

Who Counts as “the people”? Measuring Contested Political Concepts from Political Speech using Masked Language Models*

Hyeongjae Lee^{1,*}, Seonwoo Ahn^{1,*}, Sunwoo Kwak²,
June Jeon¹, Jisung Yoon^{3,†}, Taegyeon Kim^{1,†}

¹School of Digital Humanities and Computational Social Sciences,
KAIST, Daejeon, Republic of Korea

²Department of Political Science and International Studies,
Yonsei University, Seoul, Republic of Korea

³KDI School of Public Policy and Management, Sejong, Republic of Korea

*Equal contribution

†Corresponding authors: jsyoon@kdischool.ac.kr, taegyeon@kaist.ac.kr

Abstract

What do political elites mean when they invoke “the people”, and how has this meaning shifted across centuries of democratic history? Political theorists have long held that “the people” is a discursively constructed and contested category. Yet empirical work remains largely qualitative or small in scale, and computational approaches have yet to address how contested concepts themselves are framed over time. We address this gap with a generalizable framework that models contested concepts as configurations of theoretically grounded frames, and apply it to American presidential rhetoric based on three theoretically grounded frames: Nation, Civic, and Global. Leveraging a fine-tuned political masked language model, we mask each occurrence of “people” and score each frame by the pseudo-log-likelihood of its probe words. We apply the framework to 3,402 U.S. presidential speeches from 1789 to 2025. The resulting measurements agree with expert human coding at a level comparable to inter-annotator agreement, while tracking major junctures such as the Civil War, the World Wars, the Civil Rights era, and a recent nationalist turn. They also covary with external indicators as expected: stronger democracy is associated with less national and more civic framing. Beyond “the people”, the framework can generalize to other contested political concepts.

1 Introduction

Across the globe, the question of who counts as “the people” has returned to the center of political conflict.

Resurgent ethno-nationalism, anti-immigration movements, and exclusionary populism are redrawing the boundaries of political membership in ways unseen for generations. Political elites, however, rarely define “the people” explicitly. Instead, as opinion leaders shaping mass political cognition (Zaller, 1992), they invoke “the people” in ways that can quietly reshape its meaning, bounding it by ancestry, by citizenship, or by appeals to a broader human community. These rhetorical shifts are not only descriptive but potentially constitutive: they reshape how citizens perceive the boundaries of their political community and how they treat those they recognize, or refuse to recognize, as civic equals (Anderson, 2006; Smith, 2003; Schmidt, 2008; Stuckey, 2023). Tracking these shifts at scale is therefore crucial for understanding how political membership is reconfigured.

Despite the centrality of “the people” to political theory, empirical study of the concept remains limited. Qualitative scholarship has documented how national identity is constructed in rich detail (Beasley, 2011; Stuckey, 2023), but has not traced these shifts systematically at scale. Existing computational tools each address only part of the task. Unsupervised topic models capture what is discussed rather than how it is framed; supervised approaches, including the fine-tuning of pre-trained models, depend on costly annotation; and zero- and few-shot LLM classification comes at the expense of interpretability and reproducibility. While computational work has advanced the broader study of political rhetoric (Card et al., 2022; Rule et al., 2015), few methods capture the competition among multiple frames over a single contested term at scale.

We address this gap with a generalizable frame-

*The manuscript was prepared with the assistance of an AI assistant. All AI-generated content was reviewed and revised by the authors, who take full responsibility for the accuracy and integrity of the final work. All data, code, and model outputs will be made publicly available at https://github.com/lhch9550/who_counts_as_people.

work that measures which of multiple theoretically grounded frames prevails in defining a contested political construct in a given text. For each occurrence of a target term in a given text, we replace it with a [MASK] token and use a domain-adapted masked language model to estimate how strongly theoretically guided probe words fit the masked position. We apply this framework to “the people” as our first analytical target, drawing on Kohn’s distinction between ethnic and civic nationalism (Kohn, 1944) and on cosmopolitan conceptions of moral community (Nussbaum, 1996) to define three frames: *Nation*, grounded in shared ancestry and bounded sovereignty; *Civic*, grounded in rights-bearing citizenship and constitutional principles; and *Global*, grounded in universalist obligations that transcend state boundaries.

We showcase the framework through an in-depth case study of U.S. presidential rhetoric from 1789 to 2025 (3,402 speeches, 29,424 contextual instances).¹ Three findings together establish the framework’s validity and offer substantive insights. First, Nation framing dominated nineteenth-century rhetoric before Civic framing overtook it in the post-Cold War era, followed by a renewed nationalist turn since 2016. Second, the framing shifts align closely with major historical turning points, including the Civil War, WWI, the Civil Rights era, and 9/11, providing strong face validity. Third, regression analyses show that the frame scores move with indicators of democratic development and macroeconomic conditions in the directions theory predicts.

Our contributions are twofold. On the methods side, we introduce a masking-based probing procedure that scores each frame by combining a fine-tuned political masked language model with a curated dictionary of theory-driven probe words, yielding interpretable frame scores without annotation. Empirically, we present a large-scale longitudinal measurement of how “the people” has been constructed in U.S. presidential rhetoric, and show that the approach can extend to other contested constructs such as “democracy”, “freedom”, and “security”.

¹The U.S. is a multi-cultural polity in which the boundaries of “the people” have been repeatedly redrawn, and a continuous archive of presidential rhetoric that enables longitudinal measurement at otherwise unavailable scale.

2 Data Collection and Preprocessing

2.1 Data Collection

To address our research questions, we constructed a dataset of presidential speeches from two sources: the Miller Center (<https://millercenter.org/>, 1,057 speeches) and the American Presidency Project (<https://www.presidency.ucsb.edu/>). From the latter, we retained only entries classified as *Address* and excluded those likely to introduce noise, such as remarks at informal events and signing statements, leaving 2,345 speeches directed at the general public. Both sources were accessed in March 2026 and span 1789 to 2025. To remove duplicates across the two sources, we computed pairwise semantic similarity using sentence embeddings and, for any pair scoring 0.99 or higher, removed one of the two speeches. This step yielded a final corpus of 3,402 unique presidential speeches.

2.2 Preprocessing

We included speeches containing the token *people*², the central object of analysis. In addition, because the meaning of *people* can vary with local context even within a single speech, we scored each occurrence from its immediate surroundings rather than from the speech as a whole. To this end, we constructed a fixed 100-token window around each occurrence of *people* (50 tokens on each side), truncating at the sequence boundary when fewer than 50 tokens were available. This yielded 29,424 contextual instances of *people* for analysis.

3 Quantifying the Frame Score

To analyze how “the people” is framed in presidential rhetoric, we introduce the Frame Score, a measure that infers a term’s association with ideological categories from the textual context in which it appears. Card et al. (2022) employed a BERT-based masked language model, masking mentions of immigrant groups to quantify the degree of dehumanization in the metaphorical framing applied to them. Building on this approach, we examine how the framing of “the people” has shifted over time in U.S. presidential speeches. Computing the Frame Score requires three components: a domain-adapted masked language model, a curated dictionary of theory-driven probe words, and a scor-

²Throughout this paper, “the people” refers to the conceptual object of analysis, whereas *people* (in italics) denotes the token as it appears in the text.

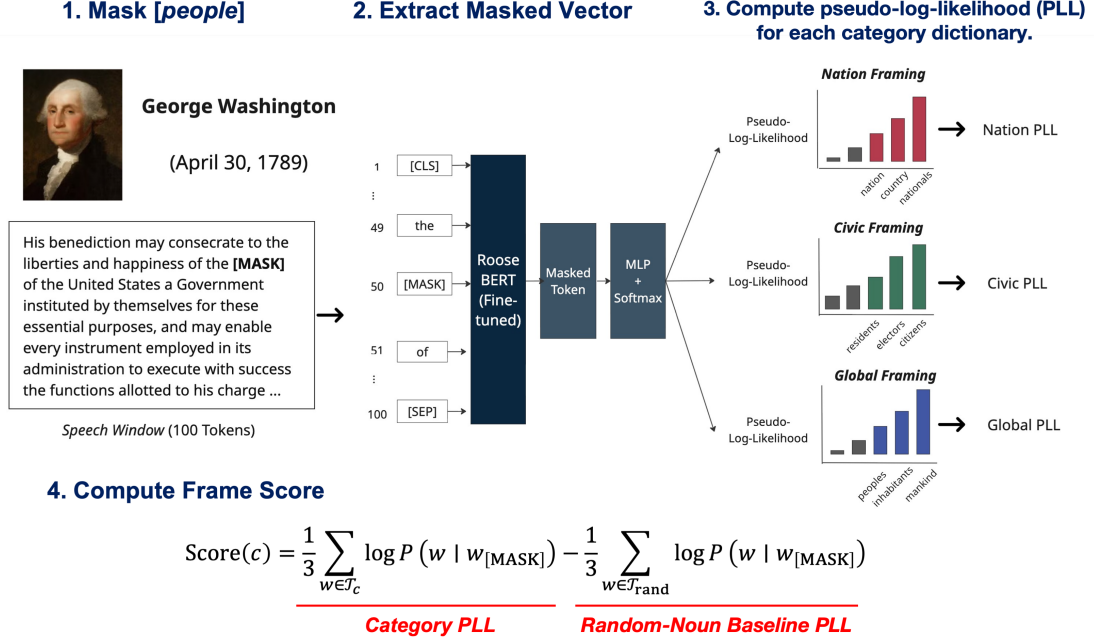


Figure 1: **Schematic illustration of the frame score computation.** For each occurrence of *people*, a 100-token context window is constructed and the target term is replaced with a [MASK] token, then encoded by RoBERTa (cased). For each of the three theory-driven frames, Nation, Civic, and Global, we compute a category-level pseudo-log-likelihood (PLL) by averaging the top-3 dictionary-word probabilities at the masked position, and subtract a random-noun baseline to obtain a baseline-corrected frame score. A higher score indicates a stronger association with the corresponding frame. The example is drawn from George Washington’s inaugural address (April 30, 1789).

ing procedure that evaluates each frame against a random-noun baseline.

3.1 Model Selection and Fine-tuning

We employed RoBERTa-cased (Dore et al., 2025) to capture the semantic properties of presidential speech. RoBERTa-cased is a domain-specific pre-trained language model built on the BERT architecture and trained on English-language corpora of political debate and speech, including U.S. presidential debates, parliamentary debates from nine countries, and United Nations proceedings.³

To confirm that RoBERTa encodes structure relevant to our analysis, we extracted a contextual embedding of every instance of “the people” from the pre-trained model and projected them into two dimensions with UMAP. The instances span a range of contexts that is structured by both era and speaker: individual presidents occupy distinct

³The developers of RoBERTa explicitly identify the analysis of political discourse and its influence on public opinion as a primary motivation for the model (Dore et al., 2025), which aligns directly with the goals of our study. RoBERTa-cased is distributed by the Inria Centre at Université Côte d’Azur for non-commercial academic research, and our use is consistent with these terms.

regions of the space, and each historical era spreads across multiple regions rather than forming a single block (Appendix A.2, Figure 5).

Because the meaning of “the people” may shift over time, our main analysis relies on era-specific models rather than a single static model. Following standard BERT practice (Devlin et al., 2019), we fine-tuned a separate model on each period’s subset of the corpus and computed the frame score for every instance using its corresponding era-specific model. All frame scores reported in the subsequent analyses are derived from these era-specific models.

3.2 Curating Dictionary

Our typology builds on an established tradition in nationalism studies. Kohn (1944) first systematized the distinction between ethnic and civic nationalism, later developed by Meinecke (1970) and Brubaker (1992). The *Nation* frame corresponds to ethnic or cultural nationalism, defining peoplehood through common ancestry, territorial origins, and ethno-cultural markers such as language, religion, and tradition (Kolstø, 2000; Janmaat, 2006;

Reeskens and Hooghe, 2010; Lubbers and Coenders, 2017; Harrison and Bruter, 2011). The *Civic* frame defines membership through shared institutional and legal norms, including constitutional rights, duties, and democratic participation (Brubaker, 1992; Jayet, 2012; Larin, 2020; Tamir, 2019). The *Global* frame extends membership beyond the state, drawing on cosmopolitan theories of moral obligation toward humanity as a whole (Kohn, 1932, 1934; Nussbaum, 1996; Appiah, 2018; de Sousa Santos, 2007; Tinsley, 2019; Mylonas and Tudor, 2021). Though the frames overlap in practice, they mark analytically distinct categories against which we measure contestation in our data.

To measure each frame, we constructed a category dictionary of probe words through a four-step expansion procedure.

1. **Seed selection.** We defined a set of seed words for each category a priori on theoretical grounds.
2. **Neighbor retrieval.** We computed each category’s centroid by averaging the static token embeddings of its seed words from RooseBERT-cased, and retrieved the 20 nearest vocabulary tokens by cosine similarity.
3. **Filtering.** We kept only common nouns, identified via NLTK part-of-speech tagging, that occurred at least 20 times in the corpus. These criteria ensure that probe words denote substantive categories rather than proper nouns or function words, and that each appears often enough to support reliable estimation.
4. **Disambiguation.** When a word fell close to more than one category, we assigned it to the frame whose centroid was nearest. For example, *nationalities* and *continents* appeared among the candidate words for both *Nation* and *Global*, but were assigned to *Global* because they were closer to its centroid.

This procedure expanded each seed set into a fuller dictionary. The complete lists are given below, with seed words in bold.

- ***Nation*:** **nation**, **country**, **state**, **nationals**, nations, countries, states, republic, province, county, continent, homeland, nationality
- ***Civic*:** **citizens**, **voters**, **taxpayers**, **electors**, **residents**, voter, elector, homeowners

Table 1: **Pairwise mean cosine similarities between category word embeddings.** Diagonal entries (in bold) reflect intra-category cohesion; off-diagonal entries report inter-category similarity.

	<i>Nation</i>	<i>Civic</i>	<i>Global</i>
<i>Nation</i>	0.276	0.285	0.327
<i>Civic</i>	0.285	0.505	0.477
<i>Global</i>	0.327	0.477	0.613

- ***Global*:** **peoples**, **mankind**, **humanity**, **humankind**, **inhabitants**, civilization, nationalities, continents

3.3 Dictionary Validation

To assess whether the resulting dictionary is coherent, we measured the cosine similarity among words within and across categories (Table 1). For the Civic and Global categories, words are more similar to others in the same category than to words in other categories, with Global the most tightly grouped (0.613). The Nation category is a notable exception: its within-category similarity (0.276) is lower than its similarity to Civic (0.285) and Global (0.327), reflecting the heterogeneous meanings its words carry. Figure 2 shows the same pattern: the Nation words spread across a broad region rather than forming a tight cluster. Several words sit at the category boundaries, with *continent*, *continents*, *nationalities*, *civilization*, and *peoples* on the Nation–Global border, and the Civic words *citizens* and *residents* bordering the Nation region near *nationals*. Overall, the three theory-based categories are broadly distinguishable in the embedding space but not sharply separated. This overlap is expected: the boundaries of “the people” are themselves contested, and it is precisely this contestation that our frame scores are designed to measure.

3.4 Computing Frame Scores

Our scoring procedure, illustrated in Figure 1, proceeds as follows. For every occurrence of *people* within its 100-token context window, we replace the target word with a [MASK] token and pass the masked window through the fine-tuned RooseBERT model to compute the pseudo-log-likelihood (PLL) of each dictionary word at the masked position (Salazar et al., 2020). We summarize each category *c* by the mean PLL of its three highest-scoring words $\mathcal{T}_c$. Because some contexts favor common

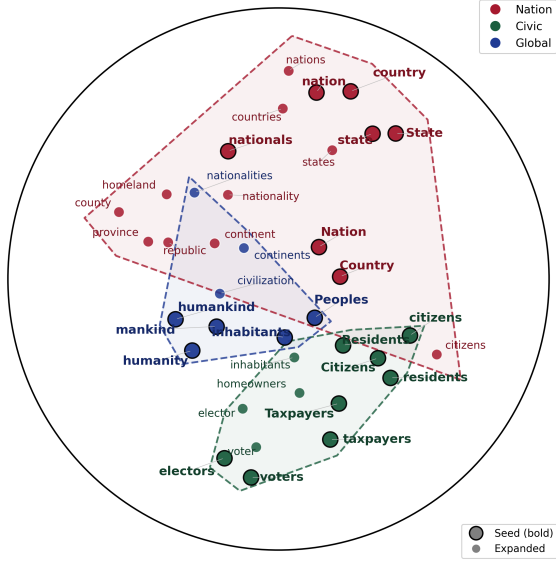


Figure 2: UMAP projection of the expanded dictionary word embeddings onto two dimensions. Each point is a dictionary word, colored by category; seed words are drawn larger with bold labels, expanded words smaller. The three categories occupy partially separable regions, with *Civic* and *Global* lying closer to each other than to *Nation*.

nouns regardless of framing, we then subtract the mean PLL of three randomly sampled nouns $\mathcal{T}_{\text{rand}}$, yielding a baseline-corrected frame score:

$$\text{Score}(c) = \frac{1}{3} \sum_{w \in \mathcal{T}_c} \log P(w | w^{[\text{MASK}]}) - \frac{1}{3} \sum_{w \in \mathcal{T}_{\text{rand}}} \log P(w | w^{[\text{MASK}]}) \quad (1)$$

where $w^{[\text{MASK}]}$ is the masked context.

A higher $\text{Score}(c)$ indicates a stronger association with frame c relative to the random-noun baseline.

3.5 Robustness Checks

We assess the reliability of our framing measurements in three ways: their agreement with human judgment, their sensitivity to the size of the context window, and their sensitivity to the number of probe words aggregated per category.

Human validation. To evaluate whether the model’s frame assignments align with human judgment, we drew a year-stratified sample of 100 speech instances and had two coders independently assign frame labels without access to model outputs. Inter-coder agreement was moderate (Cohen’s $\kappa = 0.354$, raw agreement 62.0%), reflecting

the interpretive difficulty of the task. The model’s agreement with each coder was comparable to the human baseline ($\kappa = 0.452$ with Rater 1, 0.355 with Rater 2). Adding the model as a third rater raised Krippendorff’s α from 0.339 to 0.376, indicating that it introduces no disagreement beyond that already present among the human coders.

On the 62 instances where both coders agreed, model–human agreement exceeded the human–human baseline ($\kappa = 0.552$, accuracy= 0.774, weighted $F_1 = 0.753$), and rose further under a binary distinction ($\kappa = 0.638$ for Nation, 0.587 for Civic). Overall, these results indicate that the model performs at or above human inter-annotator reliability, supporting its use as a scalable proxy for human coding.

Context-window sensitivity. Because subtler framing may require a wider context to identify the intended frame, the choice of window size could affect our measurements. Hence, we varied the context window from 25 to 200 tokens around each occurrence of *people* and computed year- and speaker-level correlations against the 100-token baseline. Agreement was highly stable, with Pearson and Spearman correlations ranging from 0.90 to 0.99 (all $p < 0.001$), confirming that our framing measurements are robust to context length.

Dictionary-size sensitivity. We assessed whether the results depend on the number of probe words aggregated per category by re-running the analysis while varying this number from the top 1 to the top 8 words. Agreement was low only for the top-1 setting (e.g., Global: $r = 0.88$ against top-3), reflecting the rank instability of a single probe word, but stabilized from top-2 onward, with top-3 correlating almost perfectly with top-4 through top-8 ($r \geq 0.92$ across all categories). We therefore adopt top-3. Full correlation matrices are reported in Appendix A.1.

4 Results

4.1 Two Centuries of “the people” Framing

We first aggregated the instance-level scores to the speech level by averaging the Nation, Civic, and Global frame scores across all occurrences of *people* within each speech. These speech-level scores were then averaged by year, and we applied a 5-year centered rolling average to smooth year-to-year fluctuations (Figure 3). The result reveals a

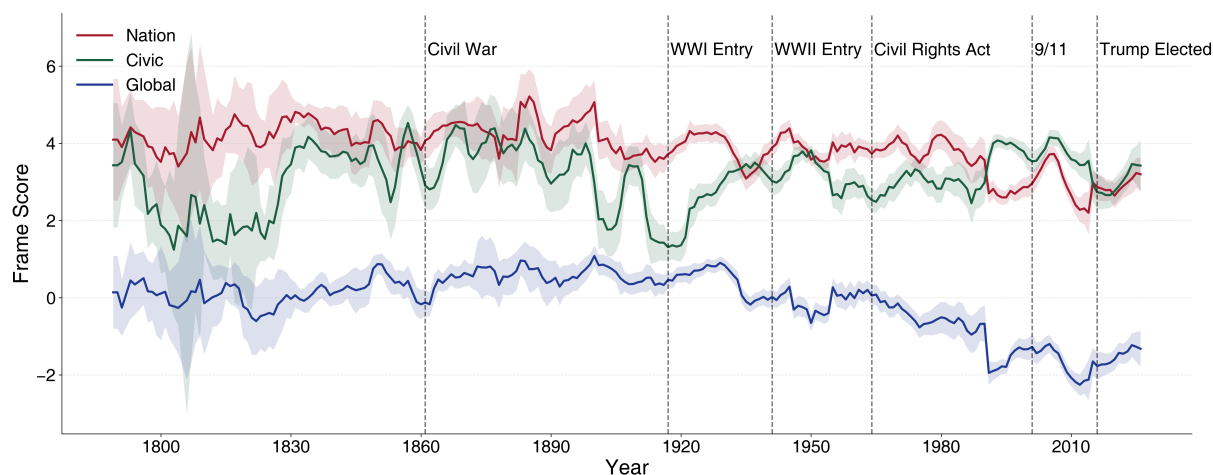


Figure 3: **Temporal evolution of Nation, Civic, and Global framing in U.S. presidential speeches (1789–2025).** Solid lines denote 5-year centered rolling averages of the baseline-corrected frame score. Shaded bands show 95% confidence intervals, and vertical dashed lines mark major historical events. Nation framing dominates the nineteenth century before Civic framing overtakes it in the post-Cold War era, with a renewed nationalist turn after 2016.

broad historical shift in how U.S. presidents have framed “the people”.

Throughout the nineteenth century, Nation framing remained dominant while Civic framing gradually expanded, as the polity absorbed new territories and more diverse populations. This expansion reflected the constitutional ties that bound a growing republic together, even as the franchise itself remained narrow (Smith, 2003). Yet this civic expansion could not avert the crisis of the Civil War (1861–1865). From the late nineteenth century, the pattern reverses: as Reconstruction faltered, mass immigration provoked a nativist backlash, and imperial expansion and WWI mobilization mounted, Nation framing rose while Civic framing receded, reaching its lowest point around U.S. entry into WWI (1917). “The people” came to be defined more by ethnicity and territory than by constitutional membership (Higham, 2002; Gerstle, 2017).

The mid-twentieth century marks a structural turning point. After WWII, Civic framing rose steadily, converging with Nation framing around the Civil Rights Act era (1964). By the 1990s, it had overtaken Nation framing, reflecting a broader shift toward more inclusive, institutional conceptions of peoplehood after the Cold War. However, this shift was not uniform. Civic framing dropped sharply after the September 11 attacks (2001), when Nation framing temporarily resurged amid national-security concerns. After Trump’s election (2016), Nation framing rebounded and Civic framing declined, marking a renewed nation-

alist turn (Bonikowski et al., 2021; Rowland, 2021).

Global framing remained relatively low throughout, consistently below the Nation and Civic frames, and turned negative in the modern era. This is largely expected, since political elites, including U.S. presidents, address a primarily domestic audience. Even so, Global framing was somewhat higher in the nineteenth and early-to-mid twentieth centuries, reflecting moments when presidents extended their appeal beyond the national community, before declining steadily in the post-Cold War decades.

4.2 Identifying Macro-level Predictors using Regression Analyses

To examine how the frame scores covary with socioeconomic and political indicators, we regressed the instance-level Nation, Civic, and Global frame scores on a set of macro-level variables using ordinary least squares. Because all windows from the same year share the same macro-level indicator values, we clustered standard errors by year.

We estimated two models for each frame. The *Long-term model* covers the full available period (1789–2025) and includes the V-Dem Liberal Democracy Index together with a binary indicator for the pre-1960 period, which accounts for structural differences in speech availability and rhetorical context before the mid-twentieth century. The *Modern model* restricts the sample to the post-1960 period (1960–2025) with complete economic data and adds log GDP per capita, trade openness (im-

Table 2: **OLS regression results, estimated at the instance (contextual-window) level.** Macro-level predictors are the V-Dem Liberal Democracy Index, a pre-1960 indicator, log GDP per capita, trade openness, and log net migration. *Long* and *Modern* denote the long-term and modern models. Standard errors, clustered by year to account for shared annual conditions, are reported in parentheses.

	Nation		Civic		Global	
	Long	Modern	Long	Modern	Long	Modern
Liberal Democracy	−3.354*** (0.472)	0.024 (2.517)	2.578*** (0.733)	6.150** (2.610)	−4.708*** (0.700)	−2.040 (2.579)
Pre-1960	−0.250 (0.198)		0.732*** (0.281)		−0.189 (0.315)	
GDP (logged)		−2.679*** (1.032)		−1.160 (0.807)		−2.403** (1.085)
Trade Openness		0.090 (0.057)		0.022 (0.045)		0.085 (0.065)
Net Migration (logged)		−0.539* (0.302)		0.748*** (0.242)		−0.322 (0.324)
Intercept	5.615*** (0.356)	13.688*** (3.613)	1.395** (0.567)	−1.057 (2.882)	2.294*** (0.532)	8.938** (4.262)
<i>N</i>	29,132	12,455	29,132	12,455	29,132	12,455
<i>R</i> ²	0.043	0.023	0.006	0.034	0.110	0.021

* $p < .1$; ** $p < .05$; *** $p < .01$

ports as % of GDP), and log net migration. Because the regression requires complete macro-level data, we drop years with missing values, leaving 29,132 instances for the long-term models and 12,455 for the modern models.

Table 2 presents the results. In the long-term models, the democracy index shows theoretically consistent associations with Nation and Civic framing: higher democracy scores are associated with greater Civic framing (2.578***) and lower Nation framing (−3.354***). Global framing is also negatively associated with the index (−4.708***). These patterns suggest that democratization corresponds to a shift toward civic-institutional conceptions of peoplehood, strengthening constitutional identification while weakening both exclusive national identity and appeals to a community beyond the nation-state.

In the modern models, the democracy index is no longer significantly associated with Nation framing (0.024) but retains a strong positive association with Civic framing (6.150**), indicating that in the post-war period greater democratic consolidation corresponds to markedly more civic-institutional rhetoric. Among the economic covariates, log GDP per capita is negatively associated with both Nation framing (−2.679***) and Global framing

(−2.403**), while net migration is positively associated with Civic framing (0.748***) and negatively, though only marginally, with Nation framing (−0.539*). Trade openness shows no significant association with any frame. These results indicate that rising prosperity coincides with less nationalist and less universalist rhetoric, while migration pressures are accompanied by more civic-institutional appeals rather than more exclusionary ones.

The instance-level models explain only a small share of the variance (R^2 ranging from 0.006 to 0.110), as expected given that macro-level indicators capture systematic shifts in the central tendency of frame scores rather than predicting individual scores. To confirm that our results hold in aggregate, we re-estimated all models by year, averaging instance-level frame scores within each year (Table 3), so that each year contributes equally regardless of its number of speeches. The negative association between the Liberal Democracy Index and Nation framing remains significant (−2.385***). The Civic and Global coefficients no longer reach significance in the smaller sample, but they retain the same sign as in the instance-level models (Civic positive, Global negative).

Taken together, these results validate the frame scores against external indicators in theoretically

expected directions, while revealing how macro-level factors shape the framing of “the people” across different time scales. The long-term model reveals century-scale patterns, with democratization robustly associated with a shift toward civic-institutional conceptions of peoplehood, weakening both exclusive national identification and appeals to a community beyond the nation-state. The modern model, in turn, highlights post-war dynamics, showing how economic openness and migration flows redraw the boundaries of “the people” in ways that complicate conventional accounts of globalization.

5 Discussion

Although political elites frequently invoke “the people”, its meaning remains contested and historically contingent. Treating political language as constitutive rather than merely descriptive, our analysis identifies systematic shifts in how “the people” is framed across three categories—Nation, Civic, and Global—and traces how these framings covary with broader political and economic developments.

Methodologically, we offer a framework for measuring a contested political construct through three components: a fine-tuned political masked language model, a curated dictionary of probe words grounded in political theory, and a masking procedure that reformulates each occurrence of a target term as a prediction problem. Replacing each occurrence of *people* with a [MASK] token and scoring how strongly each frame’s probe words fit the masked position relative to a random-noun baseline, the framework produces a baseline-corrected frame score at the instance level. Because these scores are anchored in theory-driven dictionaries, the resulting measurements are interpretable and suited to contested concepts, whose meanings vary across contexts and historical periods in ways that topic modeling or lexical frequency counts are ill-equipped to capture.

Substantively, our analysis reveals a long-running contest among civic, ethnic, and cosmopolitan conceptions of “the people” in U.S. presidential rhetoric. This contest shifts with episodes of war, immigration, civil rights, and populist resurgence, and it tracks broader trajectories of democratic development and economic openness. Higher levels of democracy are associated with stronger Civic framing and weaker Nation and Global framing, consistent with a shift toward civic-institutional conceptions of peoplehood. Although

our analysis cannot establish causality, these associations suggest that presidential rhetoric can both reflect and constitute changes in political membership. How political elites define “the people”, explicitly and implicitly, remains a central and evolving question in politics, one that continues to shape the boundaries of political community and the distribution of rights and recognition.

More broadly, the framework can generalize beyond “the people”. The same combination of a fine-tuned masked language model, theory-guided dictionaries, and masked prediction can be applied to other foundational categories in political life, such as “democracy”, “freedom”, and “security”, whose meanings are equally contested and historically variable. The framework enables the scalable, longitudinal study of how core political concepts are reshaped over time in dialogue with the qualitative scholarship.

Limitations

Our study has several limitations. First, the frame score relies on a curated dictionary of category words, so the resulting scores could in principle be sensitive to the choice of these words. We address this concern in two ways. We validate the dictionary through embedding-based similarity search and cosine similarity analysis (Section 3.3), and we show that the scores are stable across dictionary sizes from the top-2 words onward (Appendix A.1). Nonetheless, the categories remain theoretically motivated constructs, and substantially different word selections could yield different results.

Second, the corpus is limited to textual records of presidential speeches, which are subject to selection biases in archiving and transcription, particularly for the earlier centuries. Earlier speeches are fewer in number and may overrepresent certain rhetorical occasions such as inaugural addresses, potentially skewing the temporal estimates for that period.

Finally, the regression analyses are associative rather than causal. While the co-variation between frame scores and macro-political indicators is theoretically interpretable, confounding factors cannot be ruled out, and the direction of influence between presidential rhetoric and broader social conditions remains an open question.

Ethical Considerations

This study analyzes publicly available presidential speeches sourced from the American Presidency Project and the Miller Center. All texts are in the public domain, and no private or personally identifiable information is collected or processed.

The framing scores reported in this paper are descriptive and associative in nature. They are not intended to evaluate, rank, or judge the political legitimacy of any president, party, or administration, and should not be used for such purposes. We caution against causal or normative interpretations of the observed historical trends.

The automated nature of the measurement framework means that outputs reflect patterns in language model pretraining data as much as historical rhetoric itself. Downstream users of this framework should exercise caution when drawing conclusions about the intentions or ideologies of individual political actors based solely on model-derived scores. We recommend that quantitative outputs always be interpreted in conjunction with qualitative historical and political expertise.

Acknowledgments

This work was supported by the National Research Foundation of Korea (RS-2022-NR068758; 2024S1A5C3A02042671) and by the Institute for AI and Social Innovation at Yonsei University. We thank Eunsam Cho and Sungkyu Park for helpful comments.

References

- Benedict Anderson. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, revised edition. Verso, London.
- Kwame Anthony Appiah. 2018. *The Lies That Bind: Rethinking Identity, Creed, Country, Color, Class, Culture*. Liveright Publishing Corporation, New York.
- Vanessa B Beasley. 2011. *You, the people: American national identity in presidential rhetoric*, volume 10. Texas A&M University Press.
- Bart Bonikowski, Yuval Feinstein, and Sean Bock. 2021. The partisan sorting of “America”: How nationalist cleavages shaped the 2016 US presidential election. *American Journal of Sociology*, 127(2):492–561.
- Rogers Brubaker. 1992. *Citizenship and Nationhood in France and Germany*. Harvard University Press, Cambridge, MA.
- Dallas Card, Serina Chang, Chris Becker, Julia Mendelsohn, Rob Voigt, Leah Boustan, Ran Abramitzky, and Dan Jurafsky. 2022. Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration. *Proceedings of the National Academy of Sciences*, 119(31):e2120510119.
- Boaventura de Sousa Santos. 2007. Beyond abyssal thinking: From global lines to ecologies of knowledges. *Review (Fernand Braudel Center)*, 30(1):45–89.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186.
- Deborah Dore, Elena Cabrio, and Serena Villata. 2025. Roosevelt: A new deal for political language modelling. *arXiv preprint arXiv:2508.03250*.
- Gary Gerstle. 2017. *American crucible: Race and nation in the twentieth century*. Princeton University Press.
- Sarah Harrison and Michael Bruter. 2011. *Mapping Extreme Right Ideology: An Empirical Geography of the European Extreme Right*, 1 edition. Palgrave Macmillan, London.
- John Higham. 2002. *Strangers in the land: Patterns of American nativism, 1860-1925*. Rutgers University Press.
- Jan Germen Janmaat. 2006. Popular conceptions of nationhood in old and new European member states: Partial support for the ethnic-civic framework. *Ethnic and Racial Studies*, 29(1):50–78.
- Cyril Jayet. 2012. The ethnic-civic dichotomy and the explanation of national self-understanding. *European Journal of Sociology / Archives Européennes de Sociologie*, 53(1):65–95.
- Hans Kohn. 1932. *Nationalism and Imperialism in the Hither East*. G. Routledge and Sons, London.
- Hans Kohn. 1934. *Orient and Occident*. John Day Company, New York.
- Hans Kohn. 1944. *The Idea of Nationalism: A Study in Its Origins and Background*. Macmillan, New York.
- Pål Kolstø. 2000. *Political Construction Sites: Nation-Building in Russia and the Post-Soviet States*. Westview Press, Boulder, CO.
- Stephen J. Larin. 2020. Is it really about values? Civic nationalism and migrant integration. *Journal of Ethnic and Migration Studies*, 46(1):127–141.

- Marcel Lubbers and Marcel Coenders. 2017. [Nationalistic attitudes and voting for the radical right in Europe](#). *European Union Politics*, 18(1):98–118.
- Leland McInnes, John Healy, and James Melville. 2018. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.
- Friedrich Meinecke. 1970. *Cosmopolitanism and the National State*. Princeton University Press, Princeton, NJ. Originally published as *Weltbürgertum und Nationalstaat*, 1907. Translated by Robert B. Kimber.
- Harris Mylonas and Maya Tudor. 2021. [Nationalism: What we know and what we still need to know](#). *Annual Review of Political Science*, 24:109–132.
- Martha C. Nussbaum. 1996. Patriotism and cosmopolitanism. In Joshua Cohen, editor, *For Love of Country: Debating the Limits of Patriotism*, pages 2–17. Beacon Press, Boston.
- Tim Reesens and Marc Hooghe. 2010. [Beyond the civic–ethnic dichotomy: Investigating the structure of citizenship concepts across thirty-three countries](#). *Nations and Nationalism*, 16(4):579–597.
- Robert C Rowland. 2021. *The rhetoric of Donald Trump: Nationalist populism and American democracy*. University Press of Kansas.
- Alix Rule, Jean-Philippe Cointet, and Peter S Bearman. 2015. Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014. *Proceedings of the National Academy of Sciences*, 112(35):10837–10844.
- Julian Salazar, Davis Liang, Toan Q Nguyen, and Katrin Kirchhoff. 2020. Masked language model scoring. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, pages 2699–2712.
- Vivien A Schmidt. 2008. Discursive institutionalism: The explanatory power of ideas and discourse. *Annual Review of Political Science*, 11(1):303–326.
- Rogers M Smith. 2003. *Stories of peoplehood: The politics and morals of political membership*. Cambridge University Press.
- Mary E Stuckey. 2023. *Defining Americans: The presidency and national identity*. University Press of Kansas.
- Yael Tamir. 2019. *Why Nationalism*. Princeton University Press, Princeton, NJ.
- Meghan Tinsley. 2019. [Decolonizing the civic/ethnic binary](#). *Current Sociology*, 67(3):347–364.
- John Zaller. 1992. *The nature and origins of mass opinion*. Cambridge university press.

A Appendix

A.1 Dictionary-Size Sensitivity Analysis

Figure 4 reports window-level Pearson correlations between framing scores computed under top- k aggregation for $k \in \{1, 2, 3, 4, 5, 6, 7, 8\}$ in each category. Scores stabilize from $k = 2$ onward, confirming that aggregating a few top words does not meaningfully affect the substantive results. We selected $k = 3$ for the main analysis (highlighted with black borders). Its correlation with $k \in \{4, 5, 6, 7, 8\}$ remains high across all three categories (Nation: $r = 0.92$ – 0.99 ; Civic: $r = 0.95$ – 0.99 ; Global: $r = 0.92$ – 0.99), and agreement with $k = 2$ exceeds $r = 0.98$. Only $k = 1$ shows substantially lower agreement (Nation: $r = 0.94$; Civic: $r = 0.87$; Global: $r = 0.88$), reflecting the rank instability of relying on a single probe word across contexts.

A.2 UMAP Projection of Speech Embeddings

Figure 5 presents the UMAP projection of all speech windows containing *people*. Each window is encoded into a fixed-length vector via the [CLS] token of RoBERTa and projected into two dimensions using UMAP (McInnes et al., 2018) with `n_neighbors` set to 15. The map shows broad era-level separation, with Nation framing concentrated among earlier administrations and Civic framing more prominent in the modern era.

A.3 Year-level Regression Results

Table 3 reports the regression models re-estimated at the year level, aggregating instance-level frame scores to annual means. The long-term association between the Liberal Democracy Index and Nation framing remains negative and significant, mirroring the instance-level results in Table 2.



Figure 4: Window-level Pearson correlation matrices across top- k choices for Nation, Civic, and Global framing categories. Black borders indicate the $k = 3$ row and column selected for the main analysis.

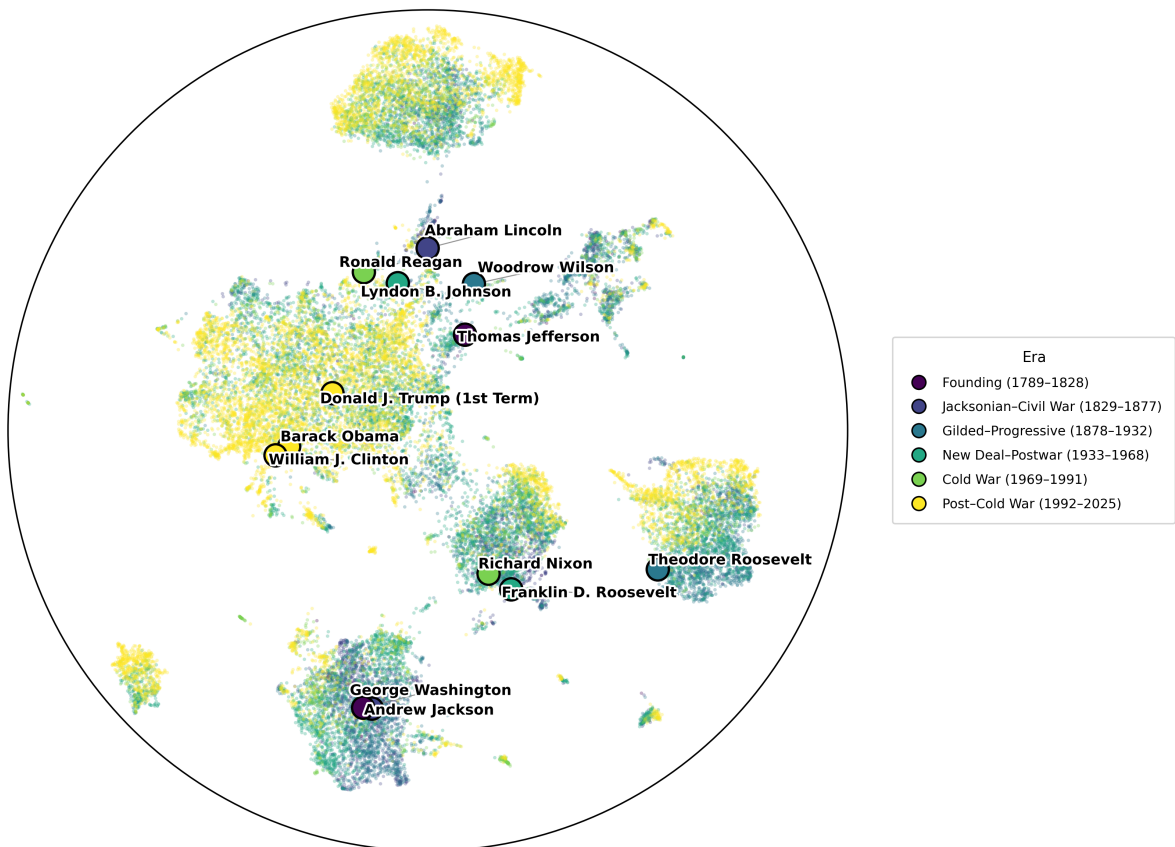


Figure 5: UMAP projection of contextual embeddings of "the people" from RooseBERT across all presidential speeches. Each dot represents the contextual window around one occurrence of "the people". Outlined markers denote the representative centroid of selected presidents, colored by historical era. The windows span a range of contexts structured by both era and speaker: individual presidents occupy distinct regions of the space, and each era spreads across multiple regions rather than forming a single block.

Table 3: **Year-level OLS regression results (robustness check), with standard errors in parentheses.** Frame scores are aggregated to annual means; *Long* and *Modern* denote the long-term ($N = 231$) and modern ($N = 53$) models. The democracy index is the V-Dem Liberal Democracy Index.

	Nation		Civic		Global	
	Long	Modern	Long	Modern	Long	Modern
Liberal Democracy	-2.385*** (0.573)	1.251 (1.896)	1.665 (1.530)	1.988 (3.331)	-1.673 (1.056)	-0.378 (2.177)
Pre-1960	-0.114 (0.243)		0.476 (0.475)		0.601 (0.487)	
GDP (logged)		-1.635* (0.845)		-1.353* (0.799)		-1.199 (0.755)
Trade Openness		0.012 (0.046)		0.072* (0.044)		-0.002 (0.046)
Net Migration (logged)		-0.279 (0.419)		0.837*** (0.223)		0.122 (0.458)
Intercept	5.104*** (0.447)	9.455*** (3.429)	1.999* (1.133)	1.314 (2.993)	0.250 (0.860)	3.130 (3.904)
N	231	53	231	53	231	53
R^2	0.261	0.266	0.021	0.309	0.459	0.244

* $p < .1$; ** $p < .05$; *** $p < .01$