Characterizing Religious Rhetoric in the U.S. Congressional Record

Lavinia Dunagan[1] , and Dallas Card[1]

[1] School of Information, University of Michigan, Ann Arbor, United States

Abstract

Christianity has historically been a potent force in American culture. In order to understand the place of religious expression in an increasingly secular and polarized society, we study the prevalence of religious rhetoric in the Congressional Record over the past three decades. We capture religious rhetoric through two distinct approaches: counting mentions of religious terms and identifying exact or approximate quotations of Bible verses using a combination of textual overlap and verse embeddings. While members of both parties routinely mention God and occasionally reference scripture, Republicans have used these signals more and more over the course of the last two decades to evoke a distinct cultural identity in addition to justifying normative commitments. Despite a pronounced decline in overall religiosity in the United States over the past fifty years, religious rhetoric remains a powerful yet flexible way for legislators, especially conservatives, to negotiate a highly partisan discursive environment while connecting their politics to a broader historical tradition.

Keywords: natural language processing, religious studies, religion in politics, text reuse

1   Introduction

In the United States, the centrality of religious identity to the national political culture is evident in the long history of religiously inflected oratory. Despite a tradition of secular government, religious rhetoric is a key component of the American state's sacred texts [28], from the Declaration of Independence to the modern Pledge of Allegiance. Whether concluding a speech with "God bless America" or using a verse from the Bible to bolster a point, politicians in the United States have engaged with Christian discourses since the origin of the country [7]. In the past half-century, however, this aspect of American political identity has been reconfigured by gradual secularization and the rise of the religious right [36]. Even as many segments of American society are reexamining religion, one may still perceive echoes of ancient scripture in political discourse. Prior work on this subject has focused on presidential speeches, but ignored more everyday expressions of religious ideas in political contexts. To fill this gap, we carry out the first large scale computational analysis of where, when, and how religious speech becomes enshrined in the official proceedings of the U.S. Congress.

Using the Congressional Record (CR) from the past three decades—comprising over one million speeches in the House and Senate—we identify major shifts in the partisan valence of religious speech, from relatively equal use by Democrats and Republicans, to a striking divergence over the past decade. In particular, we make use of lexical markers and recognizable references to the Bible, found via a combination of embedding- and token-based similarity metrics, to measure the use of religious rhetoric over time and across speakers. Doing so reveals sparse but increasingly common religious expression in Congress. Our results also suggest that Republicans emphasize a

potent mixture of cultural, religious, and military values while advancing their vision of America. We make all code necessary to replicate our experiments available at the link below.[1]

2 Background

As in many countries, religion occupies a contested space in American politics. Robert Bellah [7] famously observed that American political culture is defined by a quasi-Christian set of norms and rituals. The U.S. is unusually religious in comparison with other developed nations [37; 44], with about 96% of members of the 119th Congress (2025–27) having some religious affiliation [15]. At the same time, the American electorate has steadily become less religious [19; 41], illustrated by a thirteen-point drop in Christian identification in the United States from 2007 to 2019 (78.4% to 65%) [33]. Secularization, referring both to a cultural shift towards "rationalization-disenchantment" and to a pattern of declining religious belief and behavior, has reshaped norms surrounding religious speech, at least outside political contexts [32; 45].

Modern scholarship examining religious rhetoric in American political speech has focused primarily on Presidential communications (e.g., [18; 24; 40]), finding that presidents have used less explicitly Christian speech over time [13], but with candidates using religious rhetoric more when campaigning in religious areas [12]. The one relevant prior study on Congress examined the use of religious speech by members of Congress on Twitter during 2018 [9]. Using a handful of religious keywords, the authors found that religious speech in this context was rare overall, but more frequently used by Republicans, men, and during times of celebration or crisis. We turn our attention to the Congressional Record itself because of its unique status as a long-running, totemic, and highly multivocal reflection of political discourse in the United States.

The role religion plays in this context also recalls the work of Benedict Anderson on nationalism. He emphasizes the importance of *shared texts*, such as newspapers and popular novels, in the creation of a canonical national culture [3]. The prominence of the Bible in America, both in political rhetoric and in newspapers themselves [34], reflects that these shared texts need not be generated by the national context itself. Although national identity is cast in opposition to preexisting religious modes of social identification by Anderson and contemporaries, elaborations on that work by other scholars of nationalism have emphasized the continuing significance of religious identity in many countries [48].

Numerous religious traditions are present in American culture; however, we focus primarily on markers of Christian religious expression. While excluding other religions limits our ability to understand their presence in American political discourse, we focus on Christianity because of both its historical and contemporary importance to U.S. politics and its overall prominence. Among those expressing some religious affiliation, approximately 91% of Members of Congress in 2025 and 88% of the broader U.S. population identify with a Christian tradition [15; 41].[2]

Our work draws on both social scientific and humanistic accounts of religious speech in political settings. As we explore throughout this paper, large-scale cultural change in the United States has not resulted in a proportional reduction in religious expression in a body that is idealized as nationally representative. Refracted through the interpretive frames of partisan politics, the use of religious symbols is loaded with a tension between universalism and those symbols' growing association with particular identities.

---

[1] `https://github.com/laviniakd/religion_in_congress/`

[2] There is also some overlap of Christian language with that of the next most prominent religions in the U.S. (Judaism and Islam).

3 Data

The Congressional Record We focus on the Congressional Record (CR)—the official proceedings of all speeches given in the U.S. House and Senate. Since our focus is on recent trends, including the emergence of the present partisan divide around religion, we limit our study to the past three decades (since 1995). The CR reflects changing political discourses in the United States—influenced by the emergence of new media and rising polarization—and has thus been previously used as a data source for computational analysis [11; 21; 25]. Although some political scientists dismiss communications in Congress as a low-impact form of speech [16; 35], we nevertheless treat it as a culturally significant archive that captures the expressed opinions of members of the legislative branch. For additional details, please refer to Appendix B.

Sermon Central Our reference for Christian religious speech in America is a dataset first presented in [8] that consists of approximately 140,000 sermons from an online sermon sharing platform named Sermon Central. It is predominantly used by Protestant preachers (at least 1,500 Catholic homilies are also included).[3] Although we cannot capture religious language in the U.S. via a single corpus, Sermon Central reflects a variety of language associated with American Christianity. Since members of recent Congresses have been majority Protestant, and at least 87% Christian [15], we posit that Sermon Central will reflect most of the religious discourse they use.

The King James Bible For the purpose of identifying references to the Bible in CR, we use the King James version, as the most influential version within the English-speaking world [6; 10]. It retains the coupled grammatical structures of the source texts [4], while its Elizabethan style has entered popular culture as that of Biblical speech [17].

4 Measuring Religious Language in Congress

We begin our analysis by measuring the prevalence of religious language in the CR. In order to do so, we construct a keyword list reflecting terms that make up contemporary Christian speech in the United States. Although this may not capture more subtle references, it is highly interpretable and efficient; it also ensures we capture the most salient signals of religious rhetoric. We aim to broadly identify who uses Christian terms, how this has changed over time, and what signals are most widely used.

Methods To construct a religious lexicon, we begin by comparing the CR to sermons from Sermon Central. In particular, we use the prior-modified log odds ratio [26], to identify words that are used much more frequently than would be expected in one compared to another, with reference to a background corpus (details in Appendix C). For a background corpus, we use the Corpus of Contemporary American English (COCA; [14]). In constructing this list, we preserved capitalization, rather than lowercasing, as religious language sometimes distinguishes meaningfully between upper- and lowercase versions of words (e.g., "Him" vs. "him"). From among the resulting terms, we filtered out stopwords and verbs; we also excluded all short tokens (< 3 characters), and those that are used less than five times in either corpus.
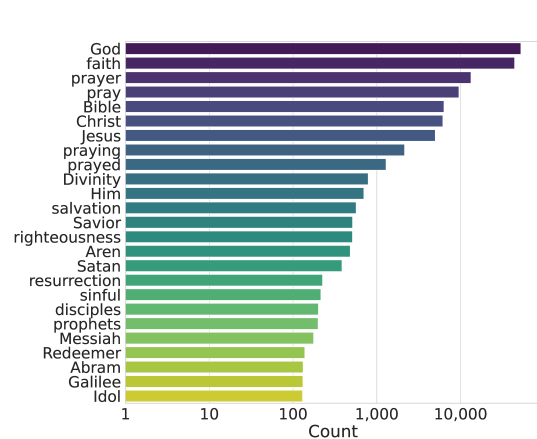
This results in a list that includes both culturally salient terminology (e.g., "sinners") and terms referring to entities in the Bible, such as "Philippian". We augment the list with one-word synonyms of "God" and "Jesus" using WordNet [20], since their relative frequency in everyday American English means they are not found by the log-odds method. We also include all words from the only other paper to study religious speech by members of Congress [9]. We keep the top 150 words
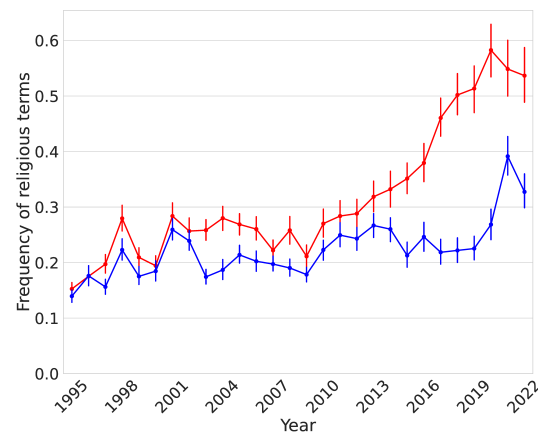
---

[3] https://www.sermoncentral.com/

from the log-odds method plus the 15 from the augmentation, with the full list given in Appendix C.

To quantify the degree of religious signaling in speech, we use two measures: i) the proportion of words in a Congressional speech that appear on our keyword list, and ii) a binary measure of whether or not a given speech contains any religious keywords. We are interested both in overall levels of religious lexicality in a given speech and whether any religious words are used; the first is a signal of a speech's focus on religion, while the second captures casual religious references which may not be related to the narrative character of a speech.

Figure 1: Counts of the 25 most common religious keywords in the Congressional Record. "God" is the most commonly used term in our lexicon, followed by "faith" and variants of "pray".
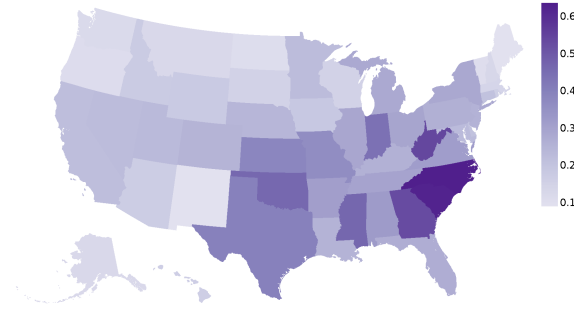

Figure 2: Average rates of religious keywords per 1000 words in the Congressional Record over time, along with standard errors. Republicans have used more religious language than Democrats recently, but both parties have increased their usage over time.

Results First examining the prominence of each of the words in our lexicon, we find that the distribution words used in the CR is extremely skewed (as shown in Figure 1). "God" is the most commonly used word in our lexicon, being mentioned around 54,000 times in the Congressional Record from 1995 to 2022; this is discussed further below. Out of the approximately 1,614 speakers represented in the preprocessed CR, only 33% of them *never* use the word God. Both Christian and nonspecific religious words are among the most prevalent terms; "Bible", "Jesus", and "Christ" each appear thousands of times.

Notably, we see a significant divergence between Democrats and Republicans since around 2001 (see Figure 2). The parties used religious terms at approximately the same rate during the 1990s. Since then, however, Republicans have used religious language with greater frequency than Democrats in every year since 2000, dramatically so since 2010. Moreover, the frequency with which Republicans use religious terms has increased three-fold since 1995. Nevertheless, the occasional use of religious terms is bipartisan. Within our sample, 72% of Democrats who speak in Congress have used at least one religious keyword, while about 79% of Republicans have. Partisan differences in both keyword rate and inclusion of at least one religious keyword in a speech are both significant per a two-tailed *t*-test and a χ^2 test ($p < 0.0001$ in both cases).

As shown in Figure 3, religious language in Congress is also to some extent a *regional* phenomenon. Although there is considerable variation even between neighboring states (with outliers in both the House and Senate), religious keywords are more frequent in speeches by members of Congress representing the South. At the extremes, religious keywords are more than six times as common in speeches by legislators representing the Carolinas and Georgia compared to Maine and

Figure 3: Average rate of religious keywords usage by state (per thousand words), showing a clear regional trend, with more frequent mentions in the U.S. South.

Vermont. This likely reflects both party dominance within states and variation in overall religiosity across states [30]. Religious language in U.S. politics is still broadly national, in that religious terms are used by speakers from every state, but the variation across states emphasizes the cultural differences that underlie certain coalitions and communities.

These results reflect both the polyvalent nature of religious language and its relationship in the United States to reaffirmations of a national community. Even though past work has identified a much earlier start to politically inflected cultural sorting in the U.S. [21], our data suggest that both parties used religious language at similar rates until 2000. The performance of partisan identity did not necessarily entail a religious component until very recently. Moreover, despite broad social trends toward secularization, we see both parties *increasing* their use of religious language over time.
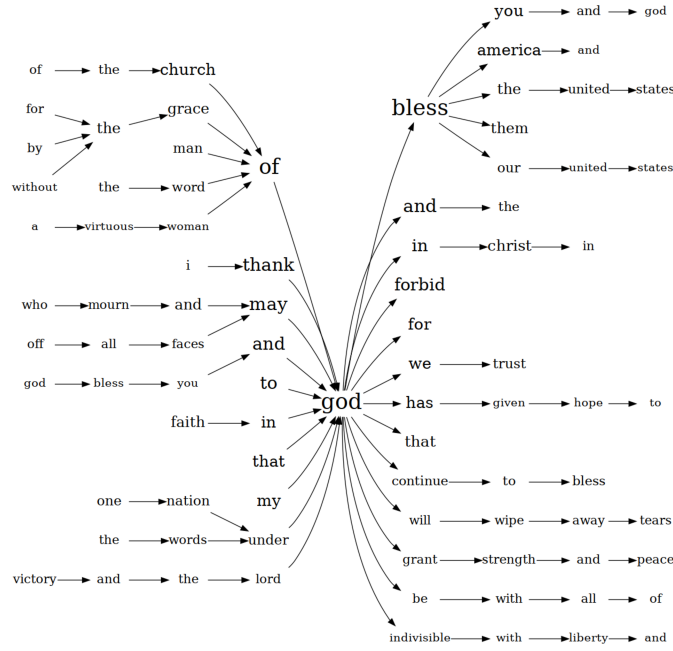
Mentions of God Given the extent to which the word "God" is present in the CR, we also briefly investigate how this term is used in the corpus. Compared to explicitly Christian terms, "God" is relevant to a much broader space of religious beliefs. At the same time, [46] note that "God" and related terms have undergone semantic bleaching in interjective settings (as in "oh my god!"). In order to understand how the term "God" is being used in Congress, we explore its usages in context. We collect all 5-grams which include the word "God" and gather together the most common phrases or subphrases. To understand the overall contexts in which the term "God" is mentioned in Congress, we use the `wordtree` algorithm to visualize the most frequent n-grams containing it [50].[4] Figure 4 shows the most common overlapping fragments of phrases that involve the word "God". Size indicates frequency, scaled using a cube root function, to accentuate differences. As shown, "thank" is the most common verb taking God as an object, and "bless" is the most common verb with God as the subject.

Many other phrases are also common, including "church of God", "faith in God", and "God forbid". The most commonly used phrase comes from the Pledge of Allegiance—"one nation under God, indivisible, with liberty and justice for all"—representing around 7% of mentions. Although many mentions of God in Congress are relatively standardized, as in phrases like "God bless America", such pro forma mentions do not capture the entirety of uses. Indeed, of the approximately 30,000 distinct five-grams centered on "God", almost 25,000 (83%) only occur once.

5   Quoting the Bible

The above analyses focused on a purely lexical method for identifying religious speech, which does not adequately distinguish between pro forma mentions, generic expressions of religiosity, and

---

[4] https://github.com/willcrichton/wordtree

Figure 4: A wordtree visualization showing the many occurrences of "God" in our dataset, with size reflecting counts of subphrases.

more explicitly Christian rhetoric. It also fails to capture the deeply textual character of Christianity [27]. In order to detect more meaningful uses of religious ideas, we develop a method to identify exact or approximate quotes from the Bible. References to the Bible are a natural target of analysis because they index both the ideas explicated by a given verse and familiarity with the text itself; the Bible is also shared across Christian denominations.

Methods   Various methods have been proposed to detect text reuse and reference, including those based on n-gram overlap [39], TF-IDF representations [34], and contextual embeddings [47]. Here, to identify whether a given sentence is a reference to a specific Bible verse, we combine two methods—one which uses cosine similarities of contextual embeddings, and another which measures the degree of overlap of their constitutive n-grams. Even though it would not be a scalable solution for our full corpus, we also experiment with prompting language models to identify references to the Bible, and consider using this to filter out false positives.

For each sentence in the CR (excluding those shorter than five tokens), we compare it to the 1,000 most frequently referenced Bible verses in the Sermon Central dataset, according to two metrics. The first metric computes the cosine similarity between vectors after embedding both the CR sentence and a Bible verse using MPNet [43]. We choose MPNet because it was trained to produce similar representations of semantically equivalent paraphrases and is thus well-suited for our goal of finding both paraphrases and different translations of Bible verses, as well as verbatim quotes. The second metric uses n-gram shingling, computing the number of unique n-grams that are common between a CR sentence and each target Bible verse, with $n = 5$.[5] A sentence is considered a reference to a particular Bible verse if it meets either the cosine similarity threshold with some verse or the shared n-gram threshold, and it has the highest similarity score across all target Bible verses.

---

[5] Although the shingling metric is sometimes normalized by the union of n-grams between the two texts, here we simply take the intersection, per [39].

In order to tune the thresholds for our method, we first scored sentences in the CR according to the embedding metric. We then sampled instances from each embedding metric decile, along with their highest-scoring Bible verse matches, and labeled whether the match was correct. This yielded 150 ground-truth triplets of sentence, label, and verse. These matches were annotated independently by the authors of this paper, resulting in an inter-annotator agreement of Cohen's $\kappa = 0.98$. Based on these annotations, we chose a cosine similarity threshold of ≥ 0.75 and shingling threshold of at least five 5-grams in common, as the combination that resulted in the best performance on the validation data ($F_1 = 0.91$, weighted by class).
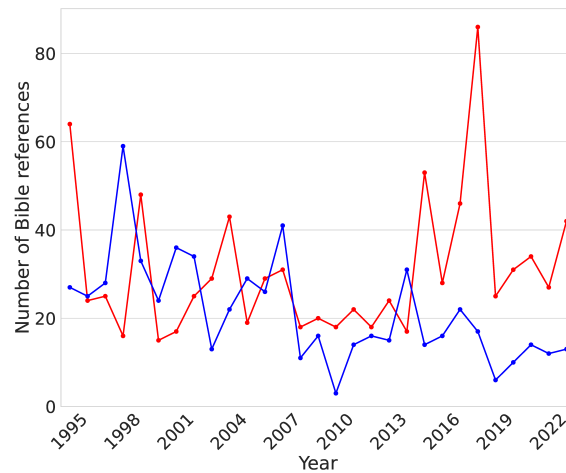
To produce a final performance estimate for our method, we collect an annotated test sample of 136 instances in the same stratified manner as above. On this, the method combining cosine similarity and n-gram overlap attained a weighted $F_1 = 0.81$ (with weighted precision of 0.87 and weighted recall of 0.84). Although the best performing prompting method (using Claude 3.7 Sonnet with a one-shot example) achieved strong results on the validation data (weighted F_1 of 0.98), it was ultimately not helpful with identifying false positives. Only a small number of matches from our method were identified as false positives by the prompting approach, and manual inspection revealed that these were overwhelmingly correctly matched (see Appendix D for full details). Prompting-based approaches could in principle lead to higher recall, but the size of the CR dataset makes this infeasible. As such, our final set of Biblical references is derived without the help of language models.

Results　Using the method above, we identify 1,555 total references to 279 different Bible verses in the CR, occurring in 898 speeches by 355 speakers (or 0.04% of all non-procedural speeches).[6] Overall, references to the Bible are rare and tend to be produced in multiples by the same speaker, sometimes within the same speech. The distribution of the verses themselves is also extremely skewed (see Appendix Figure 11). Of the small set of verses referred to many times, most are from the New Testament and several are from the Gospel of John, a book of the Bible which is frequently emphasized in American Christianity [31].
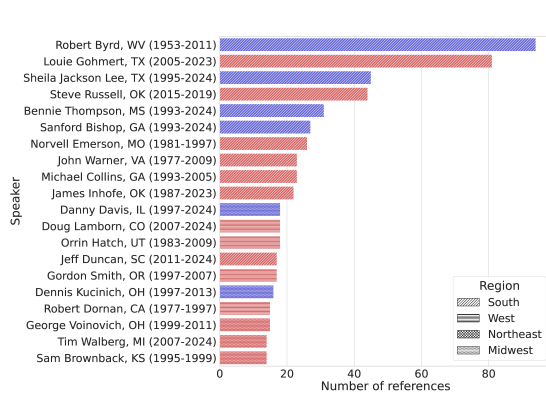
As with religious keywords, Republicans have increased their use of Bible quotations over time, with a difference between parties emerging in the past decade (see Figure 5). In fact, the divergence in the two parties' use of the Bible is even greater than that of religious terms. Democrats have actually made less frequent references to the Bible over time ($\rho = -0.62$, $p < 0.001$), in step with the broader pattern of secularization, while Republicans have relied more and more on the moral authority attributed to Scripture. Among Republicans, 2018 is a notable outlier. Manual inspection revealed that many verse usages were prompted by the death of a famous evangelical preacher, Billy Graham. This speaks to a common mode of Biblical citation—to memorialize or otherwise honor an individual (see Qualitative Examination below).

We also find several long-serving members of Congress from both parties who are notable users of Biblical references. As shown in Figure 6, there are six Democrats and fourteen Republicans among the twenty most prolific users of the Bible; however, there is once again a prominent regional pattern in who cites Scripture most frequently. Of the speakers shown, most are from the South or Midwest, including all six of the Democrats.
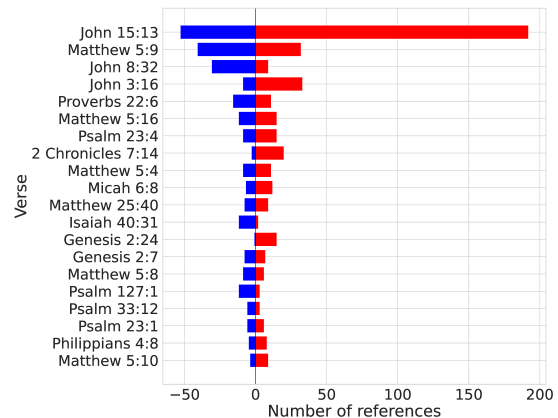
---

[6] For the most frequently cited verses, and a random sample of matching CR sentences, please refer to Appendix D.

Figure 5: Number of identified references to the Bible per year, split by party. As with religious keywords, we see a pattern of polarization between the parties in the past decade, with Democrats making less frequent references to the Bible over time.

Figure 6: The 20 most prolific users of exact or approximate quotations from the Bible in our dataset.

Figure 7: Counts of Republican and Democratic uses of the 20 most popular verses overall.

John 15:13 is by far the most popular Bible verse, which in the King James Version reads: "Greater love hath no man than this, that a man lay down his life for his friends." As one might expect, manual inspection reveals that this verse features prominently in discussion of the military, as a means of affirming the virtue of both military service and the military as an institution. The Bible's utility for addressing this issue is elucidated further by the *second* most popular reference: Matthew 5:9. It reads: "Blessed are the peacemakers: for they shall be called the children of God." This verse also lends itself to endorsement of certain beliefs about foreign policy, but it is usually used to affirm diplomacy and nonviolent resolutions to conflict. Republicans are *far* more likely to refer to John 15:13 than Democrats (see Figure 7), reflecting its more hawkish implications.[7] This suggests that even memorializing speeches about specific constituents—e.g., those who serve in the military—have some partisan valence. As Benedict Anderson recognized, modern national identity as a *cultural* object tends to involve an ethic of sacrifice which explicitly glorifies the

---

[7] Although many verses show an association with party, based on a binomial test ($p < 0.05$), only John 15:13 remains significant after applying a Bonferroni correction for the 279 verses mentioned in the CR.

military.

Although not significant after correcting for multiple comparison, other verse usage patterns suggest partisan divides on domestic issues. Genesis 2:24 is referenced 15 times by Republicans and reads, "Therefore shall a man leave his father and his mother, and shall cleave unto his wife: and they shall be one flesh." This verse is frequently used to justify a conception of marriage as an exclusively heterosexual institution [49]. Indeed, the only Democrat who used this verse is Robert Byrd, a conservative West Virginia senator who co-sponsored the Defense of Marriage Act.

Qualitative examination of Bible quotes   Given the sparser nature of actual quotations of Scripture, we can understand Bible use in terms of rhetorical strategies rather than merely keywords; drawing on previous work [42], we examine the *proximal contexts* of references to the Bible—not the entire speech, but rather the immediate environment in which a quote is being used. We begin with a random set of 32 verified references to the Bible from the corpus.

A typology emerges from a close examination of these speeches. First, representatives frequently commemorate individuals by speaking their achievements into the Congressional Record. These individuals range from deceased constituents to public figures with national profiles. The second type of use entails references to the Bible in the context of celebrations of Christianity like the National Prayer Breakfast. The third and arguably most important use of Bible verses is constituted by references to Scripture as authoritative with respect to a policy issue. Verses are a flexible means of articulating an underlying value and its relationship to the speaker's argument while presupposing the universality of the value. For example, Rick Allen (R-TX) described Exodus in a 2018 speech more broadly celebrating National Bible Week: "God created the law, the Ten Commandments... Government was instituted by God for one purpose: to restrain evil and promote good" (1 Peter 2:13-14).

The flexibility of the Bible is sometimes evident in the use of verses to support diametrically opposed positions on a given issue. Roscoe Bartlett (R-MD) referred to the Bible verses used in the New England Primer (a popular educational booklet during the colonial era) in order to support the use of religious language, including prayer, in schools. Meanwhile, Bob Wise (D-WV) used Matthew 6:6 ("and when thou prayest, pray to thy father in private and he shall hear you") in 1998 in order to advocate against a resolution that would endorse prayer in schools.

The inclusive mode's use of Bible verses varies, from more universalizing strategies to such phrases as "in the Christian faith" (deployed by Cory Booker, D-NJ, in a 2021 speech memorializing a police officer who died as a consequence of January 6th); the latter qualifier signals the specificity of the words on which Booker draws. In contrast, the universality claimed by Republican speakers is usually being deployed in order to assert the essential truth of Christian teachings. As Roscoe Barlett put it, "The sacred rights of mankind... are written as with a sunbeam in the whole volume of human nature by the hands of the divinity itself and can never be erased or obscured by mortal power." We can understand references to the Bible in our corpus as tied to a desire, irrespective of party, to establish links between fleeting legislative speech and the American project's place in history. As Shalev [38] observed in the texts of the early United States, a Biblical style is a means of conveying such timelessness; it is also a way for insurgent political currents, especially ones that claim to hearken back to antiquity, to ground themselves in recognizable motifs.

6 Discussion

Religious rhetoric shows no sign of disappearing from the American political scene. At one time in the not so distant past, it was primarily a nonpartisan means of connecting everyday American political life to grand historical concerns. American civil religion entailed a quiet pluralism—one

that did not explicitly assume the correctness of one religious tradition [7]. This mode of politics uses the Bible more as a common reference, like the mass print media Benedict Anderson sees as essential for nationalism, than as an instruction manual, and Christian rhetoric more as a flourish than as an appeal to identity. Nonetheless, numerous other approaches also characterize religious expression in this sphere. Some conservatives have historically referred to the Bible in order to justify restrictive social norms [22; 23]. For these political actors, Christianity is embedded in a very particular set of goals for government; Christian language is a means of communicating a blend of nationalism and political alignment with other Christians [36].

On a methodological note, the techniques used in this paper are reflective of a philosophy of triangulation of different quantitative approaches in order to approximate qualitative insight at scale. This paper entails the description of a multifaceted phenomenon—religious discourse in American politics—which operates through explicit lexical signals, implicit motifs, and textual references that can be either obvious or subtle. We thus chose to use multiple methods, rather than one, to operationalize these different linguistic practices. As we discuss in Appendix A, our methods are imperfect. Nevertheless, our findings indicate real change in the status of religious speech in the public sphere.

Both the Congressional Record and the Bible are important textual landmarks in American political culture. Rather than tracking the observed decline in religiosity among Americans, we find that religious terms are *more prevalent over time* in Congressional speeches by both parties, although the effect is much greater for Republicans. The use of the Bible is more clearly differentiated by ideology; since Obama's inauguration, Republicans have continued to refer to it steadily, while Democrats have begun to do less frequently. A universalist mode of Biblical reference is residually present in our corpus, as both Republicans and Democrats use the Bible to articulate moral foundations for policy, but Republicans also use the Bible to argue for specifically conservative beliefs.

We hypothesize that this arises from broader cultural trends in the United States: polarization and aforementioned secularization. Religiosity is more associated with partisanship than ever; this has produced a cultural environment in which Christian cultural markers are progressively more enmeshed with conservative political ideology [36]. In other words, the meaning of the Bible in political contexts has been reconfigured by changes in American society. In the Congressional Record—a corpus which arguably functions as *the* massively multivocal symbol of political discourse in America—we find evidence of political actors expressing different and sometimes opposing positions via a text whose long history and claim to cosmic relevance coexist with evolving interpretations and cultural connotations. Future work on this subject may deploy similar methods to explore other types of intertextual discourse, religious or not, surrounding authoritative texts like the Constitution in Congress and beyond.

7   Conclusion

The preceding analyses constitute the first large-scale exploration of religious (especially Christian) language in the U.S. Congressional Record over the past 30 years. Religious rhetoric has been examined in prior work as a significant component of American textual culture writ broadly and political discourse in particular, but its realization in the diverse and multivocal speech of the Congressional Record over a longer period has been unexplored. We establish the existence of relative similarity between Democrats and Republicans in the 1990s followed by divergence in the 2000s; Democrats followed the secularizing path of American culture while Republicans claimed Christianity as a component of their sociocultural identity. Furthermore, we find evidence of Members of Congress using the Bible to affirm the traditional relationship between it and American political discourse *and* to assert the truth of particularistic moral injunctions from Scripture. More broadly, we show that the decline of Christian belief in America has not led to a commensurate decline in

Christian language in the most hallowed halls of American politics. A religious vision of the world instead appears to have attained greater prominence in Congress as legislators articulate their place and purpose in the 21st century.

Acknowledgements

References

[1] Albertson, Bethany L. "Dog-Whistle Politics: Multivocal Communication and Religious Appeals". In: *Political Behavior 37*, no. 1 (Mar. 2015), pp. 3–26. ISSN: 0190-9320, 1573-6687. DOI: `10.1007/s11109-013-9265-x`. (Visited on 10/19/2023).

[2] Allen, Levi G. and Olson, Shayla F. "Racial Attitudes and Political Preferences Among Black and White Evangelicals". en. In: *Politics and Religion 15*, no. 4 (Dec. 2022), pp. 631–648. ISSN: 1755-0483, 1755-0491. DOI: `10.1017/S1755048322000074`.

[3] Anderson, Benedict. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso, 1983.

[4] Assis, Elie. "Chiasmus in Biblical Narrative: Rhetoric of Characterization". In: *Prooftexts 22*, no. 3 (2002), pp. 273–304. ISSN: 0272-9601. DOI: `10.2979/pft.2002.22.3.273`. (Visited on 08/23/2023).

[5] Auerbach, Erich. *Mimesis: The Representation of Reality in Western Literature - New and Expanded Edition*. Trans. by Willard Trask. REV - Revised. Princeton, NJ, USA: Princeton University Press, 2003. ISBN: 9780691160221. URL: `http://www.jstor.org/stable/j.ctt3fgz26` (visited on 09/29/2024).

[6] Bagley, Ellie G. "Writing the History of the English Bible: A Review of Recent Scholarship". en. In: *Religion Compass 5*, no. 7 (2011), pp. 300–313. ISSN: 1749-8171. DOI: `10.1111/j.1749-8171.2011.00286.x`.

[7] Bellah, Robert N. "Civil Religion in America". In: *Daedalus 96*, no. 1 (1967), pp. 1–21. ISSN: 0011-5266. (Visited on 07/25/2023).

[8] Boussalis, Constantine, Coan, Travis G., and Holman, Mirya R. "Political Speech in Religious Sermons". In: *Politics and Religion 14*, no. 2 (June 2021), pp. 241–268. ISSN: 1755-0483, 1755-0491. DOI: `10.1017/S1755048320000334`. (Visited on 08/23/2023).

[9] Bramlett, Brittany H. and Burge, Ryan P. "God Talk in a Digital Age: How Members of Congress Use Religious Language on Twitter". In: *Politics and Religion 14*, no. 2 (June 2021), pp. 316–338. ISSN: 1755-0483, 1755-0491. DOI: `10.1017/S1755048320000231`. (Visited on 08/02/2023).

[10] Byrd, James P. "America's Public Bible: Biblical Quotations in U.S. Newspapers". In: *Journal of American History 104*, no. 4 (Mar. 2018), pp. 1097–1098. ISSN: 0021-8723. DOI: `10.1093/jahist/jax556`. (Visited on 09/17/2023).

[11] Card, Dallas, Chang, Serina, Becker, Chris, Mendelsohn, Julia, Voigt, Rob, Boustan, Leah, Abramitzky, Ran, and Jurafsky, Dan. "Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration". In: *Proceedings of the National Academy of Sciences 119*, no. 31 (2022), e2120510119. DOI: `10.1073/pnas.2120510119`. URL: `https://www.pnas.org/doi/abs/10.1073/pnas.2120510119`.

[12] Coe, Kevin and Chapp, Christopher B. "Religious rhetoric meets the target audience: Narrowcasting faith in presidential elections". In: *Communication Monographs 84*, no. 1 (Jan. 2017), pp. 110–127. ISSN: 0363-7751. DOI: `10.1080/03637751.2016.1250932`.

[13] Coe, Kevin and Chenoweth, Sarah. "Presidents as Priests: Toward a Typology of Christian Discourse in the American Presidency". In: *Communication Theory 23*, no. 4 (Nov. 2013), pp. 375–394. ISSN: 1050-3293. DOI: `10.1111/comt.12020`.

[14] Davies, Mark. "The Corpus of Contemporary American English (COCA)". Online. Available online at `https://www.english-corpora.org/coca/`. 2008.

[15] Diamant, Jeff. "Faith on the Hill". en-US. Online. Jan. 2025. URL: `https://www.pewresearch.org/religion/2025/01/02/faith-on-the-hill-2025/`.

[16] Doerfler, Ryan D. "Who Cares How Congress Really Works". In: *Duke Law Journal 66*, no. 5 (2017), p. 979.

[17] Dreisbach, Daniel L. "The Bible in the Political Rhetoric of the American Founding". In: *Politics and Religion 4*, no. 3 (Dec. 2011), pp. 401–427. ISSN: 1755-0491, 1755-0483. DOI: `10.1017/S1755048311000423`. (Visited on 08/23/2023).

[18] Endy, Melvin B. "Abraham Lincoln and American Civil Religion: A Reinterpretation". In: *Church History 44*, no. 2 (1975), pp. 229–241. DOI: `10.2307/3165195`.

[19] Fahmy, Dalia. "Americans are far more religious than adults in other wealthy nations". Online. July 2018. URL: `https://www.pewresearch.org/short-reads/2018/07/31/americans-are-far-more-religious-than-adults-in-other-wealthy-nations/`.

[20] Christiane Fellbaum, edited by. *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press, 1998.

[21] Gentzkow, Matthew, Shapiro, Jesse M., and Taddy, Matt. "Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech". In: *Econometrica 87*, no. 4 (2019), pp. 1307–1340. DOI: `https://doi.org/10.3982/ECTA16566`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA16566`.

[22] Harding, Susan Friend. "The Book of Jerry Falwell: Fundamentalist Language and Politics". In: *The Book of Jerry Falwell*. Princeton, NJ, USA: Princeton University Press, June 2018. DOI: `10.1515/9780691190464`. (Visited on 10/10/2023).

[23] Hollinger, David A. "Christianity's American Fate: How Religion Became More Conservative and Society More Secular". In: *Christianity's American Fate*. Princeton, NJ, USA: Princeton University Press, Oct. 2022. DOI: `10.1515/9780691233895`. (Visited on 10/10/2023).

[24] Hughes, Ceri. "Thou art in a Deal: The Evolution of Religious Language in the Public Communications of Donald Trump". In: *International Journal of Communication 14*, no. 0 (Sept. 2020), p. 22. ISSN: 1932-8036. (Visited on 07/24/2023).

[25] Jensen, Jacob, Naidu, Suresh, Kaplan, Ethan, Wilse-Samson, Laurence, Gergen, David, Zuckerman, Michael, and Spirling, Arthur. "Political Polarization and the Dynamics of Political Language: Evidence from 130 Years of Partisan Speech". Online. 2012. URL: `http://www.jstor.org/stable/41825364` (visited on 10/02/2025).

[26] Jurafsky, Daniel and Martin, James H. "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models". In: 4th. Online manuscript released August 24, 2025. See URL: Published online, 2025. Chap. 22. Lexicons for Sentiment, Affect, and Connotation. URL: `https://web.stanford.edu/~jurafsky/slp3/22.pdf`.

[27] Juzwik, Mary M. "American Evangelical Biblicism as Literate Practice: A Critical Review". en. In: *Reading Research Quarterly 49*, no. 3 (2014), pp. 335–349. ISSN: 1936-2722. DOI: `10.1002/rrq.72`.

[28] Kalscheur, Gregory A. "Christian Scripture and American Scripture: An Instructive Analogy?" In: *Journal of Law and Religion 21*, no. 1 (2005), pp. 101–142. ISSN: 07480814. URL: `http://www.jstor.org/stable/30040559` (visited on 07/17/2025).

[29] Lienesch, Michael. "Contesting Civil Religion: Religious Responses to American Patriotic Nationalism, 1919-1929". In: *Religion and American Culture 28*, no. 1 (Dec. 2018), pp. 92–134. ISSN: 1052-1151, 1533-8568. DOI: `10.1525/rac.2018.28.1.92`. (Visited on 08/02/2023).

[30] Lipka, Michael and Wormald, Benjamin. "How religious is your state?" Online. Feb. 2016. URL: `https://www.pewresearch.org/religion/wp-content/uploads/sites/7/2016/02/how-religious-is-your-state-tables.pdf` (visited on 02/29/2016).

[31] Malina, Bruce J. and Rohrbaugh, Richard L. *Social-Science Commentary on the Gospel of John*. en. Minneapolis, MN, USA: Fortress Press, 1998. ISBN: 978-1-4514-1703-6.

[32] Marshall, Douglas A. "Behavior, Belonging, and Belief: A Theory of Ritual Practice". In: *Sociological Theory 20*, no. 3 (2002), pp. 360–380. ISSN: 07352751. URL: `http://www.jstor.org/stable/3108616` (visited on 09/29/2024).

[33] Mitchell, Travis. "In U.S., Decline of Christianity Continues at Rapid Pace". en-US. Online. Oct. 2019. URL: `https://www.pewresearch.org/religion/2019/10/17/in-u-s-decline-of-christianity-continues-at-rapid-pace/`.

[34] Mullen, Lincoln A. *America's Public Bible: A Commentary*. Redwood City, CA, USA: Stanford University Press, 2023. DOI: `10.21627/2022apb`. URL: `https://americaspublicbible.org`.

[35] Parker, Glenn R. "Can Congress Ever Be a Popular Institution?" In: *The House at Work*, ed. by Joseph Cooper and G. Calvin Mackenzie. New York, USA: University of Texas Press, 1981. Chap. 1, pp. 31–55. ISBN: 9780292771574. DOI: `doi:10.7560/730175-003`. URL: `https://doi.org/10.7560/730175-003`.

[36] Perry, Samuel L. "Mating Call, Dog Whistle, Trigger: Asymmetric Alignments, Race, and the Use of Reactionary Religious Rhetoric in American Politics". In: *Sociological Theory 41*, no. 1 (Mar. 2023), pp. 56–82. ISSN: 0735-2751. DOI: `10.1177/07352751231153664`. (Visited on 07/24/2023).

[37] Schnabel, Landon and Bock, Sean. "The Persistent and Exceptional Intensity of American Religion: A Response to Recent Research". In: *Sociological Science 4* (Nov. 2017), pp. 686–700. ISSN: 2330-6696. DOI: `10.15195/v4.a28`. (Visited on 05/05/2023).

[38] Shalev, Eran. ""Written in the Style of Antiquity": Pseudo-Biblicism and the Early American Republic, 1770-1830". In: *Church History 79*, no. 4 (2010), pp. 800–826. DOI: `10.1017/S0009640710001034`.

[39] Smith, David A., Cordell, Ryan, and Mullen, Abby. "Computational methods for uncovering reprinted texts in antebellum newspapers". In: *American Literary History 27*, no. 3 (June 2015), E1–E15. ISSN: 0896-7148. DOI: `10.1093/alh/ajv029`.

[40] Smith, Gary Scott. *Faith and the Presidency: From George Washington to George W. Bush*. Oxford, UK: Oxford University Press, Oct. 2006. DOI: `10.1093/acprof:oso/9780195300604.001.0001`. URL: `https://doi.org/10.1093/acprof:oso/9780195300604.001.0001`.

[41] Smith, Gregory A., Cooperman, Alan, Alper, Becka A., Mohamed, Besheer, Rotolo, Chip, Tevington, Patricia, Nortey, Justin, Kallo, Asta, Diamant, Jeff, and Fahmy, Dalia. "Decline of Christianity in the U.S. Has Slowed, May Have Leveled Off". Online. Feb. 2025. URL: `https://www.pewresearch.org/religion/2025/02/26/decline-of-christianity-in-the-us-has-slowed-may-have-leveled-off/`.

[42] So, Richard Jean, Long, Hoyt, and Zhu, Yuancheng. "Race, Writing, and Computation: Racial Difference and the US Novel, 1880-2000". In: *Journal of Cultural Analytics 3*, no. 2 (Jan. 2019). DOI: `10.22148/16.031`.

[43] Song, Kaitao, Tan, Xu, Qin, Tao, Lu, Jianfeng, and Liu, Tie-Yan. "MPNet: Masked and permuted pre-training for language understanding". In: *Proceedings of the 34th International Conference on Neural Information Processing Systems*. NIPS'20. Vancouver, BC, Canada: Curran Associates Inc., 2020. ISBN: 9781713829546.

[44] Stark, Rodney. "Secularization, R.I.P." In: *Sociology of Religion 60*, no. 3 (1999), pp. 249–273. ISSN: 1069-4404. DOI: `10.2307/3711936`. (Visited on 05/05/2023).

[45] Swatos, William H. and Christiano, Kevin J. "Introduction — Secularization Theory: The Course of a Concept". In: *Sociology of Religion 60*, no. 3 (Oct. 1999), pp. 209–228. ISSN: 1069-4404. DOI: `10.2307/3711934`. URL: `https://doi.org/10.2307/3711934`.

[46] Tagliamonte, Sali A. and Jankowski, Bridget L. "Golly, Gosh, and Oh My God! What North American Dialects can Tell Us about Swear Words". In: *American Speech 94*, no. 2 (May 2019), pp. 195–222. ISSN: 0003-1283. DOI: `10.1215/00031283-7251241`.

[47] Thai, Katherine, Chang, Yapei, Krishna, Kalpesh, and Iyyer, Mohit. "RELiC: Retrieving Evidence for Literary Claims". In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, ed. by Smaranda Muresan, Preslav Nakov, and Aline Villavicencio. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 7500–7518. DOI: `10.18653/v1/2022.acl-long.517`. URL: `https://aclanthology.org/2022.acl-long.517/`.

[48] Tiryakian, Edward A. "The missing religious factor in imagined communities". In: *American Behavioral Scientist 55*, no. 10 (2011), pp. 1395–1414.

[49] Warner, Megan. ""Therefore a Man Leaves His Father and His Mother and Clings to His Wife": Marriage and Intermarriage in Genesis 2:24". In: *Journal of Biblical Literature 136*, no. 2 (2017), pp. 269–288.

[50] Wattenberg, Martin and Viégas, Fernanda B. "The Word Tree, an interactive visual concordance". eng. In: *IEEE transactions on visualization and computer graphics 14*, no. 6 (2008), pp. 1221–1228. ISSN: 1077-2626. DOI: `10.1109/TVCG.2008.172`.

A Limitations

We used a set of complementary techniques to identify religious speech, each with its own advantages and disadvantages. The use of keywords is highly interpretable and easy to validate, but may miss more subtle forms of religious expression. Although our list is more extensive than has been used by past work on this topic (e.g., [9]), the use of keywords still entails a lack of sensitivity to context. Our keyword list is also based on a comparison to exclusively Christian and mostly Protestant sermons, and thus does not capture terms specific to other religions. We expect non-Christian religious speech to be rare in Congress, given the overwhelmingly Christian makeup of its members. Our verse retrieval method is similarly specific to the Christian Bible, and only captures references to specific passages. The Bible's influence can also be found in various motifs and archetypes [5]. During validation, we identified a handful of references to other Christian texts, such as phrases from popular hymns (e.g., George W. Bush's reference to "wonder-working power" [1]) which is beyond the purview of our detection system.

We also restrict our study to the past few decades of the Congressional Record. This does not include the period when Christian conservatism was first ascendant, or the period when the latest broad-based liberal movement which used widespread religious language—the civil rights movement—was active. Ultimately, we chose to focus on the recent past due to a combination of data compatibility and a desire to avoid major problems with language change.

Finally, it is worth noting that Congress is only one arena in which notions of religious destiny and American nationalism are negotiated. As noted above, considerable focus has been directed towards *presidential* speech. The only previous quantitative paper on speech by members of Congress, [9], focused on their speech on Twitter rather than in the halls of the Capitol. Other work in political science and religious studies [2; 22; 29] has explored how political speech in American religious spaces informs the voting behavior and political beliefs of religious adherents, especially when the religion in question is felt to be in opposition to the majority of society. Our work is complementary to these, and future work could investigate how the religious speech found in Congress is both informed by and influences the religious and political rhetoric of other domains.

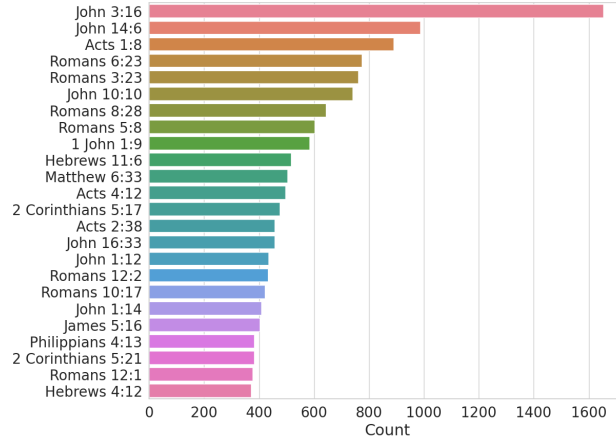## B Additional Details on Data

Congressional Record: We obtain the Congressional Record using publicly accessible code for downloading and processing the online editions: `github.com/unitedstates/congress`. Using this approach, we obtain all speeches given in the House and Senate from 1995 through 2022—the second half of the Clinton administration through the first half of the Biden administration—comprising 2.4 million speeches in total. Using the model from [11], we filter out procedural speeches (e.g., "without exception it is so ordered"), leaving 1.4 million non-procedural speeches. Speaker metadata is obtained from the same source, including speaker gender, party, chamber, and state.

Sermon Central: We obtained a copy of the Sermon Central dataset from the authors of [8]. Each sermon is typically associated with a particular verse of the Bible, the most common of which are given in Figure 8. As can be seen, the distribution is skewed towards certain popular verses.

C Religious Keywords

Lexicon creation: As described in the main paper we compare Sermon Central to the CR, looking for words that are used disproportionately more in the former, with COCA as a background corpus. Specifically, given a vocabulary of N words $\{w_1...w_N\}$ and corresponding counts of

Figure 8: Counts of the 25 most popular verses in Sermon Central.

occurrences for each corpus, $\{n_1^S...n_N^S\}$, $\{n_1^C...n_v^C\}$, and $\{n_1^B...n_N^B\}$ for Sermon Central, the Congressional Record, and the background corpus, respectively, the scoring function for a word w_i is given by
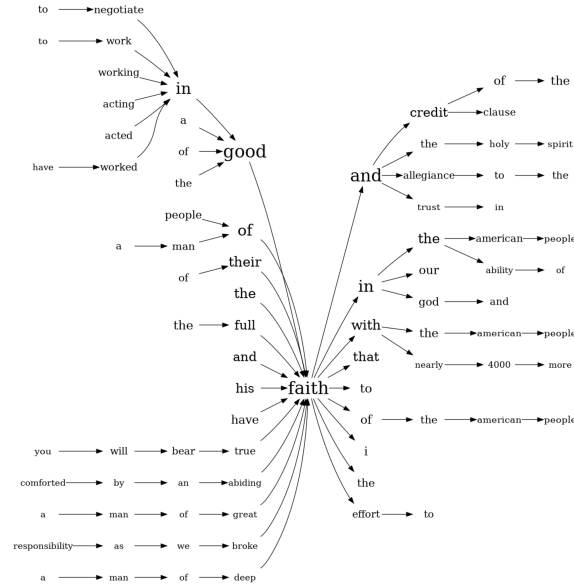
$$f(w_i) = \log\left(\frac{n_i^S + \alpha \cdot n_i^B}{-n_i^S - \alpha \cdot n_i^B + \alpha_0 + \sum_{j=1}^N n_j^S}\right)$$

$$- \log\left(\frac{n_i^C + \alpha \cdot n_i^B}{-n_i^C - \alpha \cdot n_i^B + \alpha_0 + \sum_{j=1}^N n_j^C}\right)$$

where α is a smoothing parameter, which we set at 0.1, and $\alpha_0 = \sum_{j=1}^N \alpha \cdot n_j^B$.

Keyword list: Counts for the most common religious keywords in the Congressional Record are shown in Figure 1 in the main paper. Note that we also include "Bible", "faith", and four variants of "pray" based on the terms from [9].

The full list is here: Abednego, abideth, Abram, adulterers, Ahab, AMEN, Ananias, apostle, apostles, Aren, ark, Baal, Baptism, Believers, believeth, Bethphage, Bible, Boaz, bridegroom, Caiaphas, centurion, Cephas, Christ, Chron, Colossians, conceited, Cor, Corinthians, covetousness, Deity, Deliverer, disciple, disciples, disobedient, dissensions, Divinity, doeth, dwelleth, Ephesians, Ephesus, faith, FAITH, Galatia, Galatians, Galilee, GENESIS, Gentiles, God, GOD, godlessness, godliness, Godliness, godly, Godly, Habakkuk, Haman, Haran, hast, Hebrews, Herod, Him, Himself, holiness, Idol, idolatry, Immortal, imperishable, iniquities, iniquity, Isa, Israelites, Jabez, Jephthah, Jesus, Jezebel, Joab, Judah, Judas, KJV, knoweth, legalism, Leprosy, Levite, liveth, Lystra, Magi, meekness, MERCY, Meshach, Messiah, NASB, Nebuchadnezzar, Nehemiah, Obedience, Olives, OUR, parable, Parable, parables, partakers, Pentecost, Pharaoh, Pharisees, Philippi, Philippian, Philippians, Philistine, Philistines, Pilate, pray, prayed, prayer, praying, prophets, Prophets, Rahab, Redeemer, repentance, Repentance, resurrection, Revelation, righteousness, Righteousness, Rom, sackcloth, saith, salvation, sanctification, Sanhedrin, Satan, Savior, Saviour, scoffers, seeth, Sermon, Shadrach, sinful, sinfulness, sinner, sinners, Sodom, sower, Sower, spake, Swindoll, tabernacle, Thessalonians, unbelief, unbeliever, unbelievers, unbelieving, undefiled, unfruitful, ungodliness, unrighteous, unrighteousness, Verily, whosoever, wickedness, william, wineskins, Yourselves, Zacchaeus, Zechariah, Zedekiah

Common terms. The most common terms from our list in the CR are "God" and "faith". In line with the wordtree diagram presented in the main paper (see Figure 4), we also created a wordtree

Figure 9: Wordtree diagram of the most common 5-grams in which the term "faith" appears.

diagram showing what the most typical uses of faith look like. As shown in Figure 9, there are a number of mentions of faith that are indeed religious; however, some other uses—see "in good faith" and "full faith and credit"—are false positives. The former phrase is not religious in nature, while the second appears in the Constitution.

D Biblical Quotations

Identifying Biblical quotations: Working with the validation data, we found both that both embedding similarity and n-grams had some value in recognizing different types of exact or approximate quotations. Using the annotated validation data (described in Section 5 in the main paper), we used a grid search over threshold values for minimum 5-gram overlap and minimum cosine similarity, and took the combination that resulted in the highest validation F_1 score. For the resulting combination (≥ 0.75 cosine similarity OR at least five unique 5-grams), there was a slight drop in performance from validation to test (from 0.91 to 0.81 F_1), suggesting some overfitting to the validation data, but overall our method performed well, and could scale to the full size of the congressional record.

Quote identification using language models: In order to evaluate the capabilities of language models on two tasks—detecting whether a sentence refers to a Bible verse as well as which verse—we used the validation and test sets discussed in Section 5 to compare different models and prompting strategies. It was not feasible to use a prompting-based method to detect Bible references across the entire corpus, given computational limitations, but we were able to evaluate various models and prompting strategies.

The experiments described in this section were done with two closed models with API access: OpenAI's `gpt-4o-2024-08-06` and Anthropic's `claude-3-7-sonnet-20250219`. They were also done with four open models: Meta's `Llama-3.2-3B-Instruct`, Meta's `Llama--`

`3.1-8BInstruct`, Mistral's `Ministral-8B-Instruct-2410`, and AI2's `OLMo-2-0325-32B--Instruct` (all taken from HuggingFace).[8] We used the validation set to choose effective prompts for each model; we then evaluated each model with its corresponding best prompt on the test set.[9] Overall, we tried 90 variations of the prompt below. Ultimately, we ruled out smaller models, partly based on noncompliance with our instructions.

Prompt: We began with a base prompt that captures two tasks of interest: identifying *whether* a sentence is a reference to a Bible verse and *which* verse it refers to, if so. We experimented with several prompting strategies that have been widely reported in other literature: trying different requested outputs, adding a chain-of-thought instruction, and few-shot provision of a positive example, a negative example, or both. The best prompt for `claude-3-7-sonnet-20250219` is shown below.

```
CONTEXT: You are tasked with determining whether a given string is paraphrasing or
quoting a Bible verse, then identifying which verse if so.

TASK 1:
For the input string, analyze it and determine if it's a Bible verse reference.
Print ``YES'' if it is a Bible verse reference, and ``NO'' if it is not. Do not
explain your answer.

TASK 2:
If the input string is a Bible verse, output the relevant citation (e.g., ``John
3:16''). If it is not a Bible verse, output ``NOT VERSE''. Do not explain your
answer.

OUTPUT FORMAT:
Separate the outputs for 1 and 2 with a newline.


The following example represents a correctly formatted input string-output pair.
EXAMPLE INPUT:
It may seem strange that any men should dare to ask a just God's assistance in
wringing their bread from the sweat of other men's faces but let us judge not that
we be not judged.

EXAMPLE OUTPUT:
Verse
Genesis 3:19

Output ``Verse'' if it is a Bible verse reference, and ``Not Verse'' if it is not.
Do not explain your answer.

INPUT STRING:
```
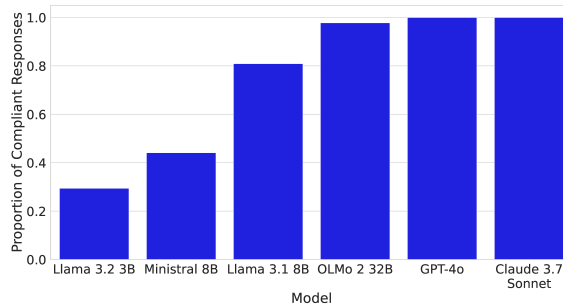
Results: We find that large closed models in particular are generally very good at detecting Bible verses (see Figure 1). The smallest models were unable to comply with instructions, as shown in Figure 10. We exclude these models from the following performance evaluation.

Among those models that had high compliance, GPT-4o and Claude 3.7 Sonnet stand out as being very good at recognizing Biblical references, as shown in Table 1.[10] The models that were

---

[8] The open models were run on A6000s GPUs. Our experiments with Claude cost about USD$50, while our experiments with GPT-4o cost about USD$30.

[9] We also ran preliminary experiments with the validation set to explore the impact of generation temperature; we found that 0.25 was effective for all models.

[10] Note that noncompliant responses were understood as predictions that a sentence was not a verse for the purposes of Table 1.

Figure 10: Compliance in output format with respect to our instructions. The smallest models frequently produced Python code, repeated instructions, or otherwise failed entirely to create usable output.

Model	F1	Precision	Recall
Baseline	0.807	0.867	0.838
Llama 3.1 8B	0.836	0.845	0.831
OLMo 2 32B	0.918	0.918	0.919
GPT-4o	0.970	0.972	0.971
Claude 3.7 Sonnet	**0.978**	0.978	0.978

Table 1: Weighted F1, precision, and recall for each model and its best prompt (with respect to the validation set) on the test set. Baseline is the method described in the main paper.

good at identifying that something was a verse were also the prompt were able to also consistently identify the correct verse. On this second task, OLMo 2, GPT-4o, and Claude 3.7 Sonnet all reach accuracy rates about 0.93.
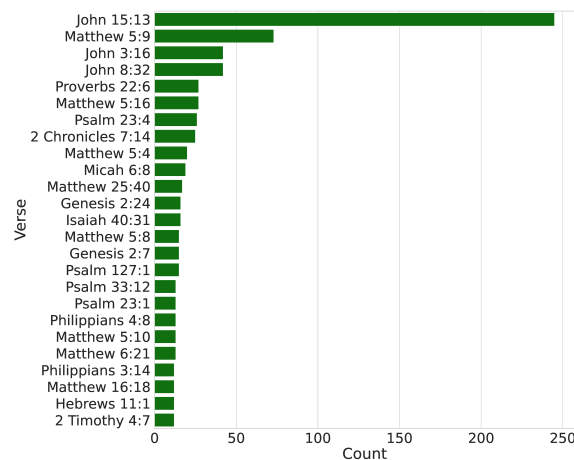
Notably, the embedding and n-gram method we use in the main paper has higher precision—only mislabeling a single positive verse—and is faster; these results are somewhat in line with other work on use of language models in historical document annotation, such as that by the Library of Congress (see `https://labs.loc.gov/work/experiments/ECD/`), that suggest a synthetic approach that incorporates both language models and structured measures of textual similarity is best for some multi-part classification and detection tasks.

Filtering baseline predictions: Given the excellent performance of Claude 3.7 Sonnet on this task, we explored using it as a way to filter out potential false positives from our baseline method's predictions. Out of 1,555 Biblical references identified my our method, 37 were identified by the prompting-based approach as false positives. However, manually inspecting these revealed that only 6 of these were actually false positives. The rest were mistakes by the language method, and tended to be references to verses that have become relatively common sayings, such as "The Good Book tells us that the truth will make you free" (John 8:32), or "We are all children of God" (Galatians 3:26).

Biblical references in the Congressional Record: Figure 11 shows the 25 most frequently quoted Bible verses in the Congressional Record, based on our analyses. To give a better sense of what verse matches look like in practice, Table 2 shows a random sample of 15 examples where our method found high cosine similarity between a sentence from the Congressional Record, and a verse from the Bible.

Text	Detected Verse	Citation	Cosine	Name	Party
And this is the condemnation: that light is come unto the world, and men loved darkness rather than light, because their deeds were evil.	And this is the condemnation, that light is come into the world, and men loved darkness rather than light, because their deeds were evil.	John 3:19	0.98	William Barrett	R
And He said unto them, Go ye into the world, and preach the gospel to every creature.	And he said unto them, Go ye into all the world, and preach the gospel to every creature.	Mark 16:15	0.99	Bennie Thompson	D
Guide us, thine own, aright Teach us by day, by night, To keep thine honor bright, For thee to fight.	Obey them that have the rule over you, and submit yourselves: for they watch for your souls, as they that must give account, that they may do it with joy, and not with grief: for that is unprofitable for you.	Hebrews 13:17	0.82	John Shimkus	R
For we know what is impossible with men is possible with God.	But Jesus beheld them, and said unto them, With men this is impossible; but with God all things are possible.	Matthew 19:26	0.81	Emanuel Cleaver	D
And remember the words of the Lord Jesus, that He said, 'It is more blessed to give than to receive.'	I have shewed you all things, how that so labouring ye ought to support the weak, and to remember the words of the Lord Jesus, how he said, It is more blessed to give than to receive.	Acts 20:35	0.83	Sanford Bishop	D
The Lord is my shepherd; I shall not want.	{A Psalm of David.} The LORD is my shepherd; I shall not want.	Psalm 23:1	0.91	Mark Pryor	D
As he said, greater love hath no man, no one, than this, that he lay down his life for his friends.	Greater love hath no man than this, that a man lay down his life for his friends.	John 15:13	0.92	Louie Gohmert	R
The verse concludes: Therefore shall a man leave his father and his mother, and shall cleave unto his wife: and they shall be one flesh.	Therefore shall a man leave his father and his mother, and shall cleave unto his wife: and they shall be one flesh.	Genesis 2:24	0.93	Addison McConnell	R
"He has told you, O mortal, what is good; and what does the Lord require of you but to do justice, and to love kindness, and to walk humbly with your God?	He hath shewed thee, O man, what is good; and what doth the LORD require of thee, but to do justly, and to love mercy, and to walk humbly with thy God?	Micah 6:8	0.85	John Boehner	R
But to do justly, to love mercy and to walk humbly with your God."	He hath shewed thee, O man, what is good; and what doth the Lord require of thee, but to do justly, and to love mercy, and to walk humbly with thy God?	Micah 6:8	0.88	Bobby Rush	D
Arise, walk through the land in the length of it and in the breadth of it; for I will give it to thee.	All that the Father giveth me shall come to me; and him that cometh to me I will in no wise cast out.	John 6:37	0.81	James Inhofe	R
The Word of God is the source of love, joy, peace, patience, kindness, goodness, and self-control.	For the word of God is quick, and powerful, and sharper than any twoedged sword, piercing even to the dividing asunder of soul and spirit, and of the joints and marrow, and is a discerner of the thoughts and intents of the heart.	Hebrews 4:12	0.82	Cathy McMorris Rodgers	R
In the Name of the Father, the Son and the Holy Spirit, Amen.	But grow in grace, and in the knowledge of our Lord and Saviour Jesus Christ. To him be glory both now and for ever. Amen.	2 Peter 3:18	0.81	Daniel Akaka	D
For we do not wrestle against flesh and blood, but against the rulers, against the authorities, against this present darkness, against the spiritual forces of evil in the heavenly places.	For we wrestle not against flesh and blood, but against principalities, against powers, against the rulers of the darkness of this world, against spiritual wickedness in high places.	Ephesians 6:12	0.96	Rick Allen	R
Today in the town of David a savior has been born to you.	For unto you is born this day in the city of David a Saviour, which is Christ the Lord.	Luke 2:11	0.84	Vicky Hartzler	R

Table 2: A random sample of 15 Congressional Record sentences whose embeddings from MPNet had high cosine similarity with that of the most similar verse.

Figure 11: Absolute counts of how many times the most popular verses are quoted.