial regulation in the US uses machine learning to replicate a five-point "discretion level" for the full text of a regulation. By focusing on a finer-grained unit of analysis—one that is much closer to what a human coder encounters—we organically replicate, similar to a human coder, measures of authority, constraint, and discretion precisely as defined by Epstein and O'Halloran (1999) and Franchino (2004, 2007). Moreover, our method is extensible to *any* context in which human coding has been profitably produced.

MEASURING AUTHORITY AND CONSTRAINT

Measures of authority and constraint in Franchino (2004, 2007) are calculated using delegation and constraint ratios as defined by Epstein and O'Halloran (1999). In both the EU and US cases, these ratios are combined to create an index of discretion. To capture the identity of agents (authority), the *delegation ratio*, Δ_i, for law i represents the number of provisions delegating authority, D_i, to national administrations or to the EC divided by the total number of provisions in the law P_i:

$$\Delta_i = \frac{D_i}{P_i}. \quad \textbf{(1)}$$

A *constraint ratio* captures restraint on the authority of national administrations or the EC to implementing policy. This ratio is constructed by dividing the total number of restraint categories used in law i, C_i, by the possible number of categories of restraint, R_i.[∅] (NB: A supplementary materials document covers a variety of technical aspects of our method and ∅ indicates specific points on which it provides helpful details.)

$$\mathcal{C}_i = \frac{C_i}{R_i}. \quad \textbf{(2)}$$

A *discretion index* δ_i, is created by subtracting the product of (1) and (2) from (1) (Franchino 2004, 283):

$$\delta_i = \Delta_i - [\mathcal{C}_i \times \Delta_i]. \quad \textbf{(3)}$$

Essentially, the authority granted to national administrations or the EC is down-weighted by the constraints placed on that authority in an enabling law. Beyond the EU, the foregoing ratios and index are used in a variety of contexts (e.g., Lavertu 2012; Clouser McCann 2016).

METHODS

We employ a *supervised* machine learning (ML) framework using gradient-boosted decision trees (GBTs), a decision tree algorithm[2] for identifying provisions within acts delegating authority or imposing constraints on national administrations and the EC. We chose GBTs for this task because they perform well in text classification contexts[3] and are among the most interpretable algorithms (Lakkaraju, Bach, and Leskovec 2016).[∅] GBTs allow us to estimate our quantities of interest directly from texts and to learn more about the *language* of authority and constraint in the EU.

ML methods learn patterns from human-coded data and have enjoyed success in a variety of classification problems in the social sciences (Benoit et al. 2016; Grimmer and Stewart 2013). Both the quality of human-labeled data and the model imposed on those data by the algorithm jointly influence these classifiers' ability to correctly identify and to reproduce human-coded labels. By training a series of GBT classifiers using labeled EU acts provided by Franchino (2004, 2007), we can verify that care was taken by human coders in labeling the data.[∅] However, because our classifiers are trained using a sample of EU acts, it is crucial to recognize their limitations and to assess whether systematic biases exist in out-of-sample predictions. Our training data, which consist of 8,417 coded provisions from 158 EU laws, have two limitations that must be carefully considered. First, Franchino (2004, 2007) collected and coded only "major" legislation (Mayhew 1991). Second, his coded legislation was enacted between 1958–93 while the laws that we examine spans 1958–2017.

These limitations suggest that predicting our quantities of interest with this classifier using out-of-sample texts may result in systematic bias that can affect the interpretation of the results. To address these concerns, we develop a discontinuity test that leverages classifier conditional probabilities and assess how these probabilities differ among different types of legislation. As we discuss below, this test reveals no evidence of systematic bias that could affect substantive interpretations.[∅] We further assess the face validity of our measures by replicating a key substantive result from Franchino (2007) using discretion indices constructed with 90% out-of-sample predictions, and reach the same substantive conclusions.

Sequence of the Estimation Framework

Our empirical strategy has two phases outlined in Table 1. In Phase I, several GBT classifiers are trained on the Franchino (2004, 2007) hand-coded dataset.[∅]

[2] Often implemented in Python and R as the software library *XGBoost*.
[3] We also compare relative performance of GBTs with other well-known machine learning algorithms such as support vector machine (SVMs), regularized logistic regression and random forests and present the results in the supplmentary materials.

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

TABLE 1. Phases and Stages of Analysis

Phase	Stage	Action	Data
I	1	Training, testing	Labeled EU provisions, 1958–93
	2	Performance	Accuracy (a_m)
			Precision (π)
			Recall (sensitivity) (σ_1)
			Combined precision & recall F_1
II	1	Classification	Unlabeled provisions from EU legislation 1958–2017
	2	Delegation ratio	$\hat{\Delta}_i$
		Constraint ratio	$\hat{C}_i$
		Discretion index estimation	$\hat{\delta}_i$

The second stage of Phase I assesses the relative performance of each trained classifier to identify its strengths and weaknesses. Performance for each classifier is captured by four metrics.[7] These performance statistics allow us to determine the extent to which the GBT classifiers are appropriate for identifying each concept in EU legislation. They also provide information about the systematic biases present in the substantive measures of authority, constraint, and discretion that we estimate in Phase II.

DATA

Our data include the full text of all 158 pieces of "major legislation" described in Franchino (2004, 2007). Each law was divided into its constituent provisions, effectively articles and sub-articles, which yielded a total of 8,417 provisions for all 158 pieces of legislation. Using Franchino's original codebook containing coding instructions for each law (Figure 1 provides a sample), each of the provisions was prepared for training with the

FIGURE 1. Sample of Franchino's Coding Guide for One Piece of Legislation

Number **360L0921**
Title
First Directive for the implementation of Article 67 of the Treaty
Description: Liberalisation of movement of some types of capital
Major provisions **M = 18**
Provisions delegating authority **D = 5**
Number of constraints **C = 2**
Provisions delegating authority to M/S **Dg = 4**
1) M/S to ensure that transfers are made at rates similar to those ruling for payments relating to current transactions – art. 2.2
2) M/S may confine application of art. 2.1 (Annex 1, List B) to some financial institutions (temporary measure, but no time limit specified) – art. 2.3
3) M/S may maintain or reintroduce exchange restrictions (Annex 1, List C) if they support economic policy objectives – art. 3.2
4) M/S to take measures of simplification – art. 5.2
Note: Article 3.2 asserts that M/S may reintroduce exchange restrictions, this is an extension of a power that would have otherwise been relinquished as a result of article 3.1.
Not included: Article 7 asks the M/S to inform the Commission of measures that go beyond obligations of directive and that amend List D of Annex 1 (not included because it is a requirement of information, it is not delegation but a sign that M/S retains their powers in these areas); General call for M/S to adopt the measure in pursuance of the directive – art. 7b; Article 5.1 asserts that M/S can verify transactions and take measures against infringements, this is a 'no prejudice provision', this power would have not relinquished as a result of this Directive.
Number of constraints to M/S **Cg = 2**
1) Consultation: Consult Commission when M/S maintains or reintroduces restrictions (Annex 1, List B) – art. 3.2
2) Rule-making requirements : Detailed list in the Annexes of the type of capital movements where M/S can act – Annexes I and II
Provisions delegating authority to Commission **Dc = 1**
1) Commission to initiate art. 169 infringement procedure (List A, Annex I) – art. 1.2
Not included: Commission to examine and issue recommendations – art. 2.2, 2.3, 3.2, 3.3; Commission to receive information – art. 7; Commission to receive a report from the monetary committee – art. 4
Number of constraints **Cc = 1**
1) Rule-making requirements : Detailed list A of Annex I of the capital movements where the Commission can act (It can also be considered an exemption, but it is only one type of constraint)
Not included: Consultation of the Monetary committee because there is no delegation of powers – art. 2.2, 2.3, 3.2
****

Notes: Source: Codebook ("Textual Analysis of Main Acts") Available at http://users2.unimi.it/fabiofranchino/

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

TABLE 2. Performance Metrics for Authority and Constraint Classifiers Ordered by F_1 Score With Training Fraction of Positive Class

Type	Accuracy	Precision	Recall	F_1	% Positive
Authority classifier					
Delegation to…					
National administrations	0.911	0.725	0.735	0.730	0.153
EC	0.959	0.703	0.812	0.754	0.064
Constraint classifiers (EU member states)					
Consultation requirements	0.996	0.619	0.867	0.722	0.007
Appeals procedures	0.996	0.636	0.583	0.609	0.004
Spending limits	0.995	0.600	0.400	0.480	0.003
Rulemaking requirements	0.893	0.386	0.400	0.393	0.084
Time limits	0.977	0.271	0.361	0.310	0.010
Reporting requirements	0.986	0.212	0.467	0.292	0.004
Executive action required	0.996	0.250	0.143	0.182	0.002
Executive action possible	0.996	0.125	0.250	0.167	0.001
Constraint classifiers (EC)					
Executive action possible	0.988	0.753	0.921	0.828	0.028
Public hearings	0.999	0.667	1.000	0.800	0.003
Legislative action possible	0.995	0.643	0.529	0.581	0.006
Consultation requirements	0.986	0.489	0.676	0.568	0.013
Rulemaking requirement	0.969	0.397	0.391	0.394	0.029
Reporting requirements	0.989	0.276	0.533	0.364	0.004
Executive action required	0.990	0.292	0.438	0.350	0.005
Time limits	0.996	0.333	0.286	0.308	0.003

ML algorithm by creating binary variables indicating whether the provision delegated authority to or imposed constraints on national administrations or the EC.

Of these 158 pieces of legislation, five were available only in the French language and were translated into English using the Google Translation API[4] and included in the training data.[5] An additional 59,423 laws enacted between 1958–2017 were collected from the *EUR-Lex* database (https://eur-lex.europa.eu/homepage.html).

PHASE I RESULTS

Prior to analysis, the text of each provision for training was pre-processed using the *quanteda* package in R (Benoit et al. 2018). Provisions were tokenized into unigrams and bigrams; special characters, punctuation, and stopwords were removed; and each term was stemmed. The cleaned, tokenized texts were then transformed into a document-feature matrix (DFM) resulting in a very sparse 8,417 × 79,088 DFM. Because tree-based methods are sensitive to poorly predictive terms, automatic sparsity reduction using a 99% threshold was imposed on the DFM, yielding an 8,417 × 1,278 DFM used for algorithm training and performance assessment. Final models were selected using a process of hyperparameter tuning with five-fold cross-validation. As Table 2 demonstrates there were significant class imbalances in the training data. These were handled via a process of re-weighting positive labels using a positive predictive-weight tuning parameter, namely, the fraction of the negative over positive examples in the training data.

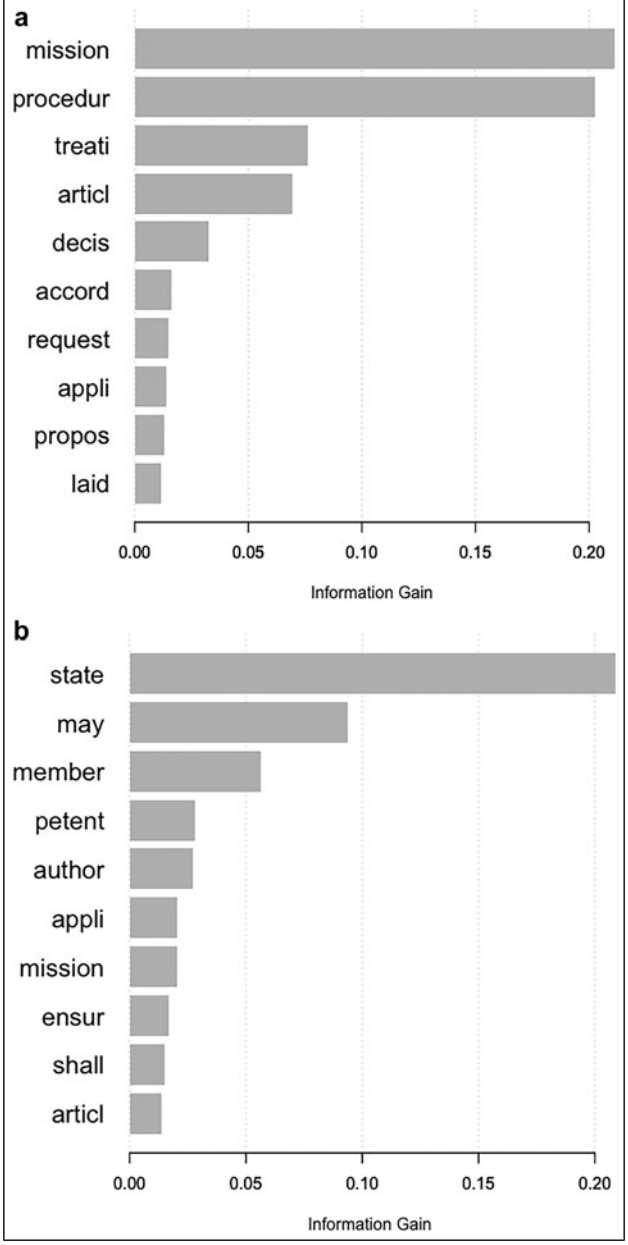
Table 2 contains performance metrics for each of the classifiers. Figures 2a and 2b contain information about term importance for the EC and national administration authority classifiers. The overall quality of each classifier is captured by the F_1 statistic, which combines precision and recall (sensitivity) in a single metric ranging from 0–1. Recall refers to classifier performance in detecting authority and constraint, while precision refers to its ability to distinguish between correctly and incorrectly labeled provisions. The precision and recall statistics in Table 2 suggest that our trained classifiers perform well for detecting authority and some categories of constraint (high recall), but that they may tend to overestimate constraint (low precision for some constraint categories). Figures 3a, 3b and 4a, 4b plot predicted versus hand-coded delegation ratios averaged over each year in the training sample, 1958–93 using 50% out of sample predictions. Both show that our classifiers accurately

[4] https://cloud.google.com/translate/docs/.

[5] Performance results for the trained classifier with and without the French language legislation are included in the supplementary materials. Including translated legislation led to improvements in the performance of the trained classifier, and so these results are reported in the main text.

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

FIGURE 2. Terms Distinguishing Provisions Which Contain (a) Delegation of Authority to the EC and (b) Delegation Of Authority to National Administrations



track aggregate trends in delegation and constraint over time.

Figures 2a and 2b contain examples of information about the importance of terms in identifying provisions delegating authority as an example.[7] The most informative terms for identifying EC delegation (including "mission," "procedure," "treaty"), and those for national administration delegation (including "state," "may," "member") make substantive sense. The former relate to formal mechanisms for authority codified in EU law, while the latter suggest permission.[7]

PHASE II RESULTS

Because provisions are difficult to parse for automated text analysis, the unit of analysis for our population of 59,423 laws is the article, which is a clearly demarcated section of EU directives and regulations addressing related topics.[7] Although multiple provisions can be nested within an article, the cleaner identification of the unit and the additional information for the classifier to process are important justifications for our choice.

We aggregate classifier predictions from articles to compute a delegation ratio ($\hat{\Delta}_i$), a constraint ratio ($\hat{C}_i$) and a discretion index ($\hat{\delta}_i$) for each of the EU laws collected. This involved three steps. First, each of the EU laws collected were parsed using NLP methods, resulting in 404,852 articles. Second, articles were cleaned and transformed into a DFM in a manner identical to that discussed in Phase I. Finally, each of the trained classifiers were applied to the articles, and each received a binary label indicating whether it contained content delegating authority to or imposing any category of constraint on national administrations or the EC. Estimates for each of the aforementioned quantities were then aggregated for each law.

Descriptives

Our trained classifiers labeled provisions from all directives and regulations available in the *EUR-Lex* database between 1958–2017. In the descriptives below, we focus on the years 1970–2017 given both the greater production of legislation and the available selection of laws in *EUR-Lex*. Figure 5 shows the growth of law availability after 1970. By 1980, the annual number of laws in our sample is above 1,000, on average.

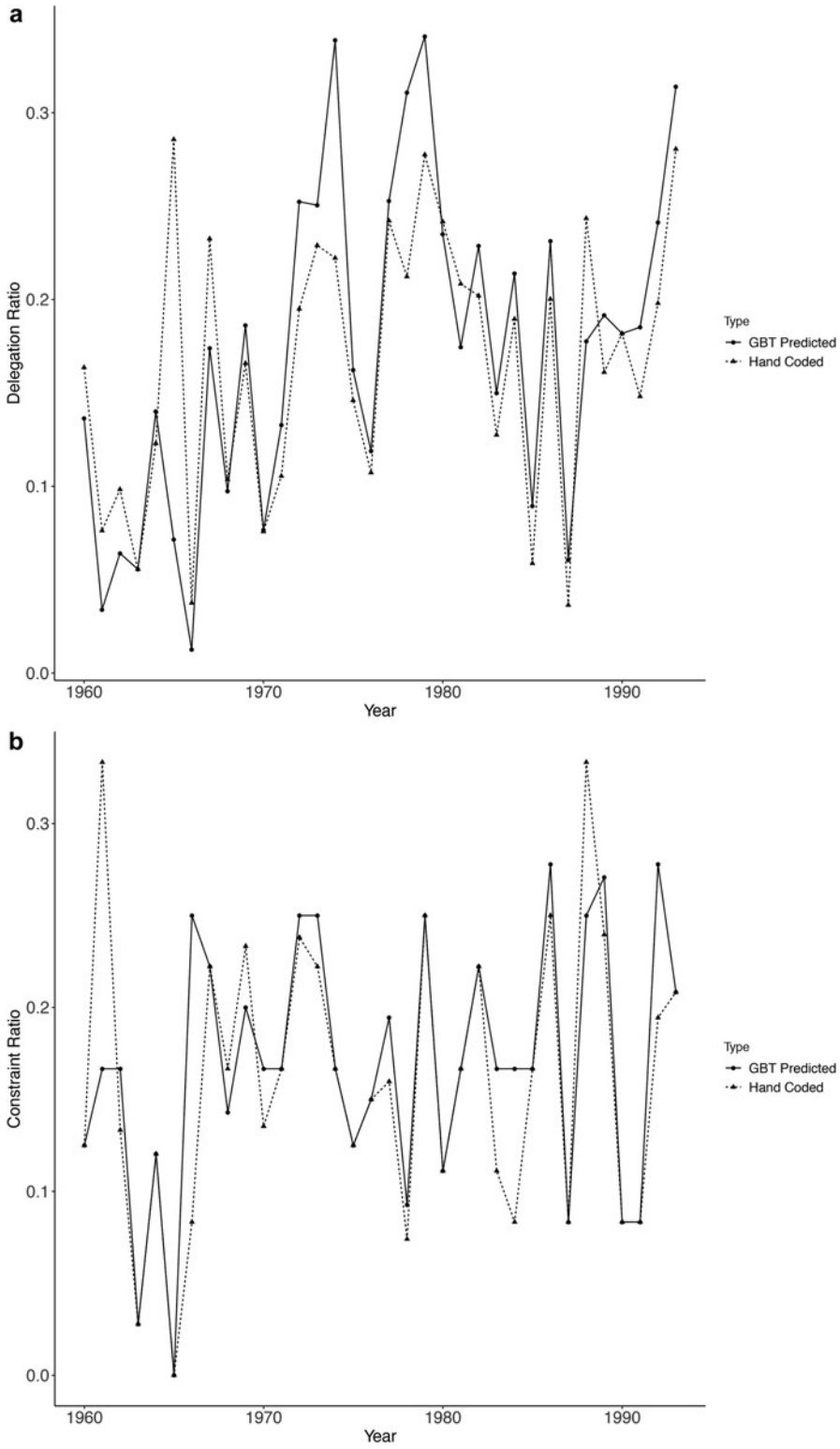
Discretion indices, delegation and constraint ratios for each law were averaged over each year to assess patterns over time. Figure 6a shows clear patterns of declines in member state discretion as the EU expanded. Of note are two periods in which large drops in discretion to national administrations occur around important events. The first was during enlargement in 1973, which added the UK, Ireland, and Denmark to the bloc. The second followed the introduction of fixed exchange rates in 1999 as the first stage in introducing the Euro.

Discontinuity Tests for Out-Of-Sample Consistency

A challenge common to many supervised ML problems is the issue of the validity of predictions outside the training sample. As noted, the hand-coded data available for training come from a sample of 158 "major" laws over 25 fewer years than our dataset covers. The principal concern in our case is that our measures of authority, constraint and discretion may be influenced by structural shifts in the language of EU lawmaking. To address this issue, we introduce a discontinuity test using conditional probabilities estimated by each of the GBT

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

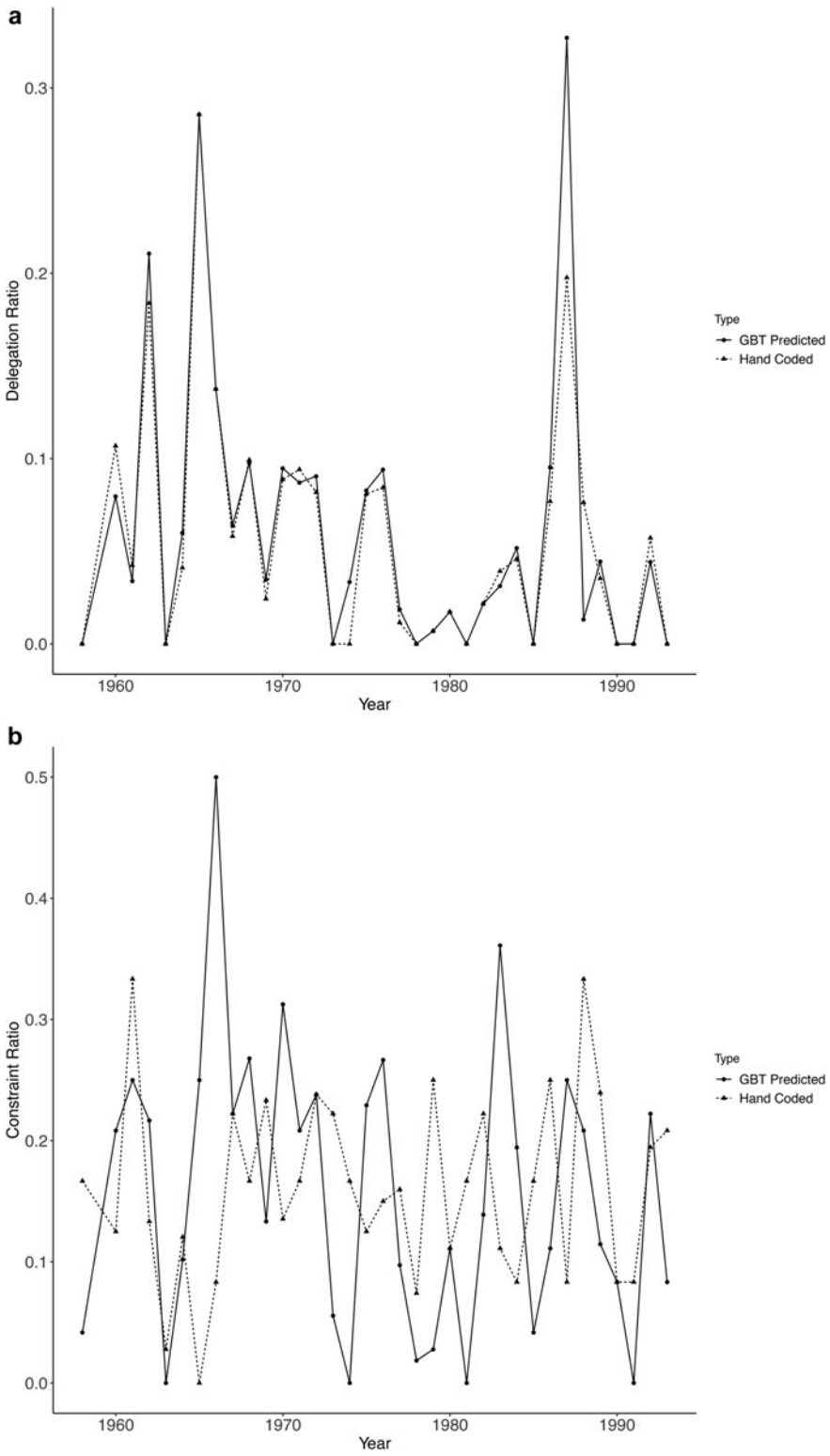
FIGURE 3. GBT Predicted versus Hand-Coded (a) Delegation and (b) Constraint Ratios for National Administrations Aggregated Annually, 1958–92



classifiers during the training process. Classification decisions at the provision level were ultimately made on the basis of whether the estimated probability of identifying authority ($\mathbb{P}[D_j|X_j]$) or constraint ($\mathbb{P}[C_{kj}|X_j]$) was above a 50% threshold based upon the words or terms in each provision X_j. All provisions *above* this

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

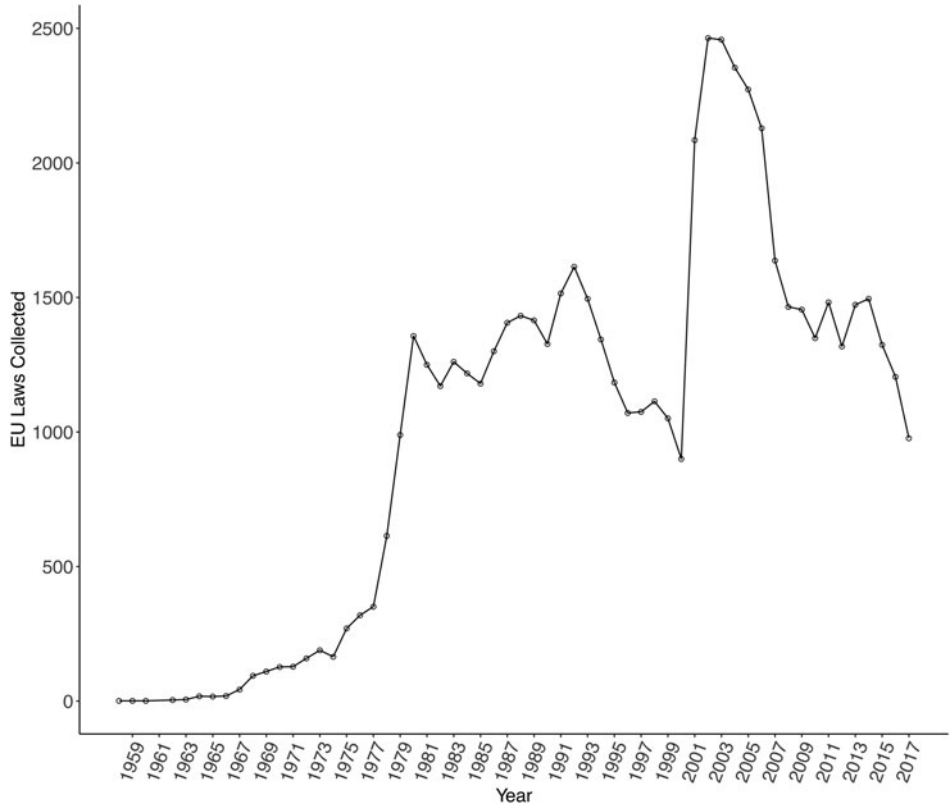
FIGURE 4. GBT Predicted versus Hand-Coded (a) Delegation and (b) Constraint Ratios for the EC Aggregated Annually, 1958–92



threshold were classified as either delegating authority or imposing one or more of the constraint types. For these provisions, then, the predicted probabilities $\mathbb{P}[D_j = 1|X_j]$ and $\mathbb{P}[C_{kj} = 1|X_j]$ provide information regarding about how *certain* or *confident* the classifier was for each of the provisions classified.[7]

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

FIGURE 5. Annual Number of EU Regulations and Directives Analyzed Between 1958–2017



Major changes in the structure of the language of delegation would be reflected by shifts in classifier certainty over time as measured by these conditional probabilities. We are particularly concerned about shifts in language in the out-of-sample years (1994–2017), which can threaten the validity of our estimates. Our discontinuity test leverages average differences in these certainty estimates with the last in-sample year, 1993, as a cutpoint. The classifier is said to "pass" our test if the average delegation and constraint certainty estimates are approximately equal around a window of the cutpoint, c. For example, the test in regard to identifying authority is as follows:

$$\mathbb{E}\left[\mathbb{P}(D_j = 1|X_j)\right]_{c-\epsilon} \approx \mathbb{E}\left[\mathbb{P}(D_j = 1|X_j)\right]_{c+\epsilon}.$$

Passing this test implies that no significant linguistic changes occurred in the out of sample years. Because we estimate a total of nine classifiers (one regarding authority and eight regarding constraints), we conduct a total of nine discontinuity tests. Our classifiers pass all nine tests.[7] To visualize the results, Figure 7 plots the discontinuity test for the national administration authority classifier.

We next use our estimated discretion index $\left(\hat{\delta}\right)$ in a replication exercise to check its face validity.

REPLICATING SUBSTANTIVE RESULTS OUT-OF-SAMPLE

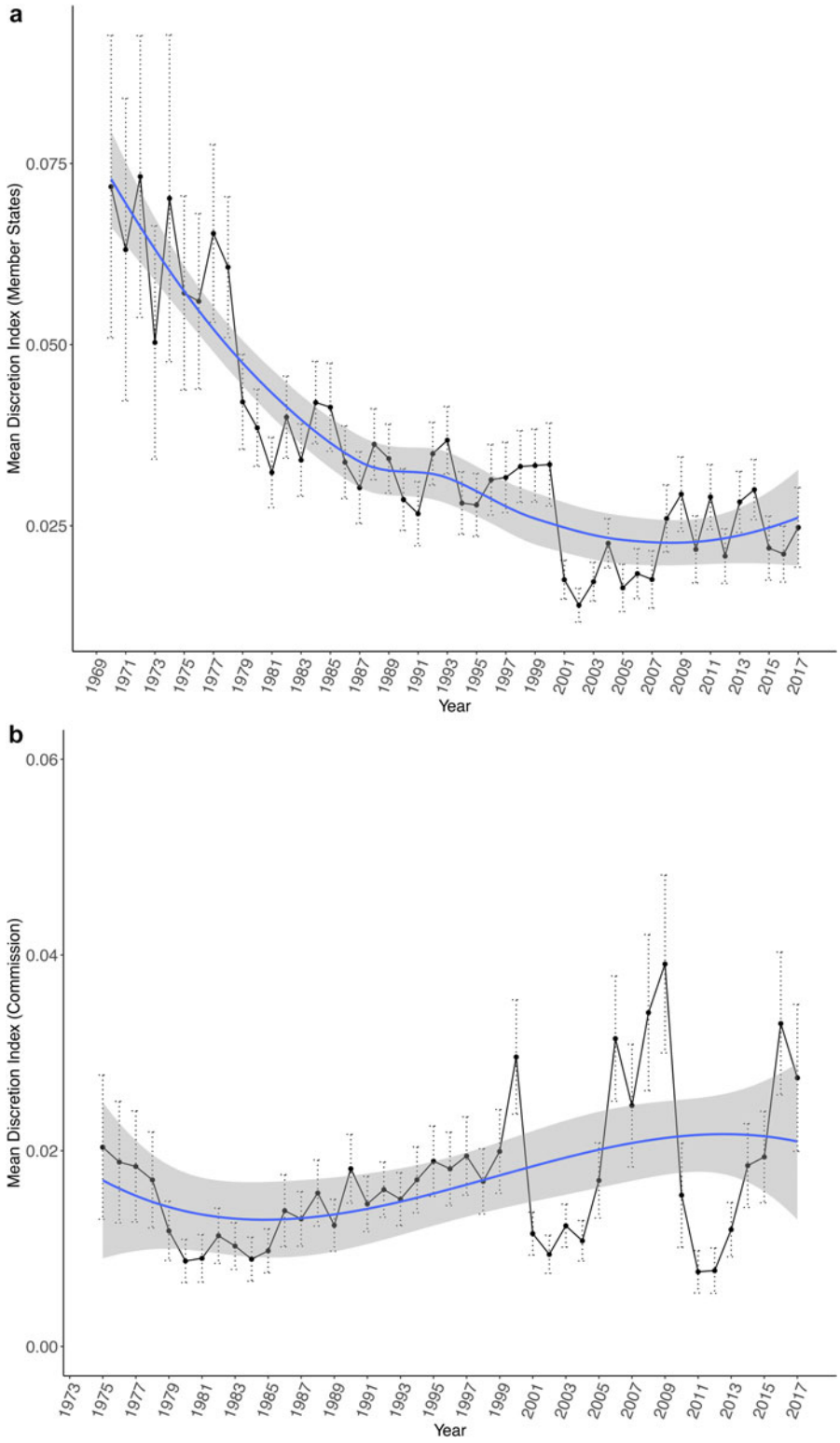
Franchino (2007) tests a hypothesis regarding the relationship between the decision rule for a law—specifically, qualified majority (QMV) versus unanimity rules—and discretion to national administrations. Proposition 2.2 from his formal theory states that the discretion of national administrations under unanimity voting is greater than or equal to that prevailing under QMV (54–55). He tests this claim in several models, but our focus is on the "summary" tobit specifications for national administrations in the right-hand column of Table 5.6 that control for other hypothesized variables (172). The dependent variable is the discretion index $\left(\hat{\delta}\right)$, while the independent variable of interest is *Decision rule*, which takes a value of 1 for QMV and zero for unanimity rules (122), and is expected to have a negative effect in Table 5.6.[6] We focus on this variable because of its measurement quality: QMV versus unanimity rule is the precise distinction that Franchino's formal theory incorporates.

As a final validity check, we seek to replicate the foregoing regression results while treating 90% of Franchino's original observations as out-of-sample. This allows us to show that even with a very limited

[6] Data drawn from the Chapter five replication file at http://users2.unimi.it/fabiofranchino/.

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

FIGURE 6. Average Yearly Discretion Indices $(\hat{\delta})$ for (a) National Administrations and (b) the EC Within all Collected EU Legislation Between 1970–2017



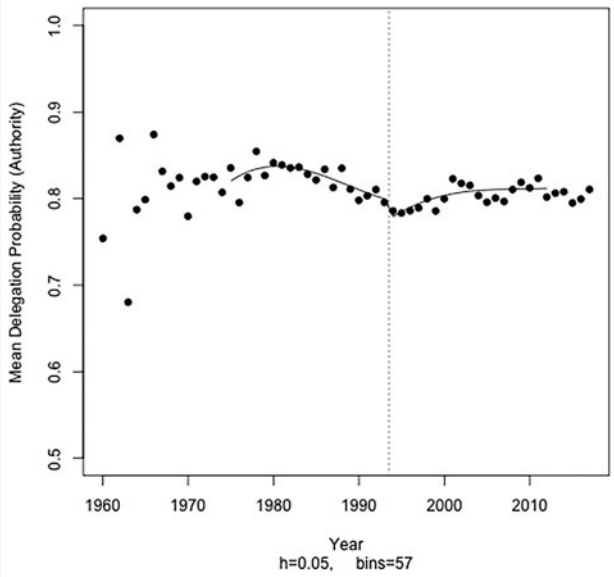
random sample of hand-coded provisions, we can achieve substantive effects that are similar in sign, significance, and magnitude to Franchino's original inferences.

To do this exercise, we, first, trained a model on a random sample of only 10% of the hand-coded provisions. Second, we made predictions on the remaining 90% of those provisions. Third, we used those

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

predictions to construct law-level predicted delegation and constraint ratios, which were combined into a predicted discretion index in the manner described above. Fourth, we used the law-level predicted discretion index ($N = 158$) in the tobit regression specification that Franchino (2007) presents in Table 5.6.

FIGURE 7. Discontinuity Test for National Administration Authority Classifier



Notes: Bins are mean values of $\mathbb{P}(D_j = 1|X_j)$. Local linear regressions estimated on each side of the cutpoint using the Imbens–Kalyanaraman bandwidth to identify discontinuity in $\mathbb{E}[\mathbb{P}(\boldsymbol{D_j} = 1|\boldsymbol{X_j})]$

Table 3 displays Franchino's original results on the left and those $\hat{\delta}$ estimated from a 10% training sample. The reduction in the training sample by 90% introduces a measurement error into the predicted discretion index. Strikingly, we recover estimates similar in sign and magnitude to Franchino's original inferences. In Table 3, the coefficient capturing the effect of a QMV rule on discretion to national administrations remains negative. The effect is only 8% smaller with a *t*-statistic that is just 2.8% larger. Because of the relatively small sample size and measurement error inherent in the policy complexity and policy position covariates, we note that the statistical significance of two CM policy conflict measures, *Integration range* and *Left–right range*, change. This may well be due to chance, but, nonetheless, the coefficient sizes and signs remain comparable.

CONCLUSION

Delegation is fundamental to governance, but presents conceptual challenges making empirical investigations time-consuming and inconsistent, and demanding strong substantive expertise. Machine learning provides an opportunity to harness the

TABLE 3. Replication of Franchino (2007, 172) Table 5.6, Rightmost Column

	Original results	90% out-of-sample $\hat{\delta}$
Programme-committee	7.877***	3.656**
	(3.97)	(2.07)
Decision rule (QMV = 1)	**−4.242****	**−3.885****
	(−2.12)	**(−2.18)**
Integration range	0.768**	0.478
	(2.14)	(1.49)
Standardized policy range	−18.047**	−20.762***
	(−2.24)	(−2.87)
Left–right range	2.284	3.896***
	(1.50)	(2.87)
Constant	7.862*	7.026*
	(1.83)	(1.84)
Observations	158	158
Pseudo *R*-squared	0.03	0.03

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.
Notes: Variable labels follow Franchino. Dependent variable is the national administration discretion index ($\hat{\delta}$) and *Decision rule* indicates QMV (=1). Additional variables described in Franchino (2007) at the specified pages, capturing conflict over EU integration, *Integration range* on the CM (130–31); policy positions, *Standardized policy range* on the CM (130–31); and left-right ideology, *Left-right range*(130–31) on the CM; and policy complexity, the existence of action programs and committee requirements *Programme–committee* (145–46, 166).

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press

expertise of delegation scholars in a consistent and efficient way. Our empirical framework captures delegation in the texts of EU legislation and constructs estimates that can be automatically updated as new legislation becomes available. It is also extensible to a variety of contexts.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0003055419000522.

Replication materials can be found on Dataverse at: https://doi.org/10.7910/DVN/FF3DQM.

REFERENCES

Ash, Elliott, Massimo Morelli, and Matia Vannoni. 2018. Civil Service Reform in US States: Structural Causes and Impacts on Delegation. Technical Report, Mimeo.

Benoit, Kenneth, Drew Conway, Benjamin E. Lauderdale, Michael Laver, and Slava Mikhaylov. 2016. "Crowd-sourced Text Analysis: Reproducible and Agile Production of Political Data." *American Political Science Review* 110 (2): 278–95.

Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Mller, and Akitaka Matsuo. 2018. "Quanteda: An R Package for the Quantitative Analysis of Textual Data." *Journal of Open Source Software* 3 (30): 774.

Clinton, Joshua D., Anthony M. Bertelli, Christian R. Grose, David E. Lewis, and David C. Nixon. 2012. "Separated Powers in the United States: The Ideology of Agencies, Presidents, and Congress." *American Journal of Political Science* 56 (2): 341–54.

Clouser McCann, Pamela J. 2016. *The Federal Design Dilemma*. New York: Cambridge University Press.

Epstein, David, and Sharyn O'Halloran. 1999. *Delegating powers: A Transaction Cost Politics Approach to Policy Making under Separate powers*. New York: Cambridge University Press.

Franchino, Fabio. 2004. "Delegating Powers in the European Community." *British Journal of Political Science* 34 (2): 269–93.

Franchino, Fabio. 2007. *The Powers of the Union: Delegation in the European Union*. Cambridge: Cambridge University Press.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21 (3): 267–97.

Huber, John D., and Charles R. Shipan. 2002. *Deliberate Discretion? The Institutional Foundations of Bureaucratic Autonomy*. New York: Cambridge University Press.

Huber, John D., Charles R. Shipan, and Madelaine Pfahler. 2001. "Legislatures and Statutory Control of Bureaucracy." *American Journal of Political Science* 45 (2): 330–45.

Lakkaraju, Himabindu, Stephen H. Bach, and Jure Leskovec. 2016. "Interpretable Decision Sets: A Joint Framework for Description and Prediction." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: ACM, 1675–84.

Lavertu, Stéphane. 2012. "Issue-Specific Political Uncertainty and Policy Insulation in US Federal Agencies." *Journal of Law, Economics, and Organization* 29 (1): 145–77.

Mayhew, David R. 1991. *Divided We Govern*. New Haven, CT: Yale University Press.

O'Halloran, Sharyn, Sameer Maskey, Geraldine McAllister, David K. Park, and Kaiping Chen. 2015. "Big Data and the Regulation of Financial Markets." In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*: ACM, 1118–24.

O'Halloran, Sharyn, Karen Chen, Rudra M. Guha Biswas, Hoon Kim, Pu Liu, YouFei Zhang, and YunPeng Zhou. 2018. "Delegating Regulation: European Union and Financial Markets." In *Annales des Mines-Realites industrielles*. Vol. 4: FFE, 91–111.

Vakilifathi, Mona. 2017. "Constraining Bureaucrats Today Knowing You'll Be Gone Tomorrow: The Effect of Legislative Term Limits on Statutory Discretion." *Policy Studies Journal*. Published online first 25 August 2017. doi: 10.1111/psj.12215.

https://doi.org/10.1017/S0003055419000522 Published online by Cambridge University Press