

Supplementary Appendix for “Keyword Assisted Topic Models” *American Journal of Political Science*

Shusei Eshima*

Kosuke Imai†

Tomoya Sasaki‡

August 8, 2022

A	Sampling Algorithm and Model Interpretation	1
B	Additional Information for the Base keyATM	4
C	Additional Information for the Covariate keyATM	19
D	Additional Information for the Dynamic keyATM	22
E	Additional Information for the Topic Matching of wLDA	27
F	Additional Information for the Validation	27

A Sampling Algorithm and Model Interpretation

A.1 The Covariate keyATM

A.1.1 Sampling Algorithm

The collapsed Gibbs sampler for the covariate keyATM is also identical to that of the base keyATM except for a couple of steps. First, we sample the topic assignment for each word i in document d from the following conditional posterior distribution that incorporates the covariates (instead of Equation (4)),

$$\Pr(z_{di} = k \mid \mathbf{z}^{-di}, \mathbf{w}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{X}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\gamma}) \propto \begin{cases} \frac{\beta_v + n_{kv}^{-di}}{\sum_v \beta_v + n_k^{-di}} \cdot \frac{n_k^{-di} + \gamma_2}{\tilde{n}_k^{-di} + \gamma_1 + n_k^{-di} + \gamma_2} \cdot \left(n_{dk}^{-di} + \exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d) \right) & \text{if } s_{di} = 0, \\ \frac{\tilde{\beta}_v + \tilde{n}_{kv}^{-di}}{\sum_{v \in \mathcal{V}_k} \tilde{\beta}_v + \tilde{n}_k^{-di}} \cdot \frac{\tilde{n}_k^{-di} + \gamma_1}{\tilde{n}_k^{-di} + \gamma_1 + n_k^{-di} + \gamma_2} \cdot \left(n_{dk}^{-di} + \exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d) \right) & \text{if } s_{di} = 1. \end{cases} \quad (\text{S1})$$

In addition, we need to sample $\boldsymbol{\lambda}$ instead of $\boldsymbol{\alpha}$ (see Equation (5)). As before, we use the unbounded slice sampler (Mochihashi, 2020) and sample each parameter λ_{mk} based on the conditional posterior

*Graduate Student, Department of Government, Harvard University. 1737 Cambridge Street, Institute for Quantitative Social Science, Cambridge MA 02138. Email: shuseieshima@g.harvard.edu URL: <https://shusei-e.github.io/>

†Professor, Department of Government and Department of Statistics, Harvard University. 1737 Cambridge Street, Institute for Quantitative Social Science, Cambridge MA 02138. Email: imai@harvard.edu URL: <https://imai.fas.harvard.edu>

‡Graduate Student, Department of Political Science, Massachusetts Institute of Technology. 77 Massachusetts Avenue E53, Cambridge MA 02142. Email: tomoyas@mit.edu

distribution, $p(\lambda_{mk} \mid \boldsymbol{\lambda}_{-[mk]}, \mathbf{X}, \mu, \sigma^2) \propto \prod_{d=1}^D \left[\frac{\Gamma(\sum_{k=1}^K \exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d))}{\Gamma(\exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d))} \frac{\Gamma(n_{dk} + \exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d))}{\Gamma(\sum_{k=1}^K n_{dk} + \exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d))} \right] \exp\left(\frac{-(\lambda_{mk} - \mu)^2}{2\sigma^2}\right)$, where $\boldsymbol{\lambda}_{-[mk]}$ is $\boldsymbol{\lambda}$ without its (m, k) element.

The model has K topics, D documents, and M covariates. The covariate matrix is given by \mathbf{X} whose d th column is denoted by \mathbf{x}_d , which is an $M \times 1$ dimensional covariate vector. In addition, $\boldsymbol{\lambda}$ is an $M \times K$ dimensional matrix for coefficients. In the estimation, we use standardized covariates, which is defined as $\tilde{\mathbf{x}}_m = \frac{\mathbf{x}_m - \bar{\mathbf{x}}_m}{\text{SD}(\mathbf{x}_m)}$. where $\text{SD}(\mathbf{x}_m)$ represents the standard deviation of \mathbf{x}_m . We can write the standardization of \mathbf{X} in the following matrix form, $\tilde{\mathbf{X}} = (\mathbf{I}_D - \frac{1}{D} \mathbf{1}_D \mathbf{1}_D^\top) \mathbf{X} \mathbf{W}$, where $\mathbf{I}_D - \frac{1}{D} \mathbf{1}_D \mathbf{1}_D^\top$ is a $D \times D$ matrix to demean \mathbf{X} , and \mathbf{W} is a $M \times M$ scaling matrix whose diagonal elements are the inverse of m^{th} covariate's standard deviation.

When we use standardized covariates, the model for θ_d becomes $\theta_d \sim \text{Dirichlet}(\exp(\tilde{\boldsymbol{\lambda}}^\top \tilde{\mathbf{x}}_d))$. And yet, we want to compute $\boldsymbol{\lambda}$ for raw covariates \mathbf{x}_d . The linear transformation of \mathbf{x} will be reflected in $\tilde{\boldsymbol{\lambda}}$, because we do not transform the assignment counts of \mathbf{z} and the only difference between the likelihood of $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{\lambda}}$ is whether or not the covariates are standardized, hence $\mathbf{X}\boldsymbol{\lambda} = \tilde{\mathbf{X}}\tilde{\boldsymbol{\lambda}}$. Now, solve it for $\boldsymbol{\lambda}$ to obtain, $\boldsymbol{\lambda} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \tilde{\mathbf{X}} \tilde{\boldsymbol{\lambda}}$. Therefore, keyATM can store $\tilde{\boldsymbol{\lambda}}$ in each iteration and rescale it to yield $\boldsymbol{\lambda}$. Using $\boldsymbol{\lambda}$, we wish to obtain a posterior predictive distribution of θ_d given a new covariate data set $\mathbf{X} = \mathbf{x}_d^*$. Then, the posterior predictive distribution is, $p(\theta_d^* \mid \mathbf{x}_d^*, \mathbf{w}) = \int p(\theta_d^* \mid \mathbf{x}_d^*, \boldsymbol{\lambda}) p(\boldsymbol{\lambda} \mid \mu, \sigma^2, \mathbf{w}) d\boldsymbol{\lambda}$. We can compute the posterior predictive distribution of the mean of θ_d^* given $\mathbf{x}_d^* \int \frac{\exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d^*)}{\sum_{k'=1}^K \exp(\boldsymbol{\lambda}_{k'}^\top \mathbf{x}_d^*)} p(\boldsymbol{\lambda} \mid \mu, \sigma^2, \mathbf{w}) d\boldsymbol{\lambda}$.

A.1.2 Model Interpretation

The derivation of the topic-word distribution is identical to the one used for the base keyATM (see Section 2.3). In addition, the covariate keyATM can characterize the relations between covariates and document-topic distributions, which are governed by the coefficients $\boldsymbol{\lambda}$. Specifically, we simply replace α_k with $\exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d)$ in the Equation (9) to obtain the marginal posterior mean of θ_{dk} ,

$$\mathbb{E}(\theta_{dk} \mid \mathbf{x}_d, \mathbf{w}) = \mathbb{E}\left(\frac{\exp(\boldsymbol{\lambda}_k^\top \mathbf{x}_d) + n_{dk}}{\sum_{k'=1}^K \exp(\boldsymbol{\lambda}_{k'}^\top \mathbf{x}_d) + n_{dk'}} \mid \mathbf{x}_d, \mathbf{w}\right). \quad (\text{S2})$$

We can also obtain the predicted topic proportion θ_{dk} by setting \mathbf{x}_d to specific values and computing the posterior distribution of its mean given the new covariate value.

Finally, although such an extension is possible, the covariate keyATM does not directly model the topic-word distributions. Nevertheless, it is possible to examine how the topic-word distributions change across different values of document-level covariates by simply computing Equation (8) with a set of documents that share the same values of document-level covariates.

A.2 The Dynamic keyATM

A.2.1 Sampling Algorithm

For fitting the dynamic keyATM, we add two steps to the sampling algorithm described in Section 2.2 (i.e., sampling of latent states and transition probabilities) and change how to sample $\boldsymbol{\alpha}$. To sample the latent state membership $\mathbf{h}_{1:T}$, we use a well-known forward-and-backward sampling procedure. Under the current setup, joint density of h from time 1 to T can be factorized as (Chib, 1996), $p(\mathbf{h}_{1:T} \mid \mathbf{z}, \boldsymbol{\alpha}, \mathbf{P}) = p(h_T \mid \mathbf{z}_T, \boldsymbol{\alpha}, \mathbf{P}) \times \dots \times p(h_t \mid \mathbf{h}_{t+1:T}, \mathbf{z}_t, \boldsymbol{\alpha}, \mathbf{P}) \times \dots \times p(h_1 \mid \mathbf{h}_{2:T}, \mathbf{z}_1, \boldsymbol{\alpha}, \mathbf{P})$. where \mathbf{z}_t represents a vector of the topic assignments for all documents that belong to time period

t . Thus, to sample from the joint distribution, we must sample latent states backwards in time, starting from h_T . Since we know $h_T = R$ and $h_1 = 1$, we sample each of the remaining states from the following conditional distribution,

$$p(h_t \mid \mathbf{h}_{t+1:T}, \mathbf{z}_t, \boldsymbol{\alpha}, \mathbf{P}) \propto p(h_t \mid \mathbf{z}_t, \boldsymbol{\alpha}, \mathbf{P})p(h_{t+1} \mid h_t, \mathbf{P}), \quad (\text{S3})$$

where the second term is given in the transition matrix \mathbf{P} . Each term in Equation (S3) is obtained using the following recursion formula (Chib, 1998, p.227), moving forward in time, $\Pr(h_t = r \mid \mathbf{z}_t, \boldsymbol{\alpha}, \mathbf{P}) = \frac{\Pr(h_t=r \mid \mathbf{z}_{t-1}, \boldsymbol{\alpha}, \mathbf{P})p(\mathbf{z}_t \mid \boldsymbol{\alpha}_r)}{\sum_{l=r-1}^r \Pr(h_t=l \mid \mathbf{z}_{t-1}, \boldsymbol{\alpha}, \mathbf{P})p(\mathbf{z}_t \mid \boldsymbol{\alpha}_l)}$, where $\Pr(h_t = r \mid \mathbf{z}_{t-1}, \boldsymbol{\alpha}_r, \mathbf{P}) = \sum_{l=r-1}^r p_{lr} \cdot \Pr(h_{t-1} = l \mid \mathbf{z}_{t-1}, \boldsymbol{\alpha}, \mathbf{P})$, which only depends on the information at time $t-1$. At last, to compute $p(\mathbf{z}_t \mid \boldsymbol{\alpha}_r)$, we integrate out $\boldsymbol{\theta}_d$ with respect to a state-specific prior $\boldsymbol{\alpha}_r$, yielding a Dirichlet-Multinomial distribution, $p(\mathbf{z}_d \mid \boldsymbol{\alpha}_{h_{t[d]}}, h_{t[d]}) = \int p(\mathbf{z}_d \mid \boldsymbol{\theta}_d)p(\boldsymbol{\theta}_d \mid \boldsymbol{\alpha}_{h_{t[d]}}, h_{t[d]})d\boldsymbol{\theta}_d$.

To sample the transition matrix \mathbf{P} , let n_{rr} be the number of transitions from state r to state r in the all sequences of latent state memberships \mathbf{h} . Then, we have, $p_{rr} \mid \mathbf{h} \sim \text{Beta}(1 + n_{rr}, 1 + n_{r,r+1})$, for $r = 1, \dots, R$, where $n_{r,r+1} = 1$, because the state always moves forward.

Finally, we sample $\boldsymbol{\alpha}_r$ from the following conditional distribution of $\boldsymbol{\alpha}_r$ using the slice sampling in Mochihashi (2020), $p(\alpha_{rk} \mid \alpha_{-[rk]}, \mathbf{z}, \mathbf{h}, \boldsymbol{\eta}) \propto \alpha_{rk}^{\eta_1-1} \exp(-\eta_2 \alpha_{rk}) \prod_{d=1}^{N_d} \left[\frac{\Gamma(\sum_{k=1}^K \alpha_{rk}) \Gamma(n_{dk} + \alpha_{rk})}{\Gamma(\alpha_{rk}) \Gamma(\sum_{k=1}^K n_{dk} + \alpha_{rk})} \right]^{\mathbb{1}(h_{t[d]}=r)}$ for $k = 1, 2, \dots, \tilde{K}$. For $k = \tilde{K} + 1, \dots, K$, the conditional distribution is identical except that $\tilde{\eta}_1$ and $\tilde{\eta}_2$ are replaced with η_1 and η_2 .

A.2.2 Model Interpretation

The topic-word distribution can be computed as before using Equation (8). To understand the dynamic trend of topic proportions, we first compute the posterior mean of topic proportion for each document d as,

$$\mathbb{E}(\theta_{dk} \mid \mathbf{w}) = \mathbb{E} \left[\frac{\alpha_{h_{t[d]},k} + n_{dk}}{\sum_{k=1}^K \alpha_{h_{t[d]},k} + n_{dk}} \mid \mathbf{w} \right], \quad (\text{S4})$$

for topic k . Then, to examine the resulting time trends, we simply compute the sample average of this quantity across all documents that belong to each time point t ,

$$\frac{1}{N_t} \sum_{d \in \mathcal{D}_t} \mathbb{E}(\theta_{dk} \mid \mathbf{w}) \quad (\text{S5})$$

where \mathcal{D}_t represents the set of documents for time t and $N_t = |\mathcal{D}_t|$ is the number of documents for time t . We can then plot this quantity to visualize the time trend of prevalence for each topic.

A.3 Keywords as Informative Prior

The incorporation of keywords in keyATM imposes an informative prior on the structure of topic-word distribution. Here, we directly connect our approach to the existing methods. Lu et al. (2011) add a pseudo counts C_v to a keyword v via,

$$\phi_k \sim \text{Dirichlet}(\{\beta + C_v\}_{v \in V})$$

where $C_v > 0$ if v is a keyword and $C_v = 0$ otherwise. Then, the mean of the topic-word distribution for a term v in a topic k is given by $\mathbb{E}[\phi_{kv}] = (\beta + C_v) / (\beta V + \sum_{v'} C_{v'})$. Since $C_v = 0$ if v is not a keyword, a keyword of a topic is more likely to appear in the topic compared to other topics on

average. In contrast, the prior distribution used in keyATM implies that $\mathbb{E}[\phi_{kv}^{\text{keyATM}}] > \mathbb{E}[\phi_{kv}^{\text{LDA}}]$ if a v is a keyword and $\mathbb{E}[\phi_{kv}^{\text{keyATM}}] < \mathbb{E}[\phi_{kv}^{\text{LDA}}]$ otherwise where $\phi_{kv}^{\text{keyATM}}$ and ϕ_{kv}^{LDA} represent the topic-word distribution under keyATM and LDA, respectively. Therefore, in terms of the expectation of prior probabilities, keyATM is similar to the model of Lu et al. (2011). In addition, Fan, Doshi-Velez and Miratrix (2019, p.215) take a similar approach by multiplying the prior β_v with a scaling factor c if v is a keyword.

The important difference between the keyATM and these two existing models is that we do not require researchers to specify the values of the hyper-parameters, which directly determine the importance of keywords. Instead, we place a restriction on the structure of the prior distribution by setting no-keyword frequencies exactly to zero in the prior distribution of $\tilde{\phi}_k$ so that the data can inform the importance of keywords.

B Additional Information for the Base keyATM

B.1 Details of Preprocessing

Documents

We preprocess the raw texts by first removing stop words via the R package `quanteda` (Benoit et al., 2018), then pruning words that appear less than 11 times in the corpus, and lemmatizing the remaining words via the Python library NLTK (Bird, Klein and Loper, 2009). We also use lemmatization. Removing stop words and low frequent terms, and lemmatizing terms help reduce the dimensionality of data (Manning, Raghavan and Schütze, 2008). We use the same preprocessing steps for all models.

CAP Codebook

First, we remove stopwords and lemmatize the remaining terms in the same way as done for the bill texts. We also remove the words and phrases that have little to do with the substantive meaning of each topic. For example, a topic description always begins with the phrase “Description: Includes issues . . .”, and hence we remove this phrase. Similarly, we exclude the phrases, “. . . related generally to . . .” and “. . . generally related to . . .,” which appear in the description of 13 different topics. Third, we use the same keywords for multiple topics only if their inclusion can be substantively justified. For example, the term “technology,” which appears in the description of several topics, is only used as a keyword for the *Technology* topic. Lastly, we limit the number of keywords to 25 per topic. We remove terms based on the proportion of keywords among all terms in the corpus if the topic contains more than 25 keywords.

B.2 The Full List of Keywords

Label	Keywords
Government operations	administrative advertising appointment attack auditing branch campaign capital census city coin collection currency mail medal mint nomination post postal registration statistic terrorist victim voter
Public lands	fire flood forest grazing historic indigenous land livestock natural parks recreation resource site staff territorial
Defense	armed base capability civilian compliance contractor coordination damage equipment foreign homeland installation intelligence material military nuclear operation personnel procurement reserve security services
Domestic commerce	account accounting bankruptcy business card commerce commercial commodity consumer cost credit disaster finance financial fraud industry insurance investment management mortgage patent promote property relief security

Law and Crime	abuse border code combat court crime criminal custom cyber drug enforcement family fine judiciary justice juvenile legal penalty police prison release representation sexual terrorism violence
Health	abuse alcohol care clinical cost cover coverage disease drug health insurance insurer liability license medical mental pay payment prescription prevention provider rehabilitation supply tobacco treatment
International affairs	aid assessment associate citizen combat committee convention country cross develop directly foreign human international monetary ocean region regional sea target terrorism treaty union world
Transportation	air airport aviation channel construction deployment freight highway infrastructure inland maintenance maritime mass pilot rail railroad ship traffic transportation travel waterway
Macroeconomics	bank budget budgeting central cost deficit growth industrial inflation interest macroeconomic manufacturing monetary price revitalization tax treasury
Environment	alternative asbestos chemical climate conservation disposal drinking endanger environment environmental hazardous laboratory performance pollution protection resource solid specie supply toxic waste wastewater water wildlife
Education	adult area bilingual college education educational elementary excellence handicapped improve language literacy loan mentally need outcome physically primary school schools secondary skill student university vocational
Energy	alternative biofuel clean coal conservation drilling electrical electricity energy gas gasification gasoline geothermal hydrogen hydropower natural nuclear oil power production renewable shortage spill utility vehicle
Technology	broadcast communication computer cooperation encourage exploration forecast form geological internet publishing radio research satellite science space speed survey technology telecommunication telecommunications telephone television transfer weather
Labor	bargaining benefit compensation debt employee employer employment fair injury insurance job labor minimum pension protection retirement standard training unemployment union wage work worker workforce youth
Foreign trade	agreement balance barrier competitiveness dispute exchange export foreign import international negotiation private productivity subsidy tariff trade treaty
Civil rights	abortion age civil contract discrimination ethnic expression franchise freedom gender group information mandatory minority participation party privacy racial religious right rights sex sexual speech voting
Social welfare	aid alleviate assess assistance association care charity child disability disable elderly family income leave parental physical poverty social volunteer welfare youth
Agriculture	agricultural agriculture animal crop farmer fish fishery food inspection labeling market pest pesticide product rancher seafood subsidy
Housing	affordability community economic family handicap homeless homelessness housing income neighborhood rural urban veteran
Immigration	refugee citizenship immigration
Culture	culture cultural

Table S1: **Keywords for the base keyATM**: Keywords used in the base keyATM application.

B.3 Top Words

Table S2 presents the top ten words with the highest estimated probability defined in Equation (7). The results show that keyATM shows better performance than wLDA in most topics.

Government operations		Public lands		Defense	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
expense	congress	land	land	military	military
appropriation	house	water	management	force	force
remain	senate	area*	area	member	system
authorize	office	management*	forest	air*	support
necessary	committee	river	indian	code*	authorization
transfer*	commission	resource	interior	authority	security
expend	representative	forest	resource	authorization	operation
exceed	congressional	authorize	park	reserve	army
office	strike	cost*	conservation	armed	committee
activity	bill	stat	within	army	air
Domestic commerce		Law & crime		Health	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
financial	financial	intelligence*	security	health	health
institution	institution	attorney	information	care	care
bank*	bank	crime	intelligence	drug	individual
insurance	company	court	homeland	payment	drug
company	insurance	enforcement	committee	medical	payment
corporation	corporation	criminal	director	individual	medical
board	board	code	system	coverage	describe
security	security	offense	foreign	medicare	respect
credit	credit	person	government	respect	social
commission	commission	justice	office	describe	part
International affairs		Transportation		Macroeconomics	

keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
assistance*	assistance	transportation	transportation	apply	transfer
foreign	foreign	highway	highway	tax	appropriation
country	country	safety	safety	amendment	emergency
international	international	carrier	vehicle	end	operation
government	president	air	carrier	taxable	military
president	government	code*	motor	strike	construction
development	committee	system	system	relate	procurement
committee	development	vehicle*	strike	income*	remain
export*	export	airport	rail	respect	maintenance
organization	organization	motor	code	case	budget
Environment		Education		Energy	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
committee*	water	education	school	energy	energy
submit	river	school	education	fuel	fuel
review	cost	educational	educational	gas	gas
later	stat	student	student	facility	change
administrator	authorize	local	child	vehicle	facility
require	carry	institution	local	oil	new
requirement	non	grant	grant	electric	vehicle
develop*	development	part	part	standard*	electric
information*	resource	high	activity	administrator	production
management*	study	eligible	eligible	power	administrator
Technology		Labor		Foreign trade	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
research	research	employee	apply	product*	air
technology	technology	benefit	tax	trade	vessel
business*	development	individual	amendment	change	airport
information*	establish	rate	end	agreement	transportation
director	committee	compensation	taxable	good	aviation
system	activity	period	respect	tobacco*	administrator
small	administrator	code*	period	head	aircraft
science	system	payment*	individual	article	carrier
development	information	determine	case	free	administration
center	carry	agreement*	relate	chapter	coast
Civil rights		Social welfare		Agriculture	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
person	person	child	expense	food	food
action	action	assistance	appropriation	agricultural	agricultural
information	regulation	individual	authorize	loan*	loan
order	require	grant	remain	agriculture	agriculture
court*	information	family	necessary	farm	payment
regulation	court	disability	expend	producer	farm
commission	order	strike	office	payment*	producer
require	rule	indian	exceed	rural*	crop
rule	commission	receive	transfer	crop	rural
provision	review	support	activity	commodity*	commodity
Housing		Immigration		Culture	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
housing	housing	security*	alien	congress	member
loan*	assistance	alien	attorney	house	strike
assistance*	loan	immigration	child	senate	code
development	development	homeland*	crime	office	force
community	family	border*	immigration	committee*	pay
mortgage*	mortgage	status	grant	representative	military
family	community	nationality	enforcement	member	officer
income	insurance	describe	person	strike	authorize
insurance*	income	individual	court	government	duty
unit	unit	employer*	offense	congressional	reserve

Table S2: **Top words of the base keyATM.** The table shows the top ten words with the highest estimated probability for each topic under each model. For keyATM, the pre-specified keywords for each topic appear in bold letters whereas the asterisks indicate the keywords specified for another topic.

B.4 ROC Curves

Figure S1 shows the ROC curves of all topics. Each line represents the ROC curve from one of the five Markov chains with different starting values for keyATM (blue lines) and wLDA (grey lines). The median AUROC indicates the median value of AUROC among five chains for each model. keyATM performs at least as well as wLDA in all topics.

B.5 Pooled Results

Figure S2 presents the pooled results after combining five Markov chains for keyATM and wLDA, respectively. Each line represents the ROC curve based on pooled five Markov chains with different starting values for keyATM (blue lines) and wLDA (grey lines). The figures show that keyATM substantially outperforms wLDA in terms of the AUROC.

B.6 Quality of Keywords and Performance of the keyATM

The poor quality of keywords for the *Government operations* topic appears to explain the failure of keyATM in terms of both topic interpretability and classification. The left panel of Figure S3 presents the histogram for the proportion of keywords among all words in each of the bills classified to the *Government operations* topic by the CBP human coders. When compared to the average of the other five topics shown in Table 2 (gray bars), the keywords for this topic (blue bars) appear much less frequently in the relevant bills. Furthermore, the right panel shows the number of unique keywords contained in each relevant bill. Again, unique keywords appear in the *Government operations* bills less frequently than the corresponding keywords do in the other five topics. The results suggest that selecting high-quality keywords is critical for the successful application of keyATM.

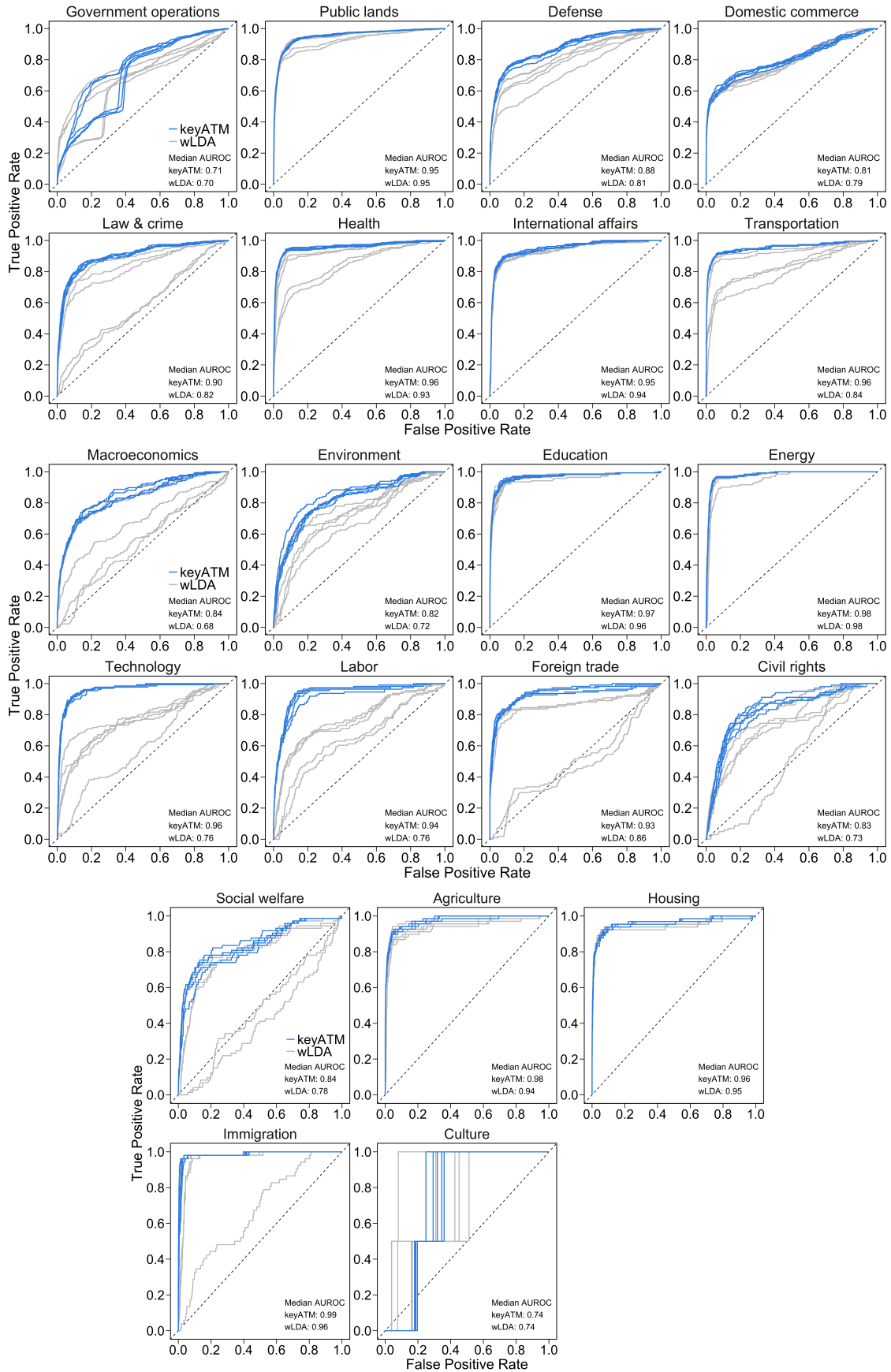


Figure S1: Comparison of the ROC curves between keyATM and wLDA.

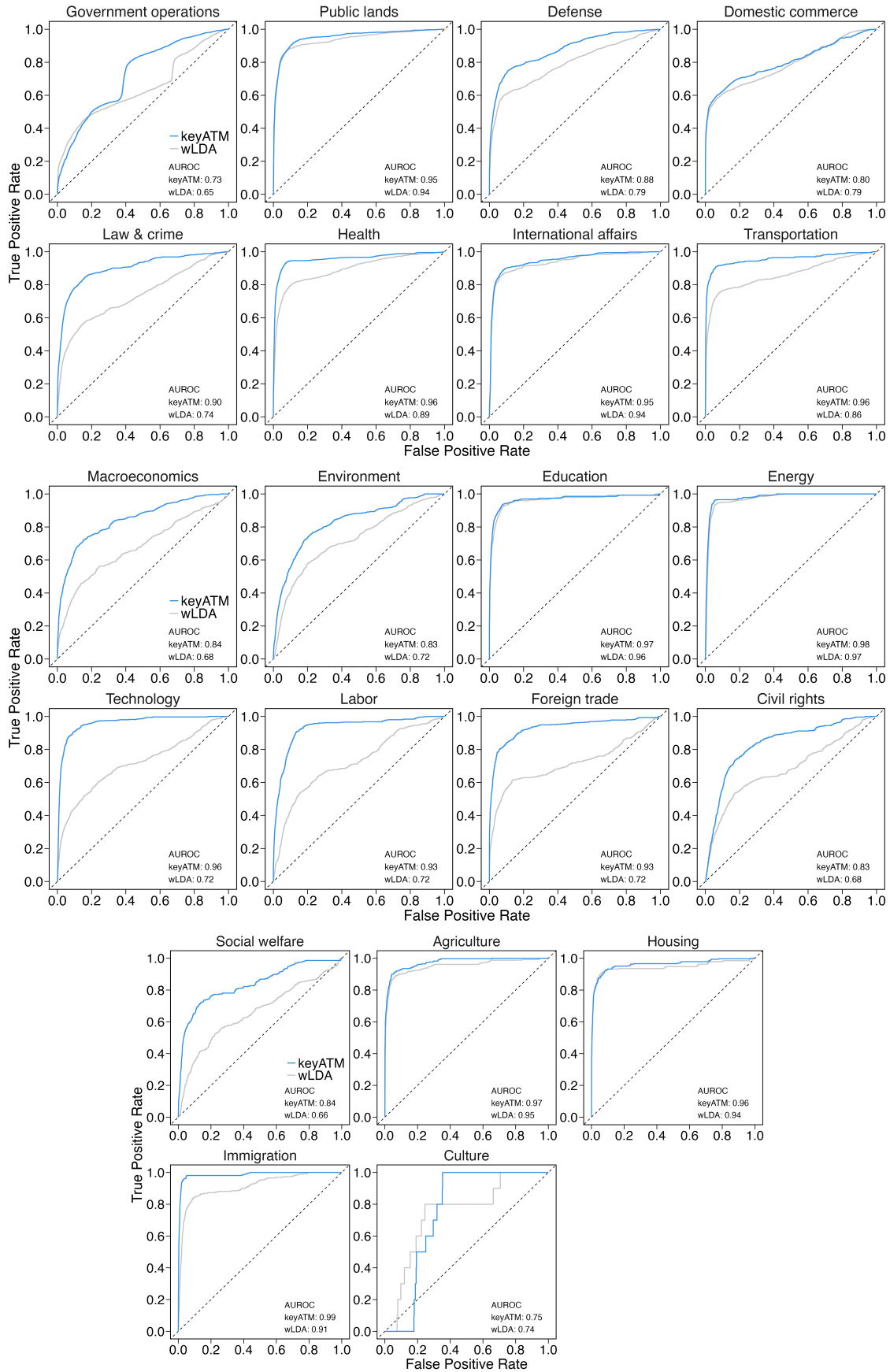


Figure S2: Comparison of the ROC curve between keyATM and wLDA based on the Combined Chains.

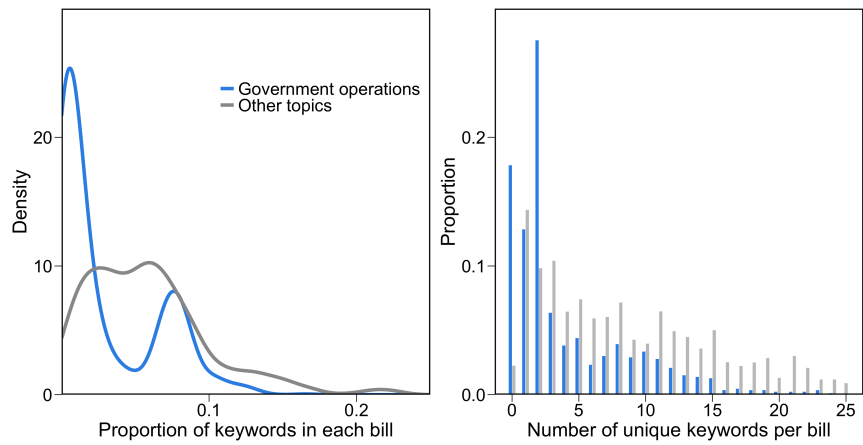


Figure S3: **Poor quality of keywords for the *Government operations* topic.** The left panel presents the histogram for the proportion of keywords in each of the bills classified to the *Government operations* topic (blue bars) by the Congressional Bills Project. Compared to the average of other five topics (gray bars) from Table 2, the keywords appear less frequently in the *Government operations* topic. The right panel presents the histogram for the number of unique keywords that appear in each of the bills classified to the *Government operations* topic. Unique keywords appear less frequently in this topic than the other five topics.

B.7 The Results based on Different Keywords

This section provides the results based a different set of keywords. As is the keyword set used for the main analysis, we derive the keywords of each topic from the description provided by the CAP. However, we do not prune any overlapping keywords and do not remove words or phrases that have little to do with the substance of each topic. Therefore, the number of keywords assigned to each topic is much larger than those used in the main analysis.

B.7.1 The Full List of Different Keywords

Label	Keywords
Government operations	administration administrative advertising appointment appropriation attack auditing branch campaign capital census city civil claim coin collection commemorative compensation constitutional construction contractor corporation currency description efficiency elsewhere employee enforcement finance fraud government holiday individual intergovernmental issue local mail management medal mention mint multiple nomination observation operations oversight pension policy political post postal process procurement property provision reform registration regulation relate relation statistic substantive system tax terrorist victim voter without
Public lands	affair assistance civil control cultural dependency description development fire flood forest grazing historic indigenous issue land livestock management natural parks people policy recreation relate research resource site staff territorial territorie water work
Defense	activity affairs agreement aid air alliance appropriation arm armed assistance base capability civil civilian claim closing collateral compensation compliance construction contract contractor control conversion coordination country courts covert damage dependent description development direct disposal domestic employment environmental equipment espionage evaluation forces foreign fraud generally hazardous homeland industry injure installation intelligence issue land manpower material military modernization non nuclear old operation oversee oversight peacekeeping personnel policy population prisoner procurement proliferation readiness relate research reserve response sale sealift security services settlement stockpile strategic support system terrorism transfer veterans war waste weapon
Domestic commerce	abuse access accounting antitrust appropriation availability bank bankruptcy business card commerce commercial commodity consumer copyright corporate cost credit description development disaster domestic exchange facilitation finance financial fitness flood fraud gambling generally governance government health impact industry institution insurance intellectual investment issue management merger mortgage municipal natural non patent personal policy preparedness promote promotion property record reform regulate regulation relate relief research safety security small sport subsidize system tourism trading
Law and Crime	abuse administration appropriation border child civil code combat control counterfeiting court crime criminal custom cyber description domestic drug effort enforcement exploitation family fine fraud guideline illegal impact international issue jail judiciary justice juvenile kidnapping launder legal mention money organize parental parole penalty police pornography pre prevention prison recidivism reduce relate release representation response security sexual specialize system terrorism traffic violence welfare white
Health	abuse alcohol ambulance appropriation availability benefit broad care change child clinical comprehensive construction cost cover coverage description development device disease drug education effect facility fraud generally government health home ill illegal industry infant insurance insurer issue lab labor liability license long malpractice manpower medical mental multiple pay payment pharmaceutical policy practice prescription prevention promotion provider quality quantity reduce reform regulation rehabilitation relate research school specific supply system tobacco topic training treatment type unfair
International affairs	abroad affairs agreement aid appropriation assessment associate bank citizen code combat committee conservation convention country court crime criminal cross debt description develop development diplomacy directly economic effort embassy europe european exploitation fight finance financial foreign genocide government human humanity implication increase individual institution international issue legal lending mechanism monetary nations ocean olympic organization passport piracy policy political red region regional related relation resource right sea see sovereign specific specifically target terrorism treaty union violation western world
Transportation	air airport appropriation availability aviation channel construction control deployment description development employment freight generally government highway infrastructure initiative inland issue maintenance maritime mass new pilot policy rail railroad regulate regulation relate research safety security ship technology traffic training transportation travel waterway work
Macroeconomics	bank budget budgeting central code control cost debt deficit description domestic effort emergency enforcement growth impact industrial inflation interest issue live macroeconomic manufacturing monetary policy price rate reduce relate revitalization tax taxis treasury unemployment wage
Environment	air airline alternative animal appropriation asbestos change chemical climate conservation contamination control description development disposal domestic drinking endanger energy environment environmental forest government hazardous illicit indoor issue laboratory land noise performance policy pollution product protection recycling regulate regulation relate research resource reuse runoff safety solid specie substance supply technology toxic trade treatment waste wastewater water wildlife
Education	adult appropriation area bilingual child college description development education educational effort elementary excellence finance foreign generally government handicapped high impact improve increase initiative issue language literacy loan math mentally need outcome physically policy primary quality reform regulate regulation relate research rural safety school schools science secondary skill special specific standard student university vocational

Energy	alternative appropriation biofuel clean coal commercial conservation description development disposal drilling efficiency electrical electricity energy gas gasification gasoline generally geothermal government home hydrogen hydropower issue natural nuclear oil policy power price production regulate regulation relate renewable research safety security shortage spill technology trade utility vehicle waste
Technology	agreement broadcast commercial communication computer cooperation description development effort encourage exploitation exploration forecast form generally geological government high industry infrastructure international internet issue mention military newspaper promotion publishing radio regulation relate research resource satellite science security space speed survey technology telecommunication telecommunications telephone television transfer weather
Labor	account adult appropriation bargaining benefit child collective compensation description development disease displace effort employee employer employment fair generally government guest injury insurance issue job labor migrant minimum overtime pension policy protection regulate relate relation retirement retrain safety seasonal standard training unemployment union wage work worker workforce youth
Foreign trade	agreement appropriation balance barrier business competitiveness control corporate description development dispute domestic exchange export foreign generally government impact import industry international investment issue negotiation payment policy private productivity promotion rate regulate regulation relate subsidy tariff tax trade treaty
Civil rights	abortion access activity age anti civil communist contract description discrimination disease ethnic expand expression franchise freedom gender generally government group handicap information issue local mandatory minority orientation participation party policy privacy racial record relate religious retirement right rights sex sexual speech type voting
Social welfare	aid alleviate assess assistance association care charity child credit dependency description direct disability disable domestic elderly family food generally government income issue leave low mental organization parental pension people physical policy poverty relate social tax volunteer welfare youth
Agriculture	agricultural agriculture animal appropriation commercial conservation consumer control crop description development disaster disease effort farmer fish fishery food foreign government impact information inspection insurance issue labeling market pest pesticide policy product promotion rancher regulation relate requirement research safety seafood subsidy trade welfare
Housing	affairs affordability assistance community description development economic effort elderly facility family generally handicap homeless homelessness housing income individual issue low military neighborhood non policy reduce relate research rural subsidy urban veteran
Immigration	citizenship description immigration issue refugee related
Culture	culture cultural

Table S3: **Different Keywords for the base keyATM:** A different set of keywords used in the base keyATM application

B.7.2 Top words

Table S4 shows the top ten words with the highest estimated probabilities for each topic and for keyATM and wLDA. The results demonstrate that the advantages of keyATM presented in the main text (Table 2) still hold with different sets of keywords: First, the *Labor* topic of keyATM lists many related words whereas the same topic for wLDA does not include any related terms. Second, wLDA cannot find meaningful terms for the *Foreign trade* topic and instead creates two topics whose top ten words are related to the *Transportation* topic (under the label of *Foreign trade* and *Transportation*). The top words of the *Foreign trade* topic for keyATM contain words related to trade. Third, keyATM can distinguish *Law & crime* and *Immigration* but wLDA cannot.

Government operations		Public lands		Defense	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
office	congress	land	land	military	military
employee	house	water	management	force	force
commission	senate	management	area	member	system
committee*	office	area*	forest	air	support
administration	committee	resource	indian	authority	authorization
district	commission	river	interior	code*	security
administrative	representative	forest	resource	authorization	operation
pay*	congressional	cost*	park	reserve	army
board	strike	authorize	conservation	armed	committee
management	bill	stat	within	army	air
Domestic commerce		Law & crime		Health	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
financial	financial	court	security	health	health
insurance	institution	attorney	information	care	care
institution	bank	person	intelligence	individual*	individual

bank	company	code	homeland	medical	drug
company	insurance	crime	committee	part	payment
board	corporation	grant	director	payment	medical
commission	board	enforcement	system	coverage	describe
security	security	offense	foreign	social*	respect
corporation*	credit	child	government	describe	social
credit	commission	victim*	office	respect	part
International affairs		Transportation		Macroeconomics	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
assistance*	assistance	transportation	transportation	expense	transfer
foreign	foreign	highway	highway	appropriation*	appropriation
international	country	carrier	safety	remain	emergency
country	international	safety	vehicle	necessary	operation
government	president	system*	carrier	authorize	military
president	government	code*	motor	transfer*	construction
committee	committee	air	system	expend	procurement
development	development	vehicle*	strike	exceed	remain
organization	export	airport	rail	activity*	maintenance
export*	organization	strike	code	pursuant	budget
Environment		Education		Energy	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
administrator	water	education	school	energy	energy
product	river	school	education	fuel	fuel
regulation	cost	educational	educational	gas	gas
requirement*	stat	student	student	facility*	change
review	authorize	child	child	vehicle	facility
drug*	carry	grant	local	oil	new
require	non	local*	grant	electric	vehicle
fee	development	part	part	power	electric
submit	resource	institution*	activity	technology	production
application	study	activity*	eligible	natural	administrator
Technology		Labor		Foreign trade	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
information*	research	benefit	apply	trade	air
system*	technology	payment*	tax	change*	vessel
technology	development	requirement*	amendment	agreement	airport
research	establish	employee	remain	good	transportation
committee*	committee	period	taxable	head	aviation
activity*	activity	rate*	respect	import	administrator
director	administrator	pay*	period	article	aircraft
development	system	determine	individual	chapter	carrier
business*	information	code*	case	free	administration
develop*	carry	provision*	relate	new*	coast
Civil rights		Social welfare		Agriculture	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
information	person	tax	expense	food	food
action	action	apply	appropriation	agricultural	agricultural
person	regulation	taxable	authorize	loan*	loan
require	require	amendment	remain	agriculture	agriculture
order	information	relate	necessary	farm	payment
procedure	court	income	expend	rural*	farm
provision*	order	end	office	crop	producer
individual*	rule	strike	exceed	payment*	crop
authority	commission	individual*	transfer	producer	rural
request	review	respect	activity	commodity*	commodity
Housing		Immigration		Culture	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
housing	housing	security*	alien	congress	member
assistance	assistance	alien	attorney	house	strike
loan*	loan	intelligence*	child	senate	code
development	development	immigration	crime	committee*	force
family	family	homeland*	immigration	strike	pay
community	mortgage	foreign*	grant	office	military
grant	community	information*	enforcement	representative	officer
mortgage*	insurance	border*	person	congressional	authorize
income	income	describe	court	bill	duty
unit	unit	attorney	offense	veteran*	reserve

Table S4: **Top words of the base keyATM with a different sets of keywords.** The table shows the top ten words with the highest estimated probability for each topic under each model. For keyATM, the pre-specified keywords for each topic appear in bold letters whereas the asterisks indicate the keywords specified for another topic.

B.7.3 ROC Curves

Figure S4 shows the ROC curves. Each line represents the ROC curve from one of the five Markov chains with different starting values for keyATM (blue lines) and wLDA (grey lines). The median AUROC indicates the median value of AUROC among five chains for each model. keyATM performs at least as well as wLDA in 19 out of 21 topics.

B.8 Randomly Removed Keywords

In this section, we present the base keyATM results where we randomly remove some keywords from the keyword sets shown in the Section B.2. In particular, we randomly remove 75% or 50% of keywords from the original keyword sets shown in Table S1.

The results are shown in Table S5. In most cases, the worst AUROC values from the base keyATM are still better than those of wLDA. Note that since the *Immigration* topic only contains three keywords under the original setting: “citizenship”, “immigration”, and “refugee”. We randomly select two out of these three keywords and fit the base keyATM for this topic. Table S5 highlights the importance of keyword selection when the number of keywords is extremely small. For example, the *Immigration* topic shows high variability in the AUROC value. The *Immigration* topic performs poorly in models with the keyword sets 1, 3, and 6. All of these sets do not contain “immigration” as a keyword.

Furthermore, we create 50 different keyword sets where we randomly remove 50% of keywords from the keyword sets shown in Table S1. We compute AUROC for each topic and present the results in Figure S5. Again, in most cases, the worst AUROC values from the base keyATM are still better than those of wLDA, except for some topics such as *Immigration*, which has only three keywords. There is a large variation in AUROC for the *Immigration* topic.

We further conduct a systematic investigation about when the *Immigration* topic performs poorly. Figure S6 shows the AUROC based on the previous simulation for the *Immigration* topic with and without the keyword “immigration” in the keyword set. The results show that the AUROC for the *Immigration* topic is substantially lower when the word “immigration” is removed from the keyword set. This result indicates that whether “immigration” is included in the *Immigration* topic determines the quality of topic classification.

B.9 Convergence

We present the Gelman-Rubin statistic for perplexity of the base keyATM with different number of iterations to assess convergence. We drop the first third of iterations as burn-in. The statistic monotonically decreases and is reduced to 1.08 after 300,000 iterations.

Topic	# of original keywords	keyATM								wLDA
		75% of original				50% of original				N/A
		1	2	3	4	5	6	7	8	
Government operations	24	0.70	0.71	0.74	0.74	0.72	0.74	0.75	0.71	0.72
Public lands	17	0.94	0.95	0.95	0.95	0.92	0.94	0.96	0.96	0.91
Defense	23	0.88	0.88	0.87	0.87	0.86	0.87	0.87	0.88	0.73
Domestic commerce	25	0.81	0.82	0.79	0.79	0.80	0.78	0.81	0.79	0.78
Law & crime	25	0.88	0.90	0.86	0.87	0.90	0.86	0.89	0.87	0.57
Health	25	0.96	0.96	0.96	0.96	0.96	0.95	0.90	0.96	0.95
International affairs	24	0.95	0.94	0.95	0.68	0.73	0.75	0.94	0.95	0.93
Transportation	21	0.94	0.96	0.95	0.96	0.96	0.95	0.94	0.95	0.80
Macroeconomics	17	0.83	0.64	0.81	0.81	0.57	0.78	0.62	0.84	0.54
Environment	24	0.89	0.80	0.80	0.93	0.75	0.84	0.88	0.83	0.72
Education	25	0.97	0.97	0.97	0.97	0.96	0.96	0.97	0.97	0.97
Energy	25	0.98	0.98	0.97	0.97	0.98	0.97	0.98	0.98	0.95
Technology	25	0.92	0.96	0.96	0.54	0.96	0.95	0.93	0.89	0.75
Labor	25	0.93	0.85	0.94	0.93	0.90	0.94	0.79	0.83	0.76
Foreign trade	17	0.94	0.92	0.94	0.92	0.89	0.72	0.91	0.94	0.47
Civil rights	25	0.84	0.81	0.84	0.80	0.74	0.74	0.69	0.86	0.73
Social welfare	21	0.81	0.83	0.81	0.85	0.71	0.84	0.81	0.88	0.47
Agriculture	17	0.96	0.97	0.95	0.98	0.98	0.97	0.97	0.96	0.98
Housing	13	0.95	0.94	0.93	0.94	0.96	0.95	0.96	0.95	0.95
Immigration	3	0.51	0.98	0.49	0.99	0.99	0.50	0.99	0.99	0.96
Culture	2	0.62	0.77	0.88	0.26	0.98	0.52	0.78	0.72	0.72

Table S5: **Comparison of AUROC with different keyword sets.** The table presents AUROC values from the base keyATM with 8 different-randomly-selected keyword sets and wLDA. We run five chains for each setting and the results are based on the MCMC draws from one of the five chains that has the median performance in terms of the overall AUROC. In most cases, the worst AUROC values from the base keyATM are still better than those of wLDA.

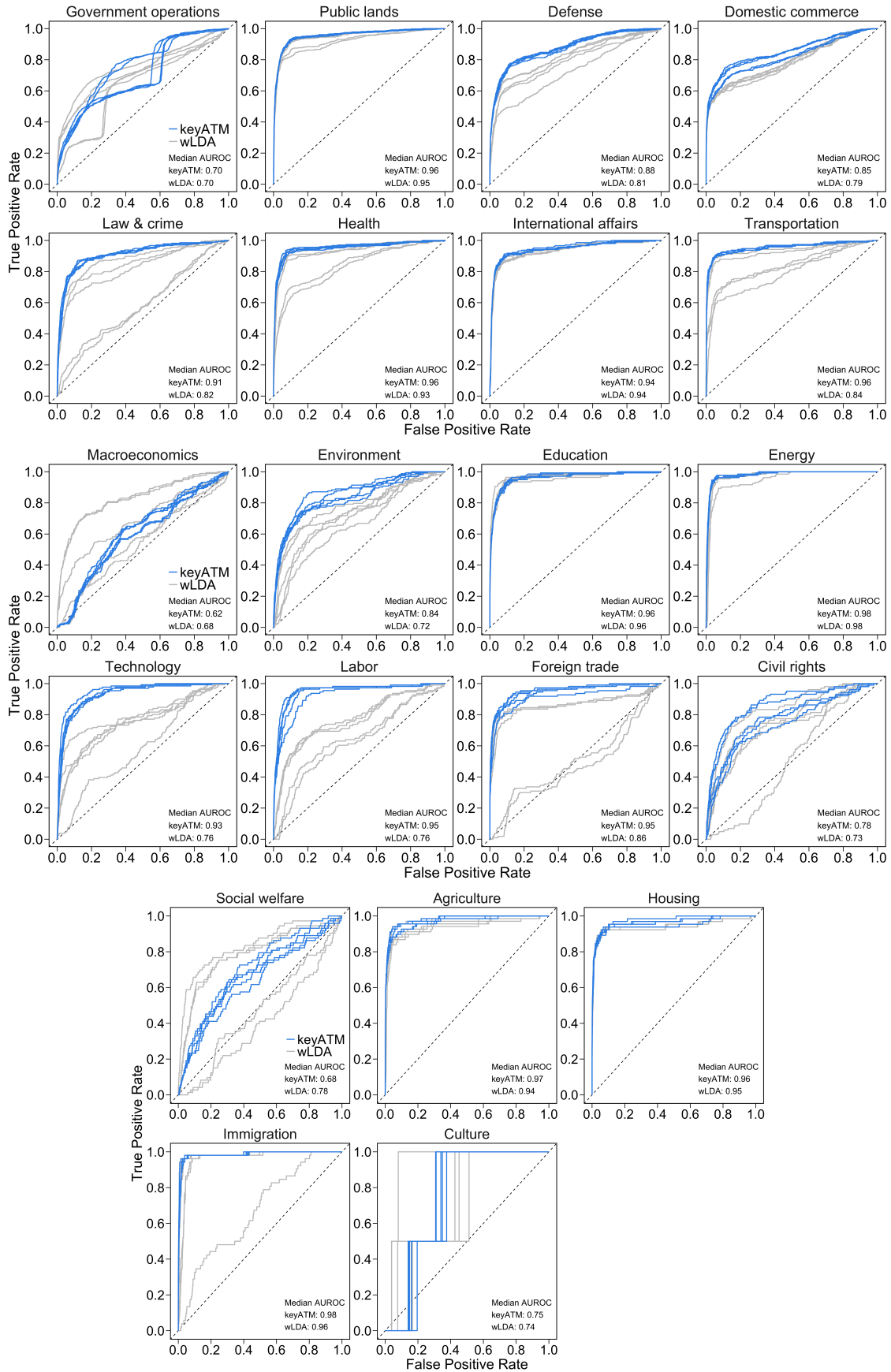


Figure S4: Comparison of the ROC curves between keyATM and wLDA with a different set of keywords

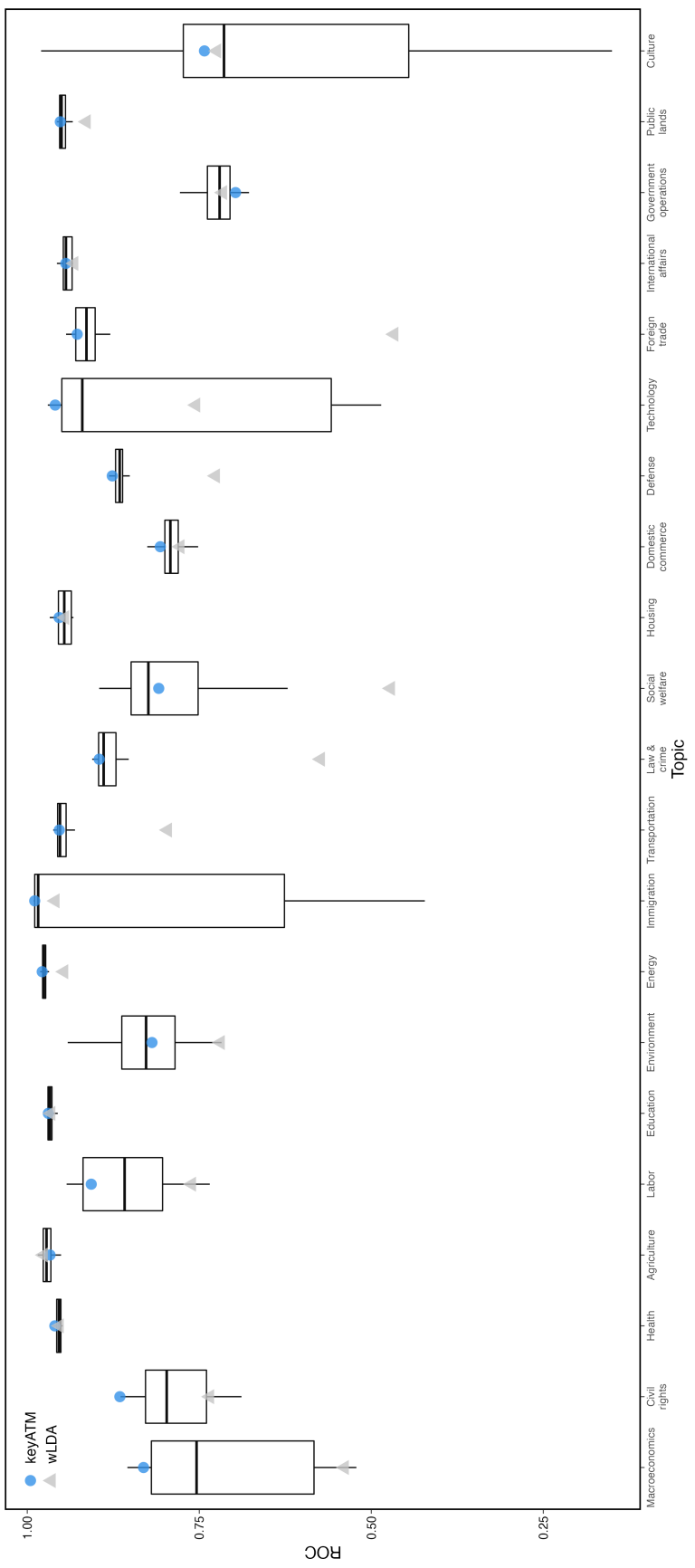


Figure S5: **AUROC with randomly selected keyword sets.** The boxplots present the AUROC with randomly selected 50 different keyword sets. Blue circle indicates the AUROC for keyATM with all keywords and grey triangles represent the AUROC for wLDA (those presented in the Figure S1).

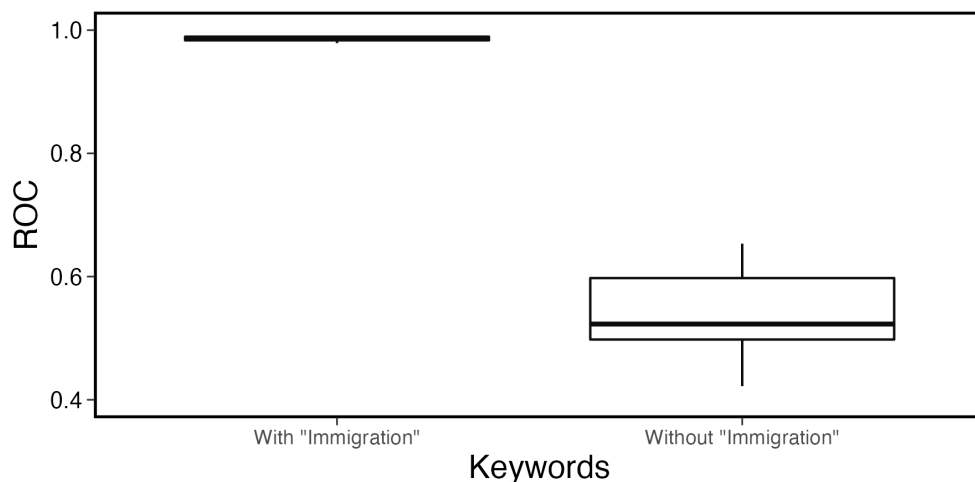


Figure S6: **Comparison of AUROC for the *Immigration* topic between the keyword set with and without “immigration”.** The figure indicates that the topic classification performance is much higher when the randomly selected keyword sets includes “immigration” as one of the keywords for the *Immigration* topic.

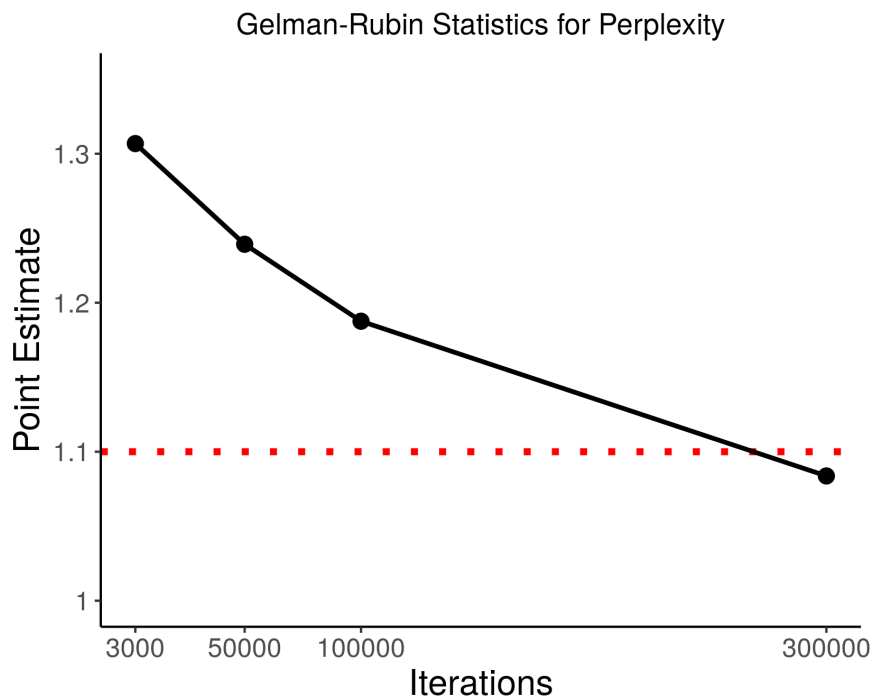


Figure S7: **Gelman-Rubin statistic for perplexity of the base keyATM with different number of iterations.** The x-axis and y-axis indicate the number of iterations and Gelman-Rubin statistic respectively. The figure shows that the statistic decreases monotonically as the number of iterations increases.

C Additional Information for the Covariate keyATM

C.1 Background

In 1994, the government reformed the electoral system that lasted since 1947. The new electoral system, which was first applied in the 1996 election, has two tiers: single-member districts (SMD) and proportional representation (PR) systems. The allocation of SMD seats is independent of that of PR seats, making the Japanese electoral institution a mixed-member majoritarian (MMM) system. Catalinac (2016) finds that the introduction of MMM changed the electoral strategies of LDP candidates, focusing more on programmatic policies rather than pork barrel.

C.2 Keyword Construction Process

The application in Section 3.2 does not have human-coded topics. Catalinac (2016) applies wLDA and then labels all topics. We do not use the most frequent words from the original output because that would lead to analyze the same data as the one used to derive keywords. Instead, we independently construct keywords using the survey questionnaires of the UTokyo-Asahi Surveys (UTAS). This section explains the three steps of obtaining keywords from UTAS: categorizing all labels in Catalinac (2016) into 21 policy areas, matching 21 policy areas with UTAS questions, and selecting keywords from the question.

Prior to every Upper and Lower House election since 2003, the UTAS has fielded a survey to all candidates (both incumbents and challengers) and published its results in the newspaper. Because of its publicity, the survey has a remarkably high average response rate of approximately 85 % (Hirano et al., 2011). Because the UTAS is designed to measure policy issues relevant across multiple elections, it serves as an ideal source of keywords.

Catalinac (2016) fits wLDA with 69 topics and uses 66 of them as pork barrel or programmatic policy topics, excluding three topics as credit claiming topics. Since the most of the topic labels are granular and contain overlapping political issues, we first categorize 66 topics into 21 policy areas to find corresponding questions in the UTAS. Tables S6 and S7 show all 66 topic labels taken from Catalinac (2016) in the middle column and 21 policy areas in the left column. Next, we find UTAS questions that represent each policy area. This process results in the removal of five policy areas (sightseeing, regional revitalization, policy vision, political position, and investing more on human capital) that do not appear in the UTAS.

Finally, we select keywords that represent each policy area from the corresponding UTAS questionnaires. Since most questions are fairly short, we choose nouns so that they match with the preprocessed texts (the preprocessing steps remove the Japanese conjugation). The full list of keywords is presented in Table 3. The original Japanese keywords are all single words, although some of them become two English words after translation.

Policy Area	Original Labels	UTAS Question (Year)
Public Works	Appropriator for the district Fixer-upper for the district Statesperson and appropriator	Subsidies to the local government should be abolished in general (2003), It is necessary to secure employment by public works (2009)
Road Construction	Transportation	We should reduce the number of new highways and make existing ones free (2003), We should privatize four public highway corporations (2003), We should keep the road budget (2009)

Sightseeing	Primary industries and tourism Health and leisure infrastructure	No corresponding questions
Regional Revitalization	Building a spiritually-rich community Hometown development Revitalizing the local community Local facilities and infrastructure Building a safe, reassuring community Benefits for organized groups Love of my hometown Catching up with the rest of Japan Agriculture, forestry and fisheries	No corresponding questions

Table S6: **Pork barrel labels and UTAS questions:** The left column lists the policy areas created by the authors, whereas the middle column shows the original topic labels taken from Catalinac (2016). The right column presents the questions from the UTAS, translated from Japanese to English by authors.

Policy Area	Original Labels	UTAS Question (Year)
Regional devolution	Regional devolution	We should reduce the state subsidies and transfer tax revenue resources to municipalities (2003), We should merge municipalities (2003)
Tax	No more unfair taxes, peace constitution No tax increases, no U.S.-Japan alliance Consumption tax is to fund the military No tax increase, no constitutional revision Tax cuts for everyone No consumption tax, no constitutional revision	We should increase the consumption tax for stable pension system (2003), Do you agree or disagree with increasing the consumption tax for social security and financial reconstruction? (2005), It is inevitable to increase the consumption tax within five years (2009)
Economic recovery	Economic recovery Economic stimulus Fiscal reconstruction	It is urgent to finish the deflation, so instead of cutting budget for fiscal reconstruction, we should implement a fiscal stimulus program to boost the economy (2003, 2005)
Global economy	Japan in global economy	We should protect the domestic industry (2009), We should promote trade and investment liberalization (2009)
Alternation of government	Political and administrative reform No more big business-favoritism Political reform Reforming Japan No more iron triangle Problems facing Japan Small government Alternation of government	What would be the ideal political framework (2005), In general alternation of government brings better politics (2009)
Constitution	Political reform, protect the constitution	Do you think we should change the constitution? (2005)
Party	No more LDP, no more public works Doing away with decayed LDP politics No other party can be trusted	Do you think it is good for your party to join the government coalition? (2009)
Postal privatization	Postal privatization No postal privatization Post offices	There are politicians who could not get an authorization from LDP because of the rebellion against postal privatization. If they form a new party, what this new party should do after the election? (2005), Do you support the postal privatization law, which was rejected in the Upper House? (2005)
Inclusive society	Building a society kind to women	The government should implement special policies to increase the number of women who have the higher positions and better jobs (2003), The basic form of the family consists of a couple and their children (2009), The forms of family should be diverse such as a single mother and DINKS (2009)
Social welfare	Welfare and medical care No reform of medical care Free medical care, no military spending Nursing care	Even if it end up with lowering the quality of social welfare, the smaller government is better (2003, 2005, 2009)
Pension	Protecting people Pensions and child allowance From roads to pension	What do you think about merging the national pension, employees' pension, and the mutual aid pension systems? (2005), We should use tax for the universal pension (2009)

Education	Investing in young people Social security and child support Better education and child-care facilities Pensions and child allowance	Education should respect the precedent methods rather than cultivating child's individuality (2009)
Environment	Saving the natural environment Earthquakes and nuclear accidents	We need to sacrifice the standard of living to protect the environment (2009), Environmental issues are not as important as to sacrifice the standard of living (2009)
Security	Foreign and national security policy No American bases Opposition to military spending Stubbornly for peace and human rights Security and reassurance Not a strong military but a kind society	Japan should strengthen its defense capacity (2003, 2005), Japan should not have nuclear weapons (2003), Japan-US alliance should be strengthened (2003, 2005), We should not hesitate a preemptive attack if there is an expected attack from (2003, 2005), Japan should join the United Nations Security Council to actively play an international role (2003, 2005), We should use dialogue rather than pressures to North Korea (2005, 2009), The government should change the interpretation of the constitution to exert the right to collective defense (2005), Japan-US alliance is the basis of the Japanese diplomacy (2009), Japanese diplomacy should be centered around the United Nations (2009)
Policy vision	Vision for Japan Politics for the civilian, not for bureaucrats	No corresponding questions
Political position	Liberal democracy is best!	No corresponding questions
Investing more on human capital	From concrete to people	No corresponding questions

Table S7: **Programmatic policy labels and UTAS questions:** The left column lists the policy areas created by the authors, whereas the middle column shows the original topic labels taken from Catalinac (2016). The right column presents the questions from the UTAS, translated from Japanese to English by the authors.

C.3 Top Words

Public works		Road construction		Regional devolution	
keyATM	STM	keyATM	STM	keyATM	STM
politic	politic	development	tax	reform	reform
Japan	Japan	road	reduced tax	rural area	administration
society*	society	city	yen	administration	tax
citizen	livelihood	construction	housing	tax*	civilian
protect	citizen	tracks	realize	Japan	consumption
livelihood	protect	budget	daily life	politic	rural area
secure	secure	realize	move forward	citizen	fiscal policy
LDP	budget	promote	city	country	devolve
works	constitution	move forward	education	society*	nursing
public	LDP	early	measure	consumption*	bureaucrat
Tax		Economic recovery		Global economy	
keyATM	STM	keyATM	STM	keyATM	STM
Japan	Japan	reform	reform	development	development
tax	citizen	measure	postal	industry	community
citizen	JCP	society*	privatize	promote	road
JCP	politic	Japan	Japan	prefecture	promote
consumption	tax	economic climate	rural area	agriculture	industry
politic	consumption	reassure	country	plan	street
tax increase	tax increase	economy	citizen	community	agriculture
oppose	oppose	institution	safe	agriculture and forestry	prefecture
business	business	safe	government	fishery	promote
protect	protect	support	pension	enrich	plan
Alternation of government		Constitution		Party	
keyATM	STM	keyATM	STM	keyATM	STM
government	yen	constitution	consumption	Japan	politic
alternation	citizen	consumption*	tax	politic	business
yen	politic	tax*	politic	JCP	LDP
citizen	Japan	protect	abolish	LDP	citizen
politic	medical	tax increase*	citizen	citizen	Japan

medical	cost	politic	LDP	business	reform
Japan	government	Japan	liberty	protect	donation
trillion	tax	livelihood	Japan	party	JCP
elderly	trillion	society*	rice	donation	plutocracy
cost	consumption	peace	agriculture	reform	party
Postal privatization		Inclusive society		Social welfare	
keyATM	STM	keyATM	STM	keyATM	STM
privatize	tax increase	politic	politic	politic	politic
tax increase*	constitution	civilian	reform	society	rich
postal	tax	society*	new	Japan	society
yen	consumption	participate	realize	reform	hometown
LDP	protect	peace	citizen	education*	building
post	oppose	welfare*	government	rich	welfare
Japan	Japan	aim	daily life	building	make effort
oppose	LDP	human rights	rural area	realize	heart
protect	deterioration	realize	corruption	century	move forward
ordinary people	yen	consumption*	change	welfare	plan
Pension		Education		Environment	
keyATM	STM	keyATM	STM	keyATM	STM
pension	pension	politic	Japan	environment	society
yen	institution	Japan	person	society*	reassure
institution	yen	person	country	education*	community
wasteful spend-	medical	children	politic	realize	building
ing					
medical	community	education	necessary	community	education
community	parenting	country	problem	institution	environment
money	wasteful spending	make	children	reassure	support
person	support	force	force	aim	measure
abolish	daily life	have	have	move forward	economic decline
realize	money	problem	future	proceed	employment
Security					
keyATM	STM				
Japan	society				
foreign policy	Japan				
peace	world				
world	economy				
economy	environment				
country	international				
citizen	education				
defense	country				
safe	peace				
international	aim				

Table S9: **Top words of the covariate keyATM.** The table shows the top ten words with the highest estimated probability for each topic under each model. For keyATM, the pre-specified keywords for each topic appear in bold letters whereas the asterisks indicate the keywords specified for another topic.

D Additional Information for the Dynamic keyATM

D.1 Introducing Time Dynamics with HMM

HMM has been used to introduce time dynamic components in various applications. For example, Zhai and Williams (2014) extends LDA to a dynamic setting with the HMM to analyze dialogues. Unlike this model, the dynamic keyATM allows multiple documents to share one state. Quinn et al. (2010) proposes an HMM-based dynamic topic model but only allows for a single topic for each document and no keyword. Although the pioneering dynamic topic model of Blei and Lafferty (2006) uses the Kalman filter, this modeling strategy does not exploit conjugacy, and hence the authors use an approximation to the posterior. Others such as Wang and McCallum (2006) model a distribution of “time stamps” for documents, but this approach is not ideal for social science research because it does not directly model time trends. Indeed, social scientists have effectively used HMM in other settings (e.g., Park, 2012; Knox and Lucas, 2021; Olivella, Pratt and Imai,

2022).

D.2 Keyword Construction Process

The procedure is similar to the one for the base keyATM application. We obtain keywords from the Supreme Court Database project issue description available at their website.¹ After scraping the description from the website, we lemmatize each word using the Python library NLTK and remove stopwords via the R package `quanteda` (Benoit et al., 2018). We then remove words and phrases that have little to do with the substance of each topic. For example, some topic description includes “Note: ”, which explains the background information that does not relate to the substance of topics or specifies the content that should be excluded from the topic. We do not include such descriptions when constructing keywords. Third, we keep the same keywords for multiple topics only if their inclusion can be substantively justified. Lastly, we limit the number of keywords to 25 per topic. We remove terms based on the proportion of keywords among all terms in the corpus if the topic contains more than 25 keywords.

D.3 The Full List of Keywords

Label	Keywords
Criminal procedure	bank call constitutional construction counsel crime criminal death evidence immunity jury justice line obtain procedure prosecution remedy review right rule search sentence statement trial witness
Civil rights	benefit civil constitutional cost counsel discrimination duty employment equal file liability national party plan political protection provision public requirement right school security subject suit
First amendment	amendment bar benefit concern cost election employee failure first form free legislative light material official party political private public regulation religious requirement school security speech
Due process	constitutional defendant due employee hear hearing impartial litigant maker notice prisoner process property resident right statutorily
Privacy	abortion contraceptive die freedom information privacy regulation right
Attorneys	admission attorney bar commercial compensation disbarment discipline employee fee license official speech
Unions	activity agency antitrust bargaining discharge dispute election employee employer fair fund health injunction labor litigation member relation representative right safety standard strike trust union work
Economic activity	business civil claim company contract damage defense determination employee employer evidence federal land liability local official power process property protection public regulation remedy right tax
Judicial power	agency appeal authority circuit claim district evidence federal file grant ground judicial jurisdiction order part party power present private procedure question review right rule suit
Federalism	air child commerce conflict dispute enforcement family federal ground interpretation interstate land legislation national natural obligation primary property regulation relationship resource rest support tax water
Interstate relations	boundary conflict dispute foreign incorporation interstate property relation territory
Federal taxation	business claim entity expense federal fiscal gift internal personal priority private professional provision revenue supremacy tax taxation
Miscellaneous	authority congress executive legislative veto
Private action	civil commercial contract evidence personal procedure property real tort transaction trust

Table S10: **Keywords for the dynamic keyATM:** Keywords used in the dynamic keyATM application

D.4 Top Words

Table S11 presents the top ten words with the highest estimated probabilities defined in Equation (7). The results show that keyATM performs better than wLDA in most topics.

Criminal procedure		Economic activity		Civil rights	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
trial	trial	federal	commission	district*	school
jury	jury	action	commerce	school	district

¹<http://www.supremecourtdatabase.org/documentation.php?var=issue>. Last accessed on March 10, 2020.

defendant*	petitioner	claim	price	discrimination	religious
evidence	evidence	damage	rate	election*	discrimination
criminal	defendant	suit*	interstate	equal	county
sentence	counsel	statute	market	county	election
petitioner	right	right	carrier	vote	vote
judge	rule	rule*	sale	plan	equal
conviction	make	plaintiff	use	one	education
counsel	judge	jurisdiction*	service	race	student
Judicial power		First amendment		Due process	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
appeal	federal	public	public	child*	sentence
district	claim	amendment	first	interest	offense
order	district	first	speech	right	death
petitioner	appeal	government	amendment	process	jury
federal	action	may	interest	may	defendant
claim	judgment	interest	party	due	crime
rule	rule	speech	may	statute	criminal
judgment	petitioner	right*	right	prison	penalty
issue	jurisdiction	can	political	person	punishment
proceeding	order	religious	government	parent	sentencing
Federalism		Unions		Federal taxation	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
land	search	employee	employee	tax	tax
water	officer	union	union	property*	property
indian	police	board	labor	pay	income
tribe	amendment	labor	employer	income	pay
right*	arrest	employer	board	payment	bank
property	warrant	agreement	agreement	interest	interest
use	fourth	employment*	contract	benefit*	corporation
reservation	person	contract*	employment	amount	payment
federal	cause	work	bargaining	plan*	amount
indians	evidence	bargaining	work	fund*	business
Privacy		Attorneys		Interstate relations	
keyATM	wLDA	keyATM	wLDA	keyATM	wLDA
search*	child	agency*	ante	commission	land
officer	benefit	regulation*	rule	commerce*	water
police	interest	use	can	interstate	indian
amendment*	medical	standard*	standard	rate	tribe
arrest	plan	congress*	whether	carrier	right
warrant	provide	require	even	railroad	reservation
fourth	parent	provide	opinion	service	property
evidence*	woman	fee	decision	new	use
person	may	rule*	dissent	gas	indians
use	statute	secretary	apply	line*	river
Miscellaneous		Private action			
keyATM	wLDA	keyATM	wLDA		
congress	congress	price	power		
statute	provision	market	right		
power*	section	sale	government		
federal*	agency	business*	congress		
government	provide	company*	federal		
provision*	federal	sell	amendment		
shall	shall	product	constitution		
section	regulation	competition	statute		
cong	statute	antitrust*	clause		
committee	committee	patent	may		

Table S11: **Top words of the dynamic keyATM.** The table shows the top ten words with the highest estimated probability for each topic under each model. For keyATM, the pre-specified keywords for each topic appear in bold letters whereas the asterisks indicate the keywords specified for another topic.

D.5 ROC Curves

Figure S8 shows ROC curves. Each line represents the ROC curve from one of the five Markov chains with different starting values for keyATM (blue lines) and wLDA (grey lines). The median AUROC indicates the median value of AUROC among five chains for each model. keyATM performs

at least as well as wLDA in 11 out of 14 topics.

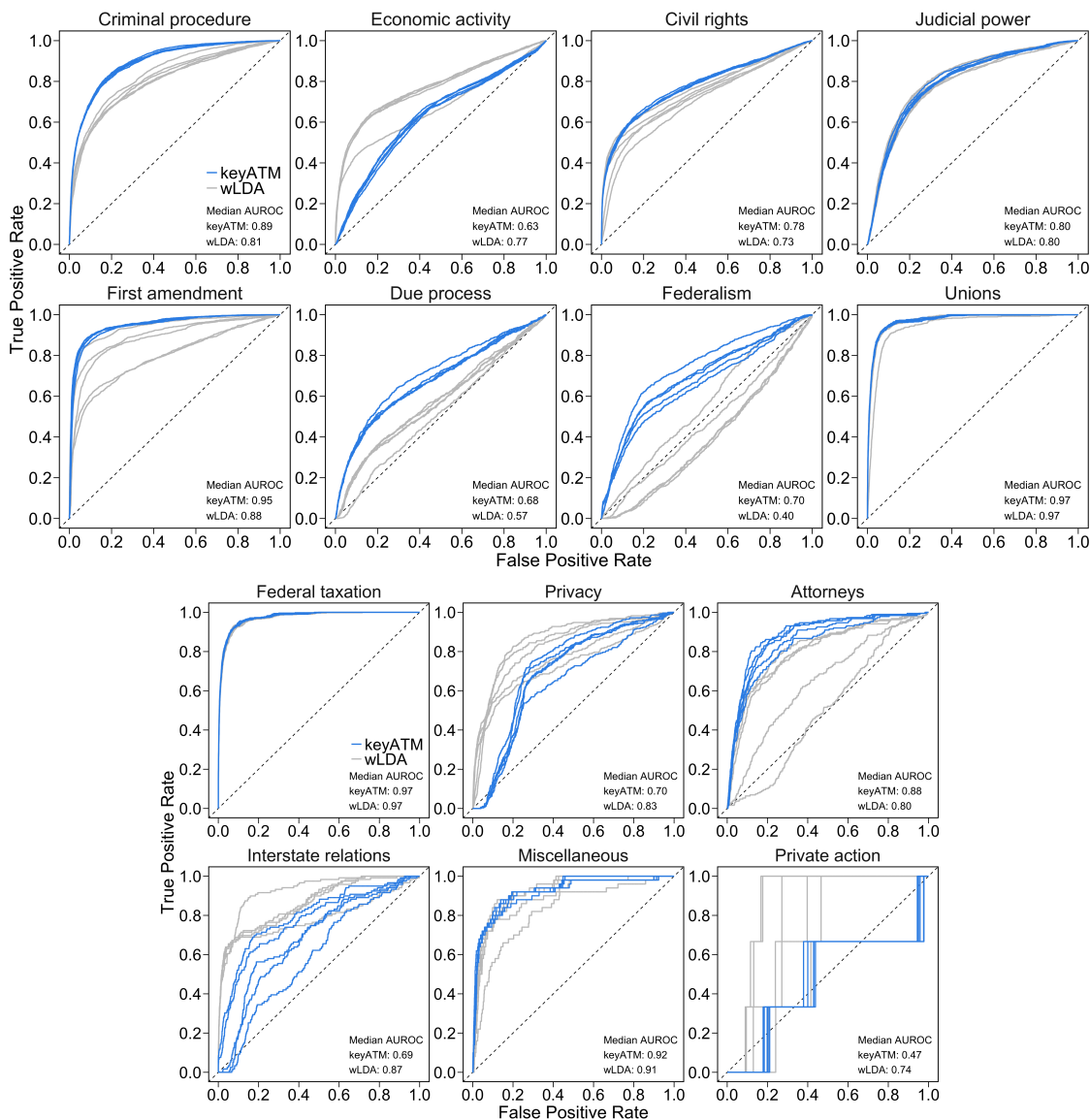


Figure S8: Comparison of the ROC curves between the dynamic keyATM and wLDA

D.6 Correlation of topic prevalence

Table S12 summarizes the results of the the correlation between the estimated topic prevalence and the SCD human coding. To compare the estimated topic prevalence and the SCD human coding, we use the standardized measure that subtracts its mean from each data point and then divide it by its standard deviation. The results show that keyATM exhibits a higher correlation with the human coding for most topics (9 out of 13) than wLDA. Note that the result for *Private action* is not shown because, according to the SCD codebook, all documents associated with this topic are from the same year (2012).

Topic	Dynamic keyATM	Dynamic wLDA
Criminal procedure	0.83	0.07
Economic activity	0.46	0.44
Civil rights	0.76	0.70
Judicial power	0.24	0.40
First amendment	0.82	0.64
Due process	0.37	0.11
Federalism	0.11	0.04
Unions	0.78	0.74
Federal taxation	0.80	0.79
Privacy	0.07	0.18
Attorneys	0.18	0.25
Interstate relations	0.24	0.34
Miscellaneous	0.12	-0.02

Table S12: **Comparison of the time trends of topic prevalence between the dynamic keyATM / wLDA and the SCD human coding.** The table shows the correlation between the estimated topic prevalence and the SCD human coding. To compare the estimated topic prevalence and the SCD human coding, we use the standardized measure that subtracts its mean from each data point and then divide it by its standard deviation.

D.7 Quality of Keywords

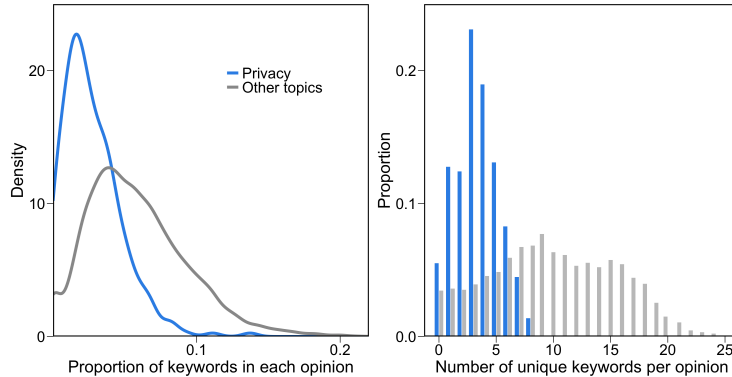


Figure S9: **Poor quality of keywords for the *Privacy* topic.** The left panel presents the histogram for the proportion of keywords among all words in each of the opinions that are classified to the *Privacy* topic (blue bars) by the Supreme Court Database (SCD) human coders. Compared to the other five topics (grey bars) from Table 6, the keywords appear less frequently in the *Privacy* topic. The right panel shows the histogram for the number of unique keywords that appear in each of the opinions associated with the *Privacy* topic by SCD. The number of unique keywords are smaller in the documents classified to the *Privacy* topic than the other five topics.

Similar to the result of Section B.6, the quality of keywords seems to matter for the poor results of the interpretability and classification by the dynamic keyATM for the *Privacy* topic. The left

panel of Figure S9 shows the histogram for the proportion of keywords in each of the opinions that are classified to the *Privacy* topic by the SCD human coding. The keywords for the *Privacy* topic appear much less frequently than the other selected topics from Table 6 (grey bars). The right panel of the figure presents the histogram for the number of unique keywords that are contained in each opinion associated with the *Privacy* topic and other five topics by SCD. The figure indicates that on average fewer keywords appear in the *Privacy* related opinions compared to the corresponding keywords for the other five topics. As similar to the application of the base keyATM, these results suggest that researchers need to pay a careful attention to the selection of keywords when fitting keyATM.

E Additional Information for the Topic Matching of wLDA

For keyATM, there is no need to label topics with pre-specified keywords after model fitting. In contrast, wLDA requires the post-hoc labeling of the resulting topics. Here, we determine the topic labels such that the document classification performance of wLDA is maximized.

We use the Hungarian algorithm to match the topic labels to the resulting topics by maximizing the area under the receiver operating characteristics (AUROC). To account for the multi-label classification structure, we calculate the harmonic mean of binary classification measures across each label with the R package `multiROC` (Wei, Wang and Jia, 2018).

F Additional Information for the Validation

F.1 Design

We recruited participants via Amazon Mechanical Turk (for English tasks) and CrowdWorks (for Japanese tasks). Participants from both platforms conducted tasks on Qualtrics, which makes recruiting platform the only difference between English and Japanese tasks.

Once they agree on the consent form, participants read the instruction and start answering them. We modify methods proposed by Ying, Montgomery and Stewart (2021) to implement our validation exercises. The package to implement their method is available at <https://github.com/Luwei-Ying/validateIt>. In both coherency design and coherency-and-label design, participants choose one option per each task. For each participant, we randomize tasks for keyATM and baseline models as well as the order of choice options in order to deal with the possible ordering effect. Each participant works on 11 tasks of the same design from a single empirical application, five for the output of keyATM, another five for the output of wLDA, and one gold-standard task to assess the worker quality. Examples of coherency design and coherency-and-label design are shown in Figures S10 and S11, respectively.

We track the quality of workers by randomly inserting a gold-standard task for each exercise. We prepare the gold-standard tasks for each of three applications separately so that the gold-standard tasks mimic texts used in three applications. We carefully chose the gold-standard tasks so that they are simple and unambiguous. Examples of gold-standard tasks are shown in Table S14. In each task, we drop all responses given by workers who fail to provide the correct answers to any of gold-standard tasks they conduct.

In addition to the coherency and coherency-and-label design, we also tried topic intrusion design (T8WSI) using the covariate keyATM application. In this task, each worker is shown an actual

document and four sets of words. Each of the four word sets contains the eight highest probability words for a topic (i.e., four sets of eight words). Three of these topics correspond to the highest probability topics for the displayed document, while one is a low probability for that document. The covariate keyATM application is the most suitable for T8WSI because reading the candidate manifestos does not require expert knowledge. Unfortunately, the number of workers who provided the correct answer to our gold-standard task is much lower for T8WSI compared to the coherency task (R4WSI) and the coherency-and-label task (modified R4WSI). Indeed, 18.7% of workers (17 out of 91) failed to provide the correct answer in topic intrusion task whereas the number is much lower for coherency task (6.7%, 6 out of 90) and coherency-and-label task (8.0%, 7 out of 87). This finding is consistent with the observation made by Ying, Montgomery and Stewart (2021) who wrote, “In our initial testing, we found that the WI and T8WSI tasks were often too difficult for coders, reducing their power to discriminate. Further, T8WSI is sensitive to the words included in the ‘top eight,’ making the results more arbitrary and again less informative.”

	Coherency	Coherency-and-label
Base model (Legislative bills)	105	109
Covariate model (Manifestos)	80	84
Dynamic model (Court opinions)	125	110

Table S13: **Number of workers assigned to each design.** Note that we ask workers to complete 11 tasks of the same kind from a single empirical application: five tasks for keyATM, five tasks for the baseline model, and one gold-standard task.

Method	Label	Option 1	Option 2	Option 3	Correct choice
Coherency	NA	health, period, medical, medicare	coverage, insurance, part, care	hospital, cost, provider, payment	energy, gas, vehicle, oil
Coherency-and-Label	Transportation	airport, car, rail, ship	safety, traffic, aviation, carry	transit, air, motor, passenger	college education, excellence, language

Table S14: **Examples of the gold-standard tasks.** For both methods, first three options are from the same topic and the last choice is from another one. Note that we randomize the order of options in the actual tasks.

F.2 Alternative design

We do not include Word Intrusion (WI) and Top 8 Word Set Intrusion (T8WSI) in our validation exercise because, as Ying, Montgomery and Stewart (2021) note, both are difficult tasks, and the latter is particularly known to be sensitive to the choice of displayed words.

We also tried T8WSI using the covariate keyATM application. In this task, each worker is shown an actual document and four sets of words. Each of the four word sets contains the eight highest probability words for a topic (i.e., four sets of eight words). Three of these topics correspond to the highest probability topics for the displayed document, while one is a low probability for that

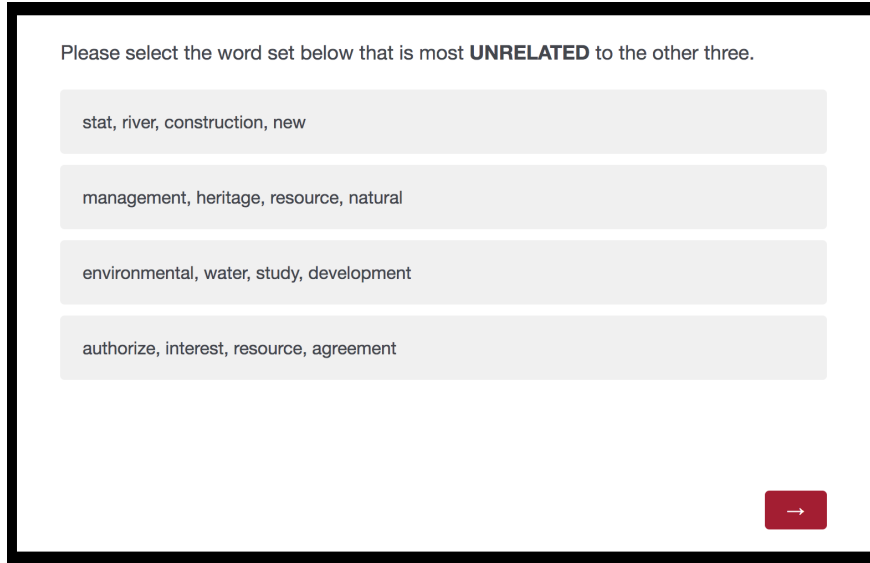


Figure S10: **Screen shot of a task from the coherency validation exercise.**

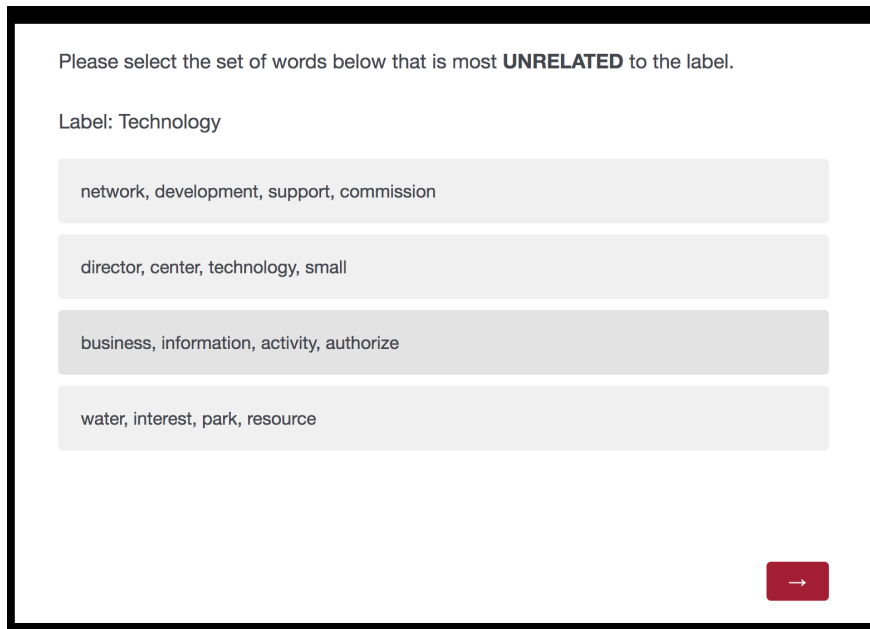


Figure S11: **Screen shot of a task from the coherency-and-label validation exercise.**

document. The covariate keyATM application is the most suitable for T8WSI because reading the candidate manifestos does not require expert knowledge. Unfortunately, the number of workers who provided the correct answer to our gold-standard task is much lower for T8WSI compared to the coherency task (R4WSI) and the coherency-and-label task (modified R4WSI). Indeed, 18.7% of workers (17 out of 91) failed to provide the correct answer in topic intrusion task whereas the number is much lower for coherency task (6.7%, 6 out of 90) and coherency-and-label task (8.0%, 7 out of 87).

F.3 Descriptive Statistics

This section lists descriptive statistics by model and exercise. Each topic has about twenty to thirty tasks. The number of tasks is not fixed because we randomly chose tasks to present. In Figures 5 and S12, we show both pooled and topic-by-topic results.

Topic	keyATM		wLDA	
	# of correct tasks	# of tasks	# of correct tasks	# of tasks
Agriculture	18	24	18	24
Civil rights	13	23	9	23
Culture	21	27	16	27
Defense	19	23	13	23
Domestic commerce	22	30	22	30
Education	19	21	16	21
Energy	40	42	32	42
Environment	21	31	21	31
Foreign trade	13	25	21	25
Government operations	17	29	25	29
Health	17	20	17	20
Housing	11	17	13	17
Immigration	12	16	8	16
International affairs	23	26	23	26
Labor	16	24	12	24
Law and crime	30	33	24	33
Macroeconomics	8	19	5	19
Public lands	13	17	14	17
Social welfare	14	23	15	23
Technology	13	18	8	18
Transportation	30	37	29	37
Total	390	525	361	525

Table S15: **Summary statistics for the coherency validation exercises with base model (legislative bills)**. This table shows the number of correct tasks and the number of total tasks for each model and for each topic.

F.4 Topic-by-topic result

Figure S12 shows the topic-by-topic results. keyATM performs better than wLDA in most topics. The performance of keyATM is worse than wLDA for some topics in the Supreme Court opinion application. The quality of these topics is lower for keyATM than wLDA. For example, *Economic activity* does not perform well in keyATM and the AUROC curve for this topic is lower for keyATM than wLDA (Figure S8). Also, the top ten frequent words for these topics are less interpretable for keyATM. As shown in Table S11, wLDA includes more meaningful words in *Economic activity* than keyATM.

Topic	keyATM		wLDA	
	# of correct tasks	# of tasks	# of correct tasks	# of tasks
Agriculture	15	18	16	18
Civil rights	13	23	12	23
Culture	11	26	4	26
Defense	20	24	20	24
Domestic commerce	24	31	19	31
Education	30	32	29	32
Energy	20	25	16	25
Environment	10	27	17	27
Foreign trade	12	26	11	26
Government operations	6	25	18	25
Health	25	27	25	27
Housing	19	22	17	22
Immigration	23	29	19	29
International affairs	25	29	25	29
Labor	18	26	13	26
Law and crime	28	30	15	30
Macroeconomics	15	28	10	28
Public lands	15	19	16	19
Social welfare	14	25	9	25
Technology	20	28	16	28
Transportation	22	25	18	25
Total	385	545	345	545

Table S16: **Summary statistics for the coherency-and-label validation exercise with base model (legislative bills)**. This table shows the number of correct tasks and the number of total tasks for each model and for each topic.

Topic	keyATM		STM	
	# of correct tasks	# of tasks	# of correct tasks	# of tasks
Alternation of government	9	30	7	23
Constitution	4	21	14	18
Economic recovery	14	26	12	27
Education	10	23	11	21
Environment	13	23	12	25
Inclusive society	12	32	6	25
Party	10	31	5	19
Pension	6	31	12	34
Postal privatization	11	32	18	35
Public works	8	19	7	18
Regional devolution	13	26	9	28
Road construction	19	26	14	27
Security	8	17	7	18
Social welfare	5	22	11	28
Tax	6	19	7	25
Trade	13	22	11	29
Total	161	400	163	400

Table S17: **Summary statistics for the coherency validation exercise with covariate model (manifestos)**. This table shows the number of correct tasks and the number of total tasks for each model and for each topic.

Topic	keyATM		STM	
	# of correct tasks	# of tasks	# of correct tasks	# of tasks
Alternation of government	11	31	12	20
Constitution	10	28	7	25
Economic recovery	15	30	16	33
Education	11	23	16	28
Environment	6	18	11	27
Inclusive society	10	18	8	23
Party	16	39	12	29
Pension	12	22	9	20
Postal privatization	18	29	10	37
Public works	7	26	7	22
Regional devolution	15	26	13	26
Road construction	26	33	5	19
Security	25	28	13	30
Social welfare	11	27	8	20
Tax	4	24	8	30
Trade	10	18	12	31
Total	207	420	167	420

Table S18: **Summary statistics for the coherency-and-label validation exercise with covariate model (manifestos)**. This table shows the number of correct tasks and the number of total tasks for each model and for each topic.

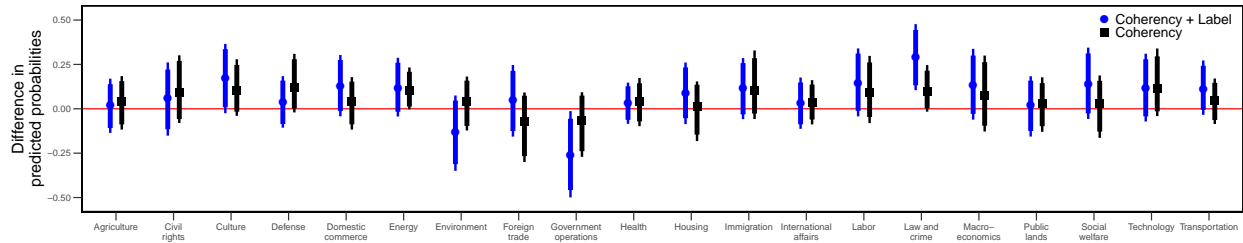
Topic	keyATM		wLDA	
	# of correct tasks	# of tasks	# of correct tasks	# of tasks
Attorneys	16	44	15	44
Civil rights	19	39	24	39
Criminal procedure	47	53	44	53
Due process	20	40	31	40
Economic activity	28	47	32	47
Federal taxation	28	39	34	39
Federalism	42	54	40	54
First amendment	24	42	24	42
Interstate relations	31	37	32	37
Judicial power	34	48	38	48
Miscellaneous	34	41	27	41
Privacy	21	30	20	30
Private action	46	63	47	63
Unions	44	48	45	48
Total	434	625	453	625

Table S19: **Summary statistics for the coherency validation exercise with dynamic model (court opinions)**. This table shows the number of correct tasks and the number of total tasks for each model and for each topic.

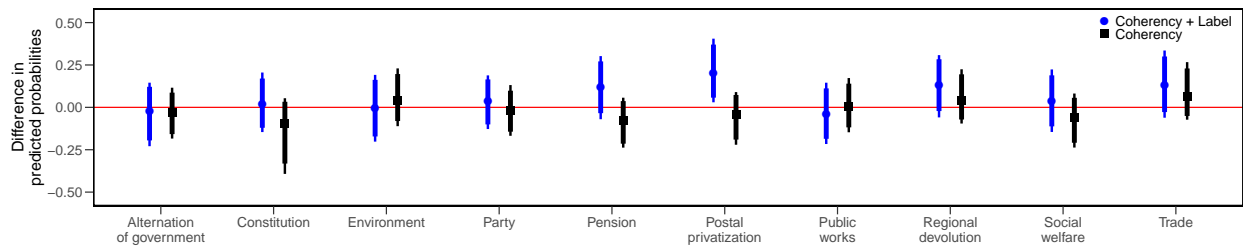
Topic	keyATM		wLDA	
	# of correct tasks	# of tasks	# of correct tasks	# of tasks
Attorneys	10	32	13	32
Civil rights	15	35	20	35
Criminal procedure	29	33	27	33
Due process	9	31	28	31
Economic activity	14	46	36	46
Federal taxation	38	48	39	48
Federalism	22	46	15	46
First amendment	37	42	27	42
Interstate relations	14	35	21	35
Judicial power	30	40	24	40
Miscellaneous	27	51	21	51
Privacy	22	37	16	37
Private action	17	33	14	33
Unions	38	41	37	41
Total	322	550	338	550

Table S20: **Summary statistics for the coherency-and-label validation exercise with dynamic model (court opinions).** This table shows the number of correct tasks and the number of total tasks for each model and for each topic.

(a) Base Model: Legislative Bills



(b) Covariate Model: Manifestos



(c) Dynamic Model: Court Opinions

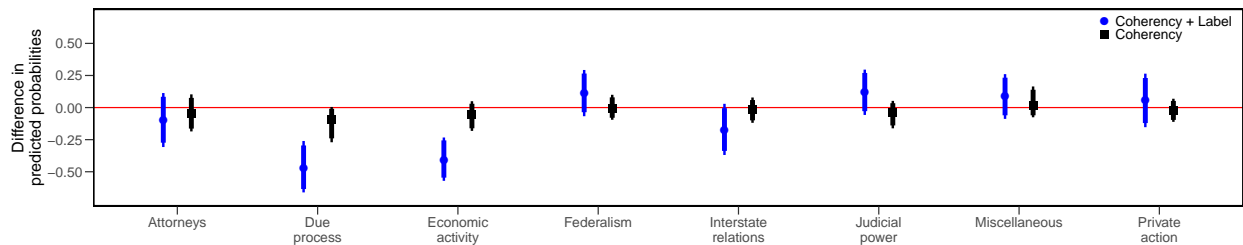


Figure S12: **Comparison of the performance of validation results between keyATM and its baseline counterparts.** Each point represents the difference in the predicted probabilities derived from the three different regression. The thick and thin vertical lines indicate the 95% and 90% credible intervals respectively. Black lines and blue lines show results from the coherency task and the coherency-and-label task respectively.

References

- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller and Akitaka Matsuo. 2018. “quanteda: An R package for the quantitative analysis of textual data.” *Journal of Open Source Software*.
- Bird, Steven, Ewan Klein and Edward Loper. 2009. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O’Reilly Media, Inc.
- Blei, David M. and John D. Lafferty. 2006. Dynamic Topic Models. In *ICML*. pp. 113–120.
- Catalinac, Amy. 2016. “From Pork to Policy: The Rise of Programmatic Campaigning in Japanese Elections.” *The Journal of Politics* 78:1–18.
- Chib, Siddhartha. 1996. “Calculating posterior distributions and modal estimates in Markov mixture models.” *Journal of Econometrics* 75:79–97.
- Chib, Siddhartha. 1998. “Estimation and comparison of multiple change-point models.” *Journal of Econometrics* 86:221–241.
- Fan, Angela, Finale Doshi-Velez and Luke Miratrix. 2019. “Assessing topic model relevance: Evaluation and informative priors.” *Statistical Analysis and Data Mining* 12:210–222.
- Hirano, Shigeo, Kosuke Imai, Yuki Shiraito and Masaaki Taniguchi. 2011. “Policy Positions in Mixed Member Electoral Systems: Evidence from Japan.” <https://imai.fas.harvard.edu/research/files/japan.pdf>.
- Knox, Dean and Christopher Lucas. 2021. “A Dynamic Model of Speech for the Social Sciences.” *American Political Science Review*.
- Lu, Bin, Myle Ott, Claire Cardie and Benjamin Tsou. 2011. Multi-aspect sentiment analysis with topic models. In *ICDM*. pp. 81–88.
- Manning, Christopher D., Prabhakar Raghavan and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press.
- Mochihashi, Daichi. 2020. “Unbounded Slice Sampling.” *ISM Research Memorandum No. 1209*.
- Olivella, Santiago, Tyler Pratt and Kosuke Imai. 2022. “Dynamic Stochastic Blockmodel Regression for Network Data: Application to International Militarized Conflicts.” *Journal of the American Statistical Association*.
- Park, Jong Hee. 2012. “A Unified Method for Dynamic and Cross-Sectional Heterogeneity: Introducing Hidden Markov Panel Models.” *American Journal of Political Science* 56:1040–1054.
- Quinn, Kevin M, Burt L Monroe, Michael Colaresi, Michael H Crespin and Dragomir R Radev. 2010. “How to analyze political attention with minimal assumptions and costs.” *American Journal of Political Science* 54:209–228.

- Wang, Xuerui and Andrew McCallum. 2006. Topics over Time: A non-markov continuous-time model of topical trends. In *International Conference on Knowledge Discovery and Data Mining*. pp. 424–433.
- Wei, Runmin, Jingye Wang and Wei Jia. 2018. *multiROC: Calculating and Visualizing ROC and PR Curves Across Multi-Class Classifications*. R package version 1.1.1.
- Ying, Luwei, Jacob M Montgomery and Brandon M Stewart. 2021. “Topics, Concepts, and Measurement: A Crowdsourced Procedure for Validating Topics as Measures.” *Political Analysis* Forthcoming.
- Zhai, Ke and Jason D. Williams. 2014. Discovering Latent Structure in Task-Oriented Dialogues. In *ACL*. pp. 36–46.