Signalling Incentives and the Quality of Legislation: A Text and Network Analysis of the U.S. Congress *

Tommaso Giommoni[†] Massimo Morelli[‡] M. Daniele Paserman[§]

September 2022

Abstract

In this paper we model and evaluate empirically the legislative production incentives of members of Congress, with and without the possibility of cosponsoring legislation drafted by others. We set up a model in which Congress members must decide whether to submit a bill to the floor, and the floor decides whether to approve the bill. Restrictions on cosponsorship induce Congress members to submit more bills, which results in a lower quality of bills and laws. This is a result of three forces: a *frivolous bills* effect, which induces Congress members to submit more bills to signal activism even if they are aware that the bills are low quality; a *screening* effect, resulting from other legislators only observing imperfectly the quality of bills, which prevents them from successfully screening out low quality bills; and a *congestion* effect, whereby Congress can spend less effort on improving the quality of the bills if many bills are submitted.

To test the predictions of the model, we exploit a reform in the 96$^{\text{th}}$ Congress (1979-1981) that removed a hard cap on the maximum number of cosponsors in the House of Representatives, while the regulations surrounding cosponsorship in the Senate remained unchanged. This allows us to set up a difference-in-differences analysis, comparing the quality of laws originating in the House to those originating in the Senate, before and after the reform. To assess empirically the quality of laws, we use a variety of indicators based on both text and network analysis. We calculate measures of readability of the text of laws, and calculate various measures of citations, network centrality, and codification by mapping the entire newtork of laws passed in Congress.

*We wish to thank the Framework for Attraction and Strengthening of Excellence (*FARE*), grant for projects on legislative complexity, and the Dondena research center. The usual disclaimer applies.

[†]ETH Zurich, Dondena and PERICLES

[‡]Bocconi University, IGIER, PERICLES and CEPR

[§]Boston University, NBER, CEPR and IZA

We find that, in the period where co-sponsorship possibilities were limited, excessive legislative activism caused production of significantly lower quality laws, demonstrating how damaging signalling incentives can be. The main mechanism driving this result is the congestion effect, while the frivolous bills and screening effects play a more limited role.

Keywords: production of laws, co-sponsorship, congestion, frivolous laws, text analysis, network analysis.

1 Introduction

A well-functioning polity depends on the quality of its laws. Ideally, laws should be clear, easy to interpret, and apply universally. But, in the process of lawmaking, legislators may have other incentives. They may be interested in signalling their positions to their constituents (Matthews, 1960) or to other legislators (Kessler and Krehbiel, 1996). Alternatively, they may be interested in signalling activism as a way to build a reputation as skillful reformers, rather than being motivated by an honest and competent assessment of the welfare consequences of the reform (Gratton, Guiso, Michelacci and Morelli, 2021). Any of these motives can lead to deterioration in the quality of proposed legislation, and by consequence, the quality of legislation that is enacted into law.

In this paper, we study the effect of signalling incentives on the quality of bills and laws using data from the U.S. Congress between the years 1973 and 1989 (93^{rd} to 100^{th} Congress). We exploit the change in the rules governing cosponsorship during this period in the House of Representatives. In the period between the 91^{st} (1969-1971) and the 96^{th} Congress (1979-1981), the House had a strict limit of a maximum of 25 signatories per bill.[1] The cap was removed completely starting with the 96^{th} Congress, and cosponsorship was allowed again without any limitation to the number of signatories. Importantly, throughout this whole period bill cosponsorship was allowed without limit in the Senate. This allows us to implement a difference-in-difference analysis, comparing the quality of legislation produced in the House and the Senate before and after the lifting of the cosponsorship cap. We supplement the simple difference-in-difference analysis with an event study analysis, to ascertain that the changes in observations in the House after the 96^{th} Congress are not the result of pre-existing trends.

To guide the empirical analysis, we introduce a simple model of the cosponsorship decision. A congressional committee is made up of two members, each drawing an "idea" for a bill, which can be good or bad. Besides deciding whether to propose their own bill, each

[1] This was already a more relaxed regime relative to the previous period, when cosponsorship was completely banned.

member must decide whether to cosponsor the other member's bill if proposed. The "floor" player imperfectly observes the quality of the drafted bills, and decides whether to pass each bill. Members draw utility from the quality of the bills that become law, but also from having their name on any bills – they are intrinsically motivated to signal legislative activism.

We show that restrictions on cosponsorship induce Congress members to submit more bills, which results in a lower quality of bills and laws. This is a result of multiple forces. Bills are of lower quality because of a *frivolous bills* effect, whereby Congress members submit more bills to signal activism, even if they are aware that the bills are low quality. The frivolous bills effect, combined with the floor's inability to perfectly observe the quality of a proposed bill (a *screening* effect), also leads to a deterioration in quality of approved laws. The deterioration in quality of laws can also be the result of a *congestion* effect, since Congress can spend less effort on improving the quality of proposed legislation if many bills are submitted.

To test these predictions, we create a novel dataset on legal production in the US. We collect detailed information on the universe of proposed bills and approved laws covering Congresses 93-100 (1973-1989). This dataset contains information on 117,456 bills and 5,670 approved laws. For each legal act, we have rich information including the title, the text, the topics discussed, the entire legislative process as well as the legislators involved. We then map the citation network of laws in the US congresses, with the goal of representing the American legal system as a network where the laws are the nodes of the network, and citations the links (Bommarito and Katz, 2010).

We rely on these rich data to create indicators of legislative quality applying text analysis and network techniques. We first construct various simple measures of the length and readability of bills and laws: number of words per sentences and average length of words or sentences. We then proceed to measure legislative quality by analyzing the full network of laws enacted by Congress. For each law, we first count the number of times the law is cited in the future by other legal acts (in-citations). As in the bibliometrics and patent literature (Narin et al., 1976; Narin, 1994; Jaffe and Trajtenberg, 2002), this is a key indicator of qual-

ity. Second, we study the centrality of each law in the entire network of laws as a measure of relevance for jurisprudence. Third, we consider the codification process, and we measure whether a law is included in one or more sections of the United States Code. The inclusion in the U.S. Code is a signal of relevance for a piece of legislation. These three groups of indicators allow us to observe the quality of U.S. laws from different perspectives.

The main results confirm the implications derived in the model. The reform in the 96^{th} Congress deeply affected the legislative process: it dramatically reduced the number of proposed bills; and also affected the number of parliamentary actions a bill undergoes, and the average probability that a bill becomes law, all in the directions predicted by our simple model. We also observe important modifications in the text of approved laws, which become longer but not necessarily more complex, as the average length of words and sentences decreases.

Second, we show that eliminating limits to cosponsorship significantly raised legislative quality in the House, as measured by our three groups of indices. First, we show that approved laws are cited by more legal acts in the following congresses. Second, laws introduced after the reform are more central in the legislative network, and their centrality increases as the network grows. Third, the reform increases the likelihood that approved laws are included into one or more sections of the United States Code. Overall, these findings document a consistent increase in the quality of laws after the lifting of cosponsorship limits in the House.

We then test the three potential mechanisms, and find mostly support for the congestion hypothesis. Using quantile regressions, we find that the quality of laws in the top 20% of the quality distribution increases significantly in the House, relative to the Senate, after the removal of cosponsorship. This result suggests that the over-production of bills when cosponsorship was limited affected the quality of laws by reducing the efficiency of Congressional revision, consistent with the congestion effect. On the other hand, we find that the quality of laws at the 30^{th} percentile of the quality distribution increases only very slightly. This suggests that the ability of the floor to filter out low-quality bills was not significantly

hindered in the no-cosponsorship regime. Rather, the screening filter, while imperfect, still provided some protection against the worst of the frivolous bills.

The fact that banning or limiting cosponsorship impacted on the production of bills had already been noted in the literature (see e.g. Cooper and Young, 1989; Thomas and Grofman, 1993; Adler and Wilkerson, 2013).[2] Among the scholars who have studied individual motivations for cosponsorship, Campbell (1982) emphasized the efforts and incentives of bill proposers to actively recruit cosponsors and focused on the position-taking incentive.[3] We focus instead on the complementary incentives of legislators to become cosponsors rather than producing individual bills, especially for legislators who are unsure about the value of the proposal they would focus on in the absence of the possibility of signalling activism through cosponsorship. The electoral connection theory championed by Mayhew (2004) and signalling theories in general (see e.g.Wilson and Young, 1993, for incentives to signal ideology, expertise and loyalty), are all consistent with our simple signalling motivation assumption in the model. Cosponsorship constitutes low-cost signalling to multiple audiences.[4] There is also a literature on what determines the choice of which bills to cosponsor. This literature mostly uses data from the 96$^{\text{th}}$ Congress (Koger, 2003) or later (Krehbiel, 1995). By contrast, our work is exclusively interested in the quantity and quality effects on bills and laws in the aggregate, and does not aim to make any contribution on the prioritization criteria that each legislator makes within any cosponsorship regime, before and after the reform.

To sum up, the literature finds multiple reasons to observe legislators' preference for the institution of cosponsorship, ranging from position taking, signalling internally and externally to different audiences, and elimination of duplicate bills (Thomas and Grofman, 1993).

[2]The quantity dynamics can be appreciated using the data provided by Thomas and Grofman (1993), and we will use them for the initial descriptive figures below.

[3]Fowler (2006) finds that an incentive to recruit co-sponsors signals centrality in the legislators' network and such centrality reflects legislative influence.

[4]Some empirical work focuses on the target audiences of legislative signalling. For example Kessler and Krehbiel (1996) and Laband, Seals and Wilbrandt (2015) find more evidence in favor of intra-legislative than extra-legislative signalling. For our model and results it makes no difference whether signalling is directed to other Congress members, to party elites, or to voters.

Less clear are the motivations for having a limit on cosponsorship.[5] In any case, our focus is on the consequences of the different regimes for the whole legislation process, and most important for the quality of laws, while the individual preferences for the different regimes is orthogonal to our research question. To the best of our knowledge, no other paper exists that traces such implications on quality for U.S. Federal legislation.[6]

Our paper is also related to the literature on measuring the quality of laws treating laws as a network. The most important references are Bommarito and Katz (2010) and (Whalen, 2016) and references therein.[7] Many studies focus on other legal systems: Koniaris, Anagnostopoulos and Vassiliou (2018) focus on the laws of the European Union, Mazzega, Bourcier and Boulet (2009) study the network on French codes, Wilsona and Zakerib (2016) analyze New Zealand, Waltl and Matthes (2014) focus on Germany and Boella, Di Caro, Violato and Robaldo (2014) treat the Italian case. We borrow from this literature on computational law the idea of representing the legal system as a network and develop a set of original indicators with the aim of measuring the quality of laws. These are measures of citations, centrality of laws in the legal network, and codification of laws. On the empirical relevance of imperfect filtering of quality of bills and congestion effect see e.g. Box-Steffensmeier, Christenson and Craig (2019).

The data on sponsorship of bills has mostly been used to analyze legislative effectiveness of legislators, which by and large is correlated with the probability of approval of a sponsored bill.[8] We do not focus on legislative effectiveness, but we do use the data on probability of

[5]In the Senate, for instance, some Senators have considered cosponsorship as the legislative equivalent of a common courtesy (Oleszek, 2018). The reasons against cosponsorship are summarized in this quote by Senator Everett Dirksen (R-IL), who proposed in February 1967 a –failed– resolution to end the practice of cosponsorship in the Senate: *"The bill may be 50 pages long, and I will not know what is in it from the enacting clause to the last period. But I do not wish to affront a Senator, and if he gives me a sufficient sales talk, I am likely to say, "All right, go ahead and add my name."* (Congressional Record, vol. 113, part 3 (February 16, 1967), p. 3725)

[6]Gratton et al. (2021) study the effects of signalling incentives on quantity and quality of laws in Italy, but the effects passed through a simultaneous deterioration of bureaucratic capacity, and hence it is difficult there to identify the direct effect of different signalling incentives.

[7]Bommarito and Katz (2010), Katz and Bommarito (2014) and Lyte, Slater and Michel (2015) study the United States code with the aim of characterizing how this evolves over time in terms of complexity while Fowler, Johnson, Spriggs, Jeon and Wahlbeck (2007) and Fowler and Jeon (2008) focus on the network of Supreme Court sentences in the United States.

[8]With minor variations of measures and emphasis, legislative effectiveness was defined that way from

approval of bills as one of the variables that are impacted by the cosponsorship reforms.

The rest of the paper is structured as follows: in Section 2 we present the institutional background on the cosponsorship process in the U.S. Congress, and how its regulation evolved over time in the House of Representatives. In Section 3 we introduce our theoretical model, and derive predictions for the effect of restricting cosponsorship on the quality of bills and laws. In Section 4 we describe our data and the various measures used to assess legislative quality. In Section 5 we present our main results, and analyze the mechanisms in Section 6. Section 7 concludes.

2 Institutional background: co-sponsorship regulation

The legislative process in the United States begins with the introduction of a bill in one of the two chambers of Congress. The bill is introduced by a single sponsor, and additional legislators may sign on to the bill as cosponsors. After the bill has been introduced in the chamber, it is referred to one or more committees for study and revisions. If it is approved by the committee(s), the bill then moves to the floor for debate and possible amendments; and finally a vote is taken. If the bill passes, it is referred to the other chamber and if it passes, it is sent to the President for approval.

In this paper we focus on the legal institution of cosponsorship, i.e. the possibility for legislators to add their name to a bill.[9] Sometimes a bill is drafted jointly by several legislators, with one of them acting as the main sponsor, while the others immediately sign their name on the bill as cosponsors (these are known as the "original cosponsors.") In other instances, cosponsors add their name to the bill later, potentially even several weeks or months after a bill was initially introduced.

The regulation of bill cosponsorship in the House of Representatives has changed over time.[10] The practice was officially banned since early twentieth century (*Congressional*

Matthews (1960) to Volden and Wiseman (2014). Gagliarducci and Paserman (2022) instead focus on the ability to recruit cosponsors as a measure of legislative effectiveness.

[9]This possibility applies to public bills, while private bills cannot have cosponsors, but may be introduced in the same way as public bills (https://www.everycrsreport.com/reports/RS22477.html).

[10]The institutional information on cosponsorship regulation is based on Thomas and Grofman (1993),

Record 1909, 3808-11). Cosponsorship was then introduced with a maximum number of 25 signatories per bill from the 91st Congress (1969-1971) (*Congressional Record 1967, 10708-12*). This cap was finally removed and cosponsorship was allowed without any limitation to the number of signatories from the 96th Congress (1979-1981) onwards (*Congressional Record 1978,34929-31*). In the Senate, on the other hand, bill cosponsorship was always allowed, without limits on the number of cosponsors.

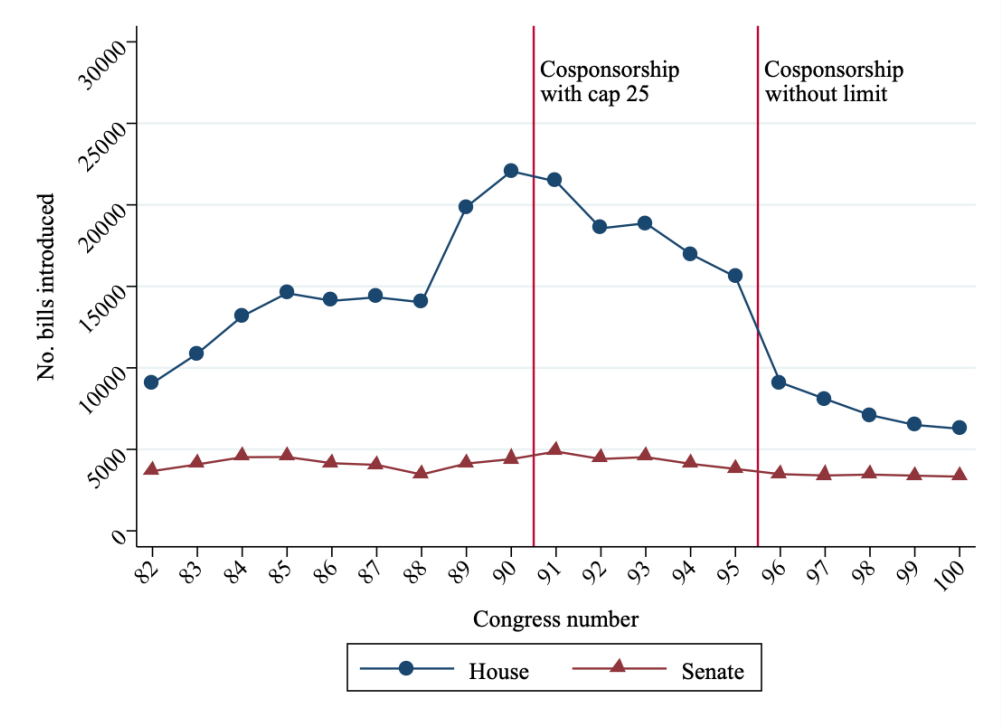
The changes in regulation modified the legislative process in the House of Representatives. First, they led to a sharp discontinuity in the total number of bills submitted in the House, in comparison with the trend in the Senate. Figure 1 shows the total number of bills introduced in the House and in the Senate during the three regimes. Before the 91st Congress, when the practice of cosponsorship was banned in the House, the number of bills introduced in the House was larger than the one in the Senate and the gap gradually increased over time. This trend is almost entirely driven by a sharp increase in the number of bills in the House, while the number of bills in the Senate stays roughly constant. The trend starts to revert after the 91st Congress, when the cosponsorship ban was partially lifted and a cap of 25 cosponsors per bill was introduced. The complete removal of the cap in the 96th Congress resulted in a sharp drop in the number of bills introduced in the House, with the gap between the House and the Senate appearing to stabilize.

Figure 2 zooms in to the period going from the 93rd to the 100th Congress, on which our analysis is focused.[11] The first subplot shows that the average number of cosponsors was relatively flat in both chambers up to the 95th, and then starts to rise rapidly after the cosponsorship upper bound is removed, with the growth in the number of cosponsors in the House far outpacing that in the Senate. This highlights that the constraint before the reform was a binding constraint. The second subplot of Figure 2 confirms this evidence by showing that the number of approved laws with more than 25 cosponsors increases at a faster pace in the House relative to the Senate.

Kessler and Krehbiel (1996) and Wilson and Young (1997).

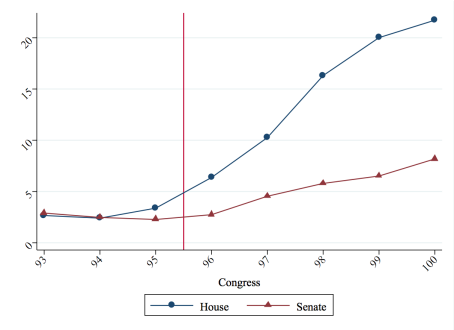
[11] We focus on this period because of our interest in the full text of laws. The first congress for which the full text is available in digitized form is the 93rd Congress.

Figure 1: Bills introduced in the US Congress over time

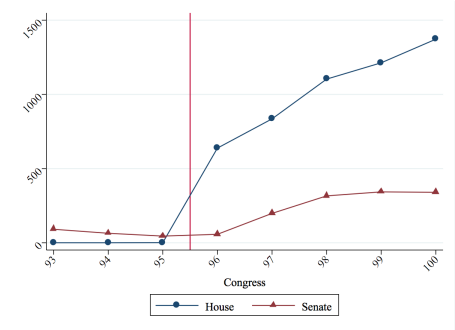


Notes: Authors' elaboration with data from Thomas and Grofman (1993).

Figure 2: Descriptive graphs on cosponsors.



(a) Av. number of co-sponsors (bills)

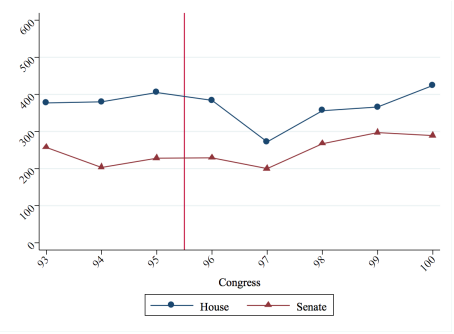


(b) Bills with more than 25 co-sponsors

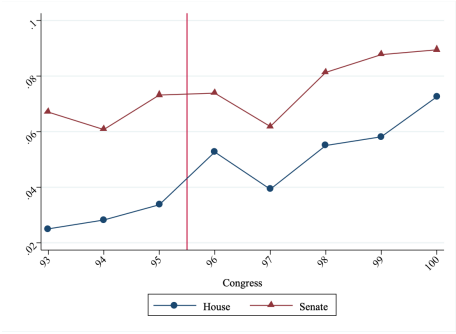
The changes in bill production also affected the production of laws. The first subplot of Figure 3 shows that the number of approved laws originating in the House declines, relative to the Senate, after the removal of the cosponsorship bound, and the second subplot shows

that the fraction of approved bills significantly increases in the House with the reform.

Figure 3: Descriptive graphs approved laws.



(a) Number laws approved



(b) Fraction bills approved

From these trends, it appears that keeping an upper bound on cosponsorship led to an over-supply of bills in the House compared to the Senate, as already argued by other scholars (Cooper and Young, 1989; Thomas and Grofman, 1993; Adler and Wilkerson, 2013).[12] What needs to be studied is the *consequence* of these distortions on the *quality* of laws, not just on the quantity. In the next sections we will first argue theoretically and then show empirically that a limit (or ban in the model) on cosponsorship did indeed affect the quality of bills and laws.

3 Model

We present here a stylized model of the process by which a bill is drafted and eventually passed into law. We model this process as a game with three participants: two committee members, labeled by $j = 1, 2$, who draw ideas for bills of heterogeneous quality, and must decide whether to draft a bill and, if the practice is allowed, to cosponsor their colleague's bill; and a "floor" player (labeled v), who imperfectly observes the quality of the bill, and must then decide whether to approve the bill. Both committee members draw utility from the quality of approved legislation, but also from having their name attached to a bill, to

[12]The same trend in the number of bills emerges even if we remove duplicate bills from the data (available upon request). Removal of the cosponsorship cap deeply affected the entire legislative production, and did not just lead to the elimination of duplicate bills.

capture the signalling value of displaying legislative initiative. Importantly, if the committee drafts only one bill, the two members can devote more time to improving its quality, so that the resulting final quality of the bill is higher than its initial quality. We consider Perfect Bayesian Equilibria in this game, and compare the equilibrium when cosponsorship is allowed to the equilibrium when it is banned.

3.1 Time line

The time line of the game is as follows:

1. Each committee member j draws an idea for a bill randomly, with quality realization $q^j \in \{h, l\}$, with $h > l$. There is a common prior probability $\pi \in (0, 1)$ that $q^j = h$.

2. The quality of each idea is observed by both members of the committee. At this point each member j of the committee decides whether to draft a bill proposal from her idea: $(s^j = 1)$ or not $(s^j = 0)$.

3. **Cosponsorship stage:** Each committee member decides whether to co-sponsor the bill proposed by the other committee member: $c^j = 1$ denotes the decision to co-sponsor the bill proposed by the other politician, and $c^j = 0$ otherwise.

4. In case the two committee members draft only one bill, we assume that together they can have more time and devote more effort to improve the quality of the unique bill, so that if the chosen bill is the one by committee member j its final quality is $q^j + E$, with $E > 0$.

5. The "floor" player observes a noisy signal on q^j: she obtains a signal $m_j \in \{h, l\}$. The signal is correct $(m_j = q^j)$ with probability $1 - z$ and mistaken with probability z, with $z \in (0, \frac{1}{2})$.

6. Given any number $S = s^1 + s^2 \in \{1, 2\}$ of bills to consider, the floor player v chooses whether to pass each bill j. Thus, for each bill j the action by v can be represented as $a^v(S, m_j) \in \{1, 0\}$—of course $a^v = 1$ (0) denotes approval (rejection).

Letting $b^j = s^j + c^j$ be the number of bills with j's name, the utility function for a committee member is as follows:

$$U_j = \bar{Q} + \theta I_{b^j>0}$$

where $\bar{Q}$ is the average quality of bills that become law (reputation effect) and $\theta > 0$ captures the reduced form signalling value of displaying legislative initiative. $I_{b^j>0}$ is a standard indicator function that takes value 1 when some legislative activity is recorded and zero otherwise.

On the other hand, the expected utility of the floor player from *each* approved bill is

$$W^v(S, m_j) = \hat{Q}^j(S, m_j),$$

where $\hat{Q}_j$ is the expected quality of the induced reform. The floor player has a simple decision rule, namely $a^v(S, m_j) = 1$ if and only if $\hat{Q}^j(S, m_j) > q_0$, where q_0 is the status quo utility without the proposed law.

We look for Perfect Bayesian Equilibria of this game.

3.2 Analysis

We first consider the case in which $q_0 < l$.

Proposition 1 *Assume $q_0 < l$. Then in equilibrium the floor player always approves bills that are proposed and*

(I) In the game with cosponsorship, the equilibrium behavior of the committee members displays $S = 1$ for all $\theta > 0$ and $E > 0$. Moreover, if $E < \frac{h-l}{2}$, only the highest quality bills are proposed in equilibrium.

(II) In the game without cosponsorship, if $E < \theta$ then whenever $q^1 = q^2$ we have $S = 2$. Moreover, if $\theta - E > \frac{h-l}{2}$, then $S = 2$ in every equilibrium, including when $q^1 \neq q^2$. Otherwise, $S = 1$ when $q^1 \neq q^2$, with only the highest quality bill being proposed.

Proof. Fix a Perfect Bayesian equilibrium. Then the floor player always approves the bills that are proposed since it is a strictly dominant strategy given that $q_0 < l$.
(I) Consider any pair with $q^i \geq q^j,\ i = 1,2,\ j = 1,2,\ j \neq i$. Then $S = 2$ cannot be an equilibrium: committee member j has incentive to deviate to $s^j = 0$ and $c^j = 1$ as it increases his payoff by at least $E > 0$. $S = 0$ cannot happen in equilibrium either: every committee member has incentive to deviate drafting abill. Therefore, in equilibrium it must be that $S = 1$ for all pairs (q^i, q^j). Moreover if $E < \frac{h-l}{2}$, in equilibrium it must be that if $q^j = h > q^i = l$, then $s^i = 0$ and $s^j = 1$. To see this, suppose $s^i = 1$. Then since $S = 1$ we have $s^j = 0$. But since $h > l + E$, by deviating to $s^j = 1$, committee member j increases his payoff since then $\bar{Q} = \frac{h+l}{2} > l + E$ (where the inequality holds as it is equivalent to $E < \frac{h-l}{2}$).
(II) Suppose that $E < \theta$ and consider any pair of qualities (q^i, q^j). Suppose $q^i = q^j$. If $S < 2$ and $s^j = 0$, then $s^j = 1$ yields a higher payoff since $q^j + \theta > q^i + E > Q_0$ (as $q^i = q^j$ and $\theta > E$). Therefore when $q^j = q^i$, we have $S = 2$ (which is easily seen to be an equilibrium). When $q^j \neq q^i$, say $q^j = l$ and $q^i = h$, we have $S = 2$ if $\theta - E > \frac{h-l}{2}$. Suppose $S = 1$ and $s^j = 0$. Then $\bar{Q} = h + E$. Deviating to $s^j = 1$ yields a higher payoff as $\bar{Q}' = \frac{l+h}{2}$, and $\bar{Q}' + \theta = \frac{l+h}{2} + \theta > h + E$ is equivalent to $\theta - E > \frac{h-l}{2}$. **QED.**

The above result shows that for θ large enough and q_0 low enough, we have $S = 1$ with co-sponsorship and $S = 2$ without co-sponsorship. Beside this sharp difference in terms of quantity of laws, it also naturally follows that quality should be higher on average when cosponsorship is allowed, both because of the E extra care effect, and because of the incentive for the legislators to propose bills regardless of their quality when cosponsorship is banned.

Even though the results above suffice to provide the main insights, let us consider now the case in which $q_0 \in (l, h)$, so that the floor player does not always approve all bills depending on his beliefs and information about the quality of the bills that are proposed.

When j's bill drafting decision is uninformative about quality because $s^j(q^j, q^i) = 1$ for all (q^j, q^i), the floor player updates his beliefs by relying on his signal m_j about the bill's quality. Using Bayes' rule, the updated probabilities $\rho(m_j)$ of the bill being of high quality

are:

$$\rho(m_h) = \frac{(1-z)\pi}{(1-z)\pi + z(1-\pi)}$$

and

$$\rho(m_l) = \frac{z\pi}{z\pi + (1-z)(1-\pi)}.$$

When there is a pooling of j's behavior, that is, $s^j(q^j, q^i) = 1$ for all (q^j, q^i), we have $\hat{Q}^j(1, m) = \rho(m)h + (1-\rho(m))l + E$ if $S = 1$, while if $S = 2$, then for each j we have in case of pooling:

$$\hat{Q}^j(2, m_j) = \rho(m_j)h + (1-\rho(m_j))l.$$

Proposition 2 *Assume $q_0 \in (l, h)$. When $\theta > (h-l) + E$, in the game without co-sponsorship the unique equilibrium is such that $s^j(q^j, q^i) = 1$ for all j, q^j, and q^i, and:*

(i) there exists z^* such that $a^v(2, m_h) = 1$ for all $z < z^*$;

(ii) If $q_0 < \pi h + (1-\pi)l$ there exists $\hat{z}$ such that $a^v(2, m_l) = 1$ for all $z > \hat{z}$ (otherwise, $a^v(2, m_l) = 0$ for all z.)

Proof. Fix a Perfect Bayesian Equilibrium. First, notice that in equilibrium we must have $s^j(q^j, q^i) = 1$ for all (q^i, q^j) as, when $\theta + l > h + E$, it is a dominant strategy for committee members in the game without co-sponsorship to always propose bills. Indeed, by proposing a bill, the minimal payoff j could obtain is $\theta + l$ (if j's bill is passed but not i's bill), while by abstaining from proposing a bill the maximal payoff obtainable is $h + E$ (when $q^i = h$ and his bill is passed).

Therefore, in all PBEs all types of committee members pool into proposing bills. Then (i) and (ii) follow from the formulas of $\hat{Q}^j(2, m)$ and $\rho(m)$ for $m \in \{h, l\}$. Indeed the floor player adopts given m_j if and only if $\rho(m_j)(h-l) > q_0 - l$. For (i), it follows as when $z \to 0$, $\rho(m_h) \to 1$ and $h - l > q_0 - l$. For (ii), when $z \to \frac{1}{2}$ we have $\rho(m_l) \to \pi$, and so $\pi(h-l) > q_0 - l$ holds if and only if $q_0 < \pi h + (1-\pi)l$. **QED.**

From the above result, we can see that for a bill of perceived low quality to be approved, the signal must be considered sufficiently noisy. The case of high complexity and hence high

noise is therefore particularly problematic because there is a higher fraction of approved low quality reforms. Proposition 2 has also implications for the distinction between bills and laws: when the floor has greater difficulty evaluating the quality of a bill (high z) the filtering is inefficient and too many laws can pass, and this could in principle also determine a higher z down the road because of an additional congestion effect.[13] Overall, the propositions suggest that without cosponsorship the number of proposed bills is strictly higher and the average quality of approved laws is lower, for any $z > 0$.

The model highlights three reasons for a lower quality of laws during the time in which co-sponsorship is banned or limited: (1) a *congestion effect* captured by the inability of committees to add the extra care E to revise and improve the quality of bills; (2) a *frivolous bills* effect, captured by the incentive to draft more bills even when clearly low quality; (3) a potential *increasing noise* with which the floor player is able to evaluate the quality of each bill (captured by a potential increase in z). In our empirical analysis we will evaluate all the predictions of the model and we will try to assess which one of these three mechanisms is quantitatively the most relevant.

For the floor player a world with cosponsorship is preferable iff $\pi h + (1-\pi)l < E$. Given the results we find next on the crucial role of the congestion effect (i.e., E seems to be high), the model and empirics together seem to suggest that the floor player should have higher expected utility after the reform implemented in the 96th Congress. However, this conjecture is hard to test, as most welfare predictions. There will be no welfare analysis therefore in what follows.

4 Data and methodology

4.1 Information on bills and laws

We assemble a rich dataset with detailed legislative information for the United States. Our main source is the web portal *congress.gov* which allows us to cover the universe of the

[13] On the possibility that screening and filtering may be endogenously affected by overproduction of laws, see (Gratton et al., 2021) and Foarta and Morelli (2021).

proposed bills and approved laws for the time span 1973-1989, corresponding to Congresses 93-100. We choose this period because it allows us to have a symmetric time window around the reform date.[14] Also, for many of our citation and codification measures it is important to take a long-run perspective – e.g., by looking at the number of citations a law receives within a time span of 10 or 20 Congresses.

We collect data on all bills that have been proposed and all laws that have been approved in the United States Congress. The data set covers 117,456 bills presented and 5,670 laws approved for the period we analyze. For these legal acts we have a rich set of information. For each bill and law, we have a set of general information, such as the title and the type of bill,[15] the chamber of origin, the date of last action, the number of days for approval (only for laws), information on duplicate bills,[16] and the number of legislative actions and committees. We focus all of our analysis on public bills only, excluding all private legislation.

We also have information on the topic discussed in the document: in particular, we have the policy area and the number of legislative subjects, as classified by the Library of Congress. Figure A.1 presents a summary of major topics addressed by laws.[17] The three most common topics are *Environment, emergency*; *Democratic institutions*, and *Commemorations*, with more than 600 laws each.

We then have information on the legislative process for each bill: whether it fails after the introduction, it passes one chamber, it passes two chambers or it becomes a law. Finally, we have many details on legislators, such as the name of the sponsors, the names and the number of cosponsors, a set of demographic characteristics (age and gender), political experience, and a set of political variables: the state and the district of election, political party and the Poole-Rosenthal DW-NOMINATE scores from `voteview.com` as measures of ideology.

[14] Importantly, our dataset starts from Congress 93 because the portal *congress.gov* contains incomplete information on laws and bills for older congresses.

[15] We include in the data set House and Senate bills as well as House and Senate joint resolutions. We exclude from the analysis concurrent resolutions and House/Senate resolutions as these acts do not have the force of law and they either regulate the functioning of the chambers or express the sentiment of them (United States Senate).

[16] In particular, we know whether the bill has duplicate bills, how many and whether it is the original bill or a copy (in this case we know the order).

[17] Relative to the list of topics available in the Library of Congress, the figure has fewer topics, as we group topics into broader categories.

footnoteLewis, Poole, Rosenthal, Boche, Rudkin and Sonnet (2021)).

Table B.1, Panel A, shows the descriptive statistics for bills. More than 75% of bills originate in the House and a small fraction of them is eventually approved as law (4.8%). Bills have on average 3.6 legislative actions and 6.7 co-sponsors. Each bill has on average 3.3 duplicated bills, but the distribution of this variable presents a long right tail. Panel B of Table B.1 shows the descriptive statistics for laws. 60% of approved laws originate in the House. Approved laws experience more legislative actions and co-sponsors than bills, respectively 19.9 and 20 on average. The average number of approval days is 212.

4.2 Indicators of legislative quality: Readability measures

We create a first set of indicators to measure law quality that relies on the texts of laws. In particular, from the complete text of approved laws, as provided by the Library of Congress, footnoteThe web portal *congress.gov* does not include texts of laws prior to the 100$^{\text{th}}$ Congress. Therefore, we recover these information from the PDF versions of the *United States Statutes at Large* for Congresses 93-100. we apply simple text analysis techniques to construct a set of indicators of the length and readability of texts. In particular, we collect for each law the number of pages, sentences, words and unique roots in the text; we compute the average length of words and sentences in terms of number of characters, and we count the number of complex words in each law.[18,19] Panel A of Table B.2 shows the descriptive statistics for the text measures for approved laws. On average, laws have 8.17 pages, 166.7 sentences and more than 2,000 words.

4.3 Indicators of legislative quality: Network measures

The second type of indicators to measure the quality of laws focuses on the interconnection between laws through the analysis of the citations network (Bommarito and Katz, 2010).

[18]Complex words are characterized in two ways: first, we consider all words that are not into the Dale-Chall list, a list of 3,000 English words that groups of fourth-grade American students could reliably understand. Second, we count the number of words that have more than three syllables.

[19]Average word and sentence length are the building blocks of various measures of readability of texts, such as the well-known Flesch-Kincaid readability test (Kincaid, Fishburne Jr, Rogers and Chissom, 1975).

Scholars in computational law analyze networks of laws and exploit specific features of them to characterize the efficiency and the complexity of a legal system.[20] We borrow from this literature on computational law the idea of representing the legal system as a network where laws are the nodes and citations are the (directed) links.[21] We develop a set of original indicators with the aim of measuring the quality of laws. These are measures of citations, centrality of laws in the legal network, and codification of laws.

4.3.1 Citation measures

The first set of measures are citation measures. In order to calculate these indices, we construct the network of laws and extract citations toward other legal acts directly from the text. Given the nature of these citations, moving from a *citing* law to a *cited* one, this is a unidirectional network. In the U.S. legal system, a law may be quoted in three distinct ways. First, it may be recalled with the name assigned to it after Congress approval, *i.e.* the *slip law* name (*e.g.* Public law 100-1). Then, the *slip law* is published into the *United States Statutes at Large*, the official source for laws and resolutions passed by Congress, so the law may be quoted with its official denomination in the Statute (*e.g.* 96 Stat. 1781). Finally, a law may be codified and included into one or more sections of the *United States Code* and it may be cited recalling the corresponding section of the United States Code (*e.g.* 49 U.S.C. 2206). We connect the three types of citations and we express all citations in terms of the slip law type.[22] In this way, we generate the first citation measure that captures for each law the number of citations toward other legal acts that are contained in the text (hereafter *out-citations*).[23] This out-citation measure captures the amount of documents that should

[20] See Whalen (2016) for a review of the literature on legal network analysis.

[21] Our analysis is dynamic as we look at the evolution of the network over time.

[22] To connect these three types of citations we conduct a set of additional analyses. First, we trace back citations to the *United States Statutes at Large*. We downloaded the texts of all the *United States Statutes at Large* from congress 82, through the website of the *U.S. Government Publishing Office*, and we link each citation to the statutes with the corresponding slip law name. Second, we connect citations to the *United States Code*. We downloaded the Code in XML and we linked each section of the code to the set of slip laws that composed it over time. Given the dynamic nature of the *United States Code*, we only assign the corresponding slip laws that were in force when the citations took place.

[23] In the construction of this measure we always eliminate duplicated citations so that the measure captures the number of unique laws cited by the text.

be considered to properly understand the meaning and the content of the law.

Table B.2, Panel B, contains the descriptive statistics for these network measures. The average number of out-citations, traced back in terms of slip laws is 70.3. Most laws do not cite directly a slip law, but rather they quote a section of the *United States Code*, whose average value is 14.4, or a law contained in the *United States Statutes at Large*, with an average value of 2.03.

Our second citation measure captures how many times a law has been cited in other future legal acts, the *in-citation* measure. We track citations to each law in the texts of future legal acts and we expand our dataset to more recent congresses so as to have the maximum number of subsequent years.[24] We generate a set of additional in-citations measures that focus on fixed time windows in order to avoid that older laws systematically receive more citations because they were written earlier: in particular, we generate the number of in-citations received in the following 5, 10, 15, or 20 congresses after approval. We also calculate the average number of in-citations per year since approval, as an alternative measure of quality that is independent of time since approval. These in-citation measures capture the extent to which a law is recalled in subsequent congresses. This may be interpreted as an indicator of how relevant or influential a law is for future jurisprudence. This interpretation of our in-citation measure is coherent with the literature studying sentences of the supreme court (Yu and Ash, 2021) but also with studies analysing the quality of patents (Lukach and Lukach, 2007; Benson and Magee, 2015; Jaffe and De Rassenfosse, 2019) and the impact of academic papers (Bornmann, Haunschild and Hug, 2018).

Table B.2, Panel B, shows the descriptive statistics for the in-citations indices: the average number of total in-citations is 56.7, while the average number within the first five congress is 20.9 and it rises progressively as the time window broadens.

[24]In particular, we downloaded and processed the text of all approved laws until the 115^{th} Congress and we trace back all citations types in terms of slip laws. The dataset where we extract in-citation measures is then the full set of laws approved in Congresses 93-115.

4.3.2 Centrality measures

The second set of indicators are centrality measures of laws. In network theory the concept of centrality captures how many connections a node has to other nodes and how central is its position in the network (Schweitzer, Fagiolo, Sornette, Vega-Redondo, Vespignani and White, 2009). This is a measure of the importance of a node relatively to all the other components of the network (Jackson, 2008). In our setting, a central law is relevant for jurisprudence as it is a reference for many other legal texts. We therefore interpret law centrality as a measure of quality. Given the dynamic nature of our dataset, we are able to track the evolution of centrality over time. This allows us to assess how the importance of a law changes over time, as future laws eventually cite it.

We use two basic centrality indices, which are standard in network theory: *degree centrality* and *betweenness centrality*. Degree centrality captures the fraction of all nodes that are connected to the node and betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes.[25]. We compute these two indicators for our set of laws and we measure the centrality of a node at the time of approval of the law, as well as at subsequent points in time (5 and 10 congresses after approval).[26]

Table B.2, Panel C, shows the descriptive statistics for the centrality indicators of laws at the time of their introduction. Degree centrality has an average value of .99, meaning that, on average, each law is connected to the 1% of the laws in the network. The average value of betweenness centrality is .018. Figure A.2 shows the average value of the two indicators as a function of the number of congresses since the law's enactment. In theory, both measures of centrality may either increase or decrease as the network becomes larger. The increase in the size of the network leads to an increase in the number of nodes, but potentially also to an increase in the number of new citations toward old laws. Hence, degree centrality can either increase or decrease over time. As for betweenness centrality, as the network grows it is more likely that an existing law is now on a path connecting other laws; on the other

[25]Betweenness centrality is the sum of fraction of all pair-shortest paths that pass through that node.

[26]Moreover, to ease results interpretation, we rescaled these centrality indicators multiplying them by one hundred.

hand, the inflow of new laws may generate alternative shortest paths that make old laws less central as bridges. In practice, both degree and betweenness centrality exhibit a strong downward trend – on average, as the network of laws grows, the centrality of a single law decreases.

4.3.3 Codification measures

As a third measure of quality of laws, we focus on the process of codification. After its approval a legal text may be included in one or more sections of the United States Code. The codification of a law and its inclusion in a section of the Code is a signal of relevance and we capture it as an additional measure of legal quality.

With this aim, we measure for each law the number of sections of the United States Code where this has been included. Table B.2, Panel D, shows descriptive statistics for this measure. The average number of sections where a law is included is around 11, but the standard deviation of this indicator is high.

A final measure of legislative quality relies on the idea that legal texts that are modified often are of lower quality. We focus therefore on subsequent modifications to sections of the Code that took place with amendments. First, we extract from the Code data on amendments for every section. Then, we count, for each law included in the Code, the occurrences of amendments in the section in which the law appears. We only count amendments that are introduced after the law has been included in that section.[27] This allows us to construct a dynamic index capturing the total number of amendments involving a law. We interpret this index as a signal of law quality, with high quality laws having lower value of the index. Similarly to the in-citation measures, we create additional indicators that rely on fixed time windows to avoid that older laws are subject to more amendments by construction: in particular, we generate the number of amendments received in the following 5, 10, 15, or 20 congresses after a law's approval.

[27] One drawback of this approach is that we cannot identify which part of the section has been amended. This means that we impute the modification to all the laws that contributed to generate that section until that moment.

Table B.2, Panel D, shows descriptive statistics for these indicators. On average, laws are subject to 40.3 amendments. Unsurprisingly, the number of amendments increases with time since a law's approval: the number of amendments within 5 congresses since a law's approval is about 18, but this number rises to almost 40 if we expand the time window to 20 congresses.

5 Results

5.1 Empirical strategy

The main goal of our analysis is to study the impact of the reform of cosponsorship in the 96$^{\text{th}}$ Congress on legislative production and on the quality of approved laws. The design of the reform allows us to implement a standard difference-in-differences design.

We compare outcomes in the House before and after the 96$^{\text{th}}$ Congress (i.e., comparing the period when the number of cosponsors was limited to 25, to that in which there was no limit on the number of cosponsors), to those in the Senate, where there was no modification in the regulation of cosponsorship. Our main identifying assumption is that there are no other factors affecting differently legislative activity originating from the two chambers at the time of the 96$^{\text{th}}$ Congress. We believe this assumption is plausible for a number of reasons. First, there was no modification in the regulation of cosponsorship in the Senate at that time. Second, the partisan configuration of the legislative and executive branch remained the same between the 95$^{\text{th}}$ and 96$^{\text{th}}$ Congress, with a Democratic President, and a Democratic majority in both chambers. This makes it less likely that there would be a major change in the way bills were drafted because of considerations related to divided government. Finally, in all our empirical specifications, we always include a large number of observable characteristics of bills and sponsors, to capture any time varying factors that may affect the outcomes and are correlated with the timing of the reform.

The second main assumption necessary for our empirical design is that there would have been no differential trends in the pattern of outcome variables between the House and the

Senate in the absence of the reform. This assumption is fundamentally untestable, but the large number of time periods before the reform allows us to examine whether the treatment and control groups were already on diverging trends before the reform took place.

Our outcome variables include indicators of bill quality, including the legislative outcomes (whether the bill passed one or more chambers, whether it became law, etc.) and indicators for the quality of approved laws, using the measures discussed in Section 4 – readability, citations, centrality and codifications indices.

The estimated model is as follows:

$$y_{ict} = \beta_0 + \beta_1 H_c + \beta_2 H_c \cdot Post_t + \gamma_t + \delta_{d(i,c,t)} + \zeta' X_{ict} + \epsilon_{ict} \tag{1}$$

y_{ict} is the dependent variable, of bill or law i, in chamber c and in year t, H_c is a dummy indicating the House, $Post_t$ is a dummy indicating the post-reform period (equal to 1 starting from Congress 96), γ_t represents a year fixed effect, and $\delta_{d(i,c,t)}$ is a set of district-decade or state fixed effects associated with the district or state of the main sponsor.[28] The vector X_{ict} includes the following set of controls: political experience of the sponsor, party, gender and age of the sponsor, ideology of the sponsor (as measured by the two Poole-Rosenthal indicators), and, for laws, subject of the law, as categorized by the Congress archive. Robust standard errors are clustered at the sponsor state-year level.

The difference-in-differences estimator is captured by the coefficient on the interaction term $H_c \cdot Post_t$. Under the assumptions described above, it captures the causal effect of the cosponsorship reform.

We also supplement the simple difference-in-differences specification with an event study design, which allows us to test for the presence of pre-trends in our setting, and to study the dynamic effect of the treatment. The event-study specification is the following:

$$y_{ict} = \beta_0 + \beta_1 H_c + \sum_{s=-3, s\neq -1}^{+5} \beta_{2s} H_c \cdot 1\{t = s\} + \gamma_t + \delta_{d(i,c,t)} + \zeta' X_{ict} + \epsilon_{ict}. \tag{2}$$

[28] For the senate, the fixed effect is a simply a state fixed effect. For the House, we include a district by decade fixed effect, because of the reapportionment of Congressional districts every 10 years.

Here the coefficients of interest are the β_{2s}'s, which capture the differences in the outcome variable both before and after the reform. All other terms are as in model 1.

5.2 Impact on legislative production

In this section we study the impact of the reform on the legislative process, according to the empirical model 1. First, we focus on proposed bills, on their characteristics and on the Congressional process; second, we analyze laws, the legislative procedure that leads to the approval and the their text.

Table 1, shows main results for the impact of the reform on the proposed bills. The first column shows that the reform did indeed raise the number of cosponsors in the House relative to the Senate. The remaining columns in the table show that the reintroduction of co-sponsorship deeply modified the Congressional process, with clear evidence of a tradeoff between the quantity and quality of bills. Bills in the House (relative to the pre-reform period and relative to the Senate) are less likely to fail after their introduction (column 2), and the probability of passing one chamber (column 3), two Chambers (column 4), and being approved as law (column 5) consistently increases. The impact on the probability to be approved as law is quite large, with the difference-in-differences coefficient representing about 39% of the mean dependent variable. Finally, the number of actions a bill undergoes also increases (column 6), suggesting that congressional revision of bills is now more accurate.

The result on the number of actions is not merely a consequence of there being more bills that advance through the legislative process. Even if we focus our attention only on the bills that eventually become laws, we see evidence of a longer and deeper congressional revision. Table 2 shows that the reform led to an increase in the number of actions and to a longer approval process in the House, even though there was no increase in the number of committees.

These results confirm the unconditional evidence from Figures 1 and 2: the elimination of restrictions on cosponsorship dramatically reduces the total number of bills, but the quality of bills improves – more bills advance through the legislative process and become laws. The

Table 1: Effects on bills

Dep. variable:	Number of cosponsors	Fails after introduction	Passes one chamber	Passes two chambers	Became law	Number of actions
	(1)	(2)	(3)	(4)	(5)	(6)
Diff-in-Diff	6.632***	-0.0423***	0.0415***	0.0183***	0.0186***	0.928***
	(0.273)	(0.00633)	(0.00633)	(0.00459)	(0.00437)	(0.127)
N	103196	103196	103196	103196	103196	103196
Dep. var. mean	6.71	0.913	0.084	0.051	0.048	3.66
Year & Sponsor-district FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The specification includes year, sponsor-district fixed effects and the following controls: political experience of the sponsor, party, gender, age of the sponsor, number of duplicates bills and measure of ideology of the sponsor. Robust standard errors, clustered at the sponsor-state-year level, are in parentheses: *: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Legislative measures for approved laws

Dep. variable:	N. of actions	N. of days for approval	N. of committees
	(1)	(2)	(3)
Diff-in-Diff	3.886***	30.17**	0.0652
	(0.879)	(12.06)	(0.0692)
N	4699	4699	4699
Dep. var. mean	19.92	211.94	2.08
Year & Sponsor-district FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: The specification includes year and sponsor-district fixed effects and it controls for party and political experience of the sponsor (number laws sponsored), the nature of the law (public or private law), the subject of the law, the gender and age of the sponsor and the measure of ideology of the sponsor. Robust standard errors, clustered at the sponsor-state-year level, are in parentheses: *: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

reform, therefore, appears to have led to the elimination of a large set of irrelevant bills that usually failed to be approved. Bills also undergo a deeper congressional revision, which likely also leads to an improvement in the quality of the approved laws.

5.3 Legislative quality

Readability of Laws

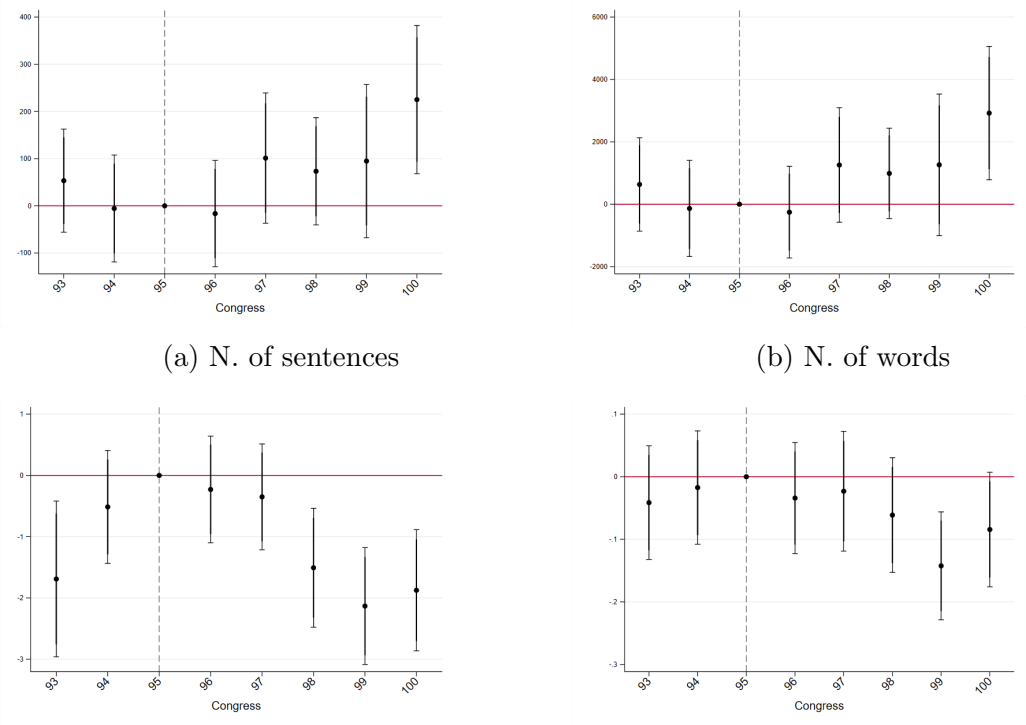
We start our analysis of how the cosponsorship reform affected legislative quality by looking at various measures of readability of the final text of approved laws. The first panel of Table 3 shows that the reform leads to the approval of laws with longer texts: in particular, these have more pages, sentences, words and characters. The second panel of the table, however, suggests that, despite the average increase in the size of texts, the length of sentences and words decreases. Figure 4 shows the event-study plot for four of the key readability measures. There is no evidence of significant pre-trend in the outcomes, but, in line with the results of the simple difference-in-difference analysis, the plots show that the number of words and sentences increases. This does not come at the cost of lower readability, as the number of words per sentence and the average word length fall after the reform.

Table 3: Readability of approved laws

	N. of pages	N. of sentences	N. of words	N. of characters (no stop-words)	N. of characters (with stop-words)
	(1)	(2)	(3)	(4)	(5)
Diff-in-Diff	2.840*	76.71**	1028.8**	7234.9**	11877.4**
	(1.493)	(30.88)	(419.2)	(2957.9)	(4916.5)
N	4699	4682	4682	4682	4682
Dep. var. mean	8.17	166.71	2096.98	14749.35	24566.65
	N. of words per sentences	Av. length of words	Av. length of sentences		
	(6)	(7)	(8)		
Diff-in-Diff	-0.496	-0.0520**	-6.459*		
	(0.311)	(0.0264)	(3.560)		
N	4682	4682	4682		
Dep. var. mean	10.34	6.84	118.07		
Year & Sponsor-district FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: The specification includes year and sponsor-district fixed effects and it controls for party and political experience of the sponsor (number laws sponsored), the nature of the law (public or private law), the subject of the law, the gender and age of the sponsor and the measure of ideology of the sponsor. Robust standard errors, clustered at the sponsor-state-year level, are in parentheses: *: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 4: Event study – readability measures



(a) N. of sentences

(b) N. of words

(c) N. words per sent.

(d) Av. word length

Notes: These plots show the outputs of the DiD estimation according to model 2. The 95th congress works as the benchmark of the analysis. The specification includes year, sponsor-district fixed effects and the following set of controls: political experience, party, gender and age of the sponsor, nature of the law, subject of the law and a measure of ideology of the sponsor. The picture includes 5% and 10% confidence intervals.

Citation measures

We next look at the set of citation measures to study the impact of the reform on legislative quality. Our hypothesis is that higher quality laws will also be more influential, and referred to more often in subsequent legislation. They may also be more connected to previous legislation and cite a larger number of previous laws.

Table 4 shows the results of the simple difference-in-difference analysis (equation 1). We first look at out-citations, the number of previous laws cited in the text (column 1). After the reform, laws cite more past legislation – they have more out-citations. The effect, however, is not statistically different from zero. This suggests that even if approved laws are longer, they do not necessarily become more intricate.

The remaining columns of the table analyze the in-citation patterns of approved laws

before and after the reform. Approved laws after the reform are cited more often by subsequent legislation. This result emerges whether we consider all future congresses (columns 2 and 3), or we restrict attention only to a limited number of future congresses (five, ten, fifteen and twenty in, respectively, columns 4 to 7).[29]

Table 4: Citation measures

Dep. variable:	Out-links total	In-links total	In-links Average (over time)	In-links within 5 congresses	In-links within 10 congresses	In-links within 15 congresses	In-links within 20 congresses
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Diff-in-Diff	14.07	13.82*	0.445**	2.694	8.032*	13.24**	13.91*
	(12.91)	(7.680)	(0.197)	(2.859)	(4.540)	(6.248)	(7.438)
N	4682	4682	4682	4682	4682	4682	4682
Dep. var. mean	70.30	56.78	1.49	20.97	33.97	46.91	55.36
Year & Sponsor-district FE Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

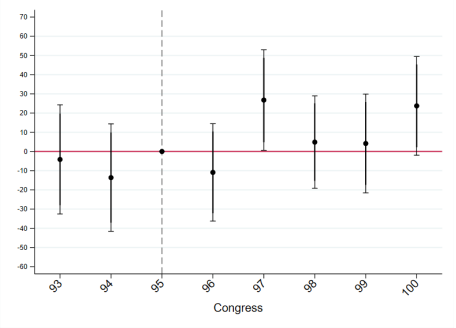
Notes: The specification includes year and sponsor-district fixed effects and it controls for party and political experience of the sponsor (number laws sponsored), the nature of the law (public or private law), the subject of the law, the gender and age of the sponsor and the measure of ideology of the sponsor. Robust standard errors, clustered at the sponsor-state-year level, are in parentheses: *: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 5, shows the event study plots for selected in-citation measures. For all of our measures, there is no evidence of any significant differences between the House and the Senate before the reform – all the pre-reform coefficients are small and statistically insignificant. After the reform, laws drafted in the House become more influential, even though the jump is not immediate. We also see that the increase in in-citations is more pronounced in the medium run (within 10 congresses, panel d in the figure) than in the short run (panel c in the figure).[30] The increase in influence of post-reform laws originating in the House is not a short-run phenomenon – rather, these laws continue to accumulate citations even in the longer run.

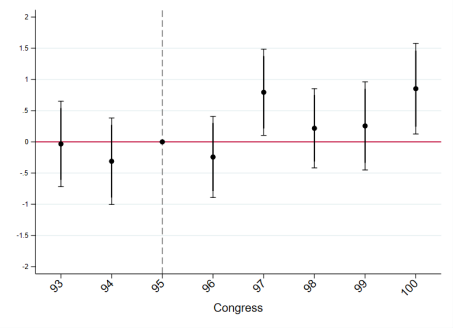
[29]It is important to mention that similar results emerge if we use as dependent variables of the analysis the number of in-citations scaled by the number of characters in the text of the law (expressed in logarithm).

[30]Importantly, event study plots for the dependent variables scaled by the number of characters are similar and are available upon request.

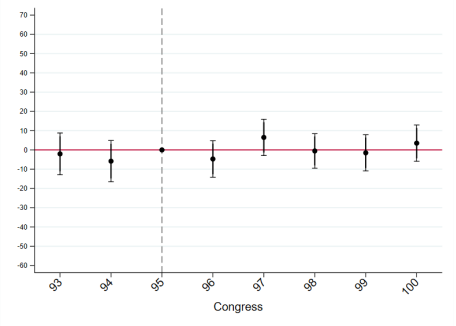
Figure 5: Event study – citation measures: in-links



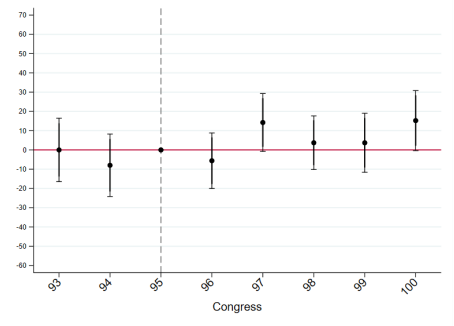
(a) Total



(b) Av. number



(c) Short-run (within 5 congresses)



(d) Medium-run (within 10 congresses)

Notes: These plots show the outputs of the DiD estimation according to model 2. The 95th congress works as the benchmark of the analysis. The specification includes year, sponsor-district fixed effects and the following set of controls: political experience, party, gender and age of the sponsor, nature of the law, subject of the law and a measure of ideology of the sponsor. The picture includes 5% and 10% confidence intervals.

Centrality measures

We next look at network centrality measures as indicators of legislative quality. Our hypothesis is that higher quality laws will be more central in the legal network. Laws are unlikely to become central in the network immediately after enactment, so we also focus on how centrality evolves over time.

Table 5, shows the main results of the difference-in-difference analysis. Panel A focuses on the degree centrality, which captures the fraction of all network nodes that are linked with a law. The effect is positive but not statistically significant if we consider the centrality when the law has been introduced (column 1) or after five congresses (column 2), but it becomes significant after ten congresses (column 3). Similar results emerge when we look at betweenness centrality (Panel B). These results suggest that the reintroduction of co-

sponsorship in the House leads to the approval of laws that are more central in the network and that become more central as the network grows. As with the other measures, these results are indicative of increased legislative quality as a consequence of the reform.

Figure 6 presents the event study plots for degree centrality, at the time of introduction, and after 10 congresses.[31] Consistent with the results in Table 5, we find no evidence of an increase in centrality at the time of approval (panel a), but a consistent positive effects on centrality in the medium run (Panel b). Reasuringly, there is no evidence of differential trends between the House and the Senate prior to the reform.

Table 5: Centrality measures

Dep. variable: centrality measures in **absolute values**			
	Network at the introduction (1)	Network after 5 congresses (2)	Network after 10 congresses (3)
Panel A: **Degree centrality**			
Diff-in-Diff	0.00585	0.128	0.172**
	(0.168)	(0.0965)	(0.0830)
N	4682	4682	4682
Dep. va. mean	1.09	0.744	0.656
Panel B: **Betweenness centrality**			
Diff-in-Diff	-0.00299	0.00310	0.00304**
	(0.0107)	(0.00266)	(0.00144)
N	4682	4682	4682
Dep. var. mean	0.021	0.006	0.003
Year & Sponsor-district FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

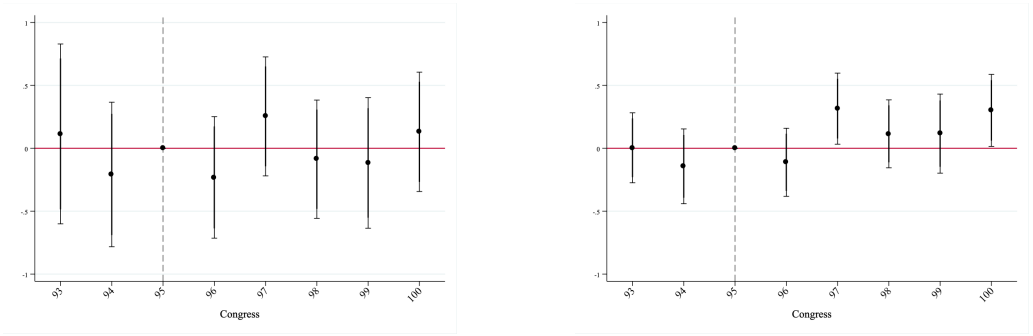
Notes: The specification includes year and sponsor-district fixed effects and it controls for party and political experience of the sponsor (number laws sponsored), the nature of the law (public or private law), the subject of the law, the gender and age of the sponsor and the measure of ideology of the sponsor. Robust standard errors, clustered at the sponsor-state-year level, are in parentheses: *: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Codification measures

Finally, we focus on the codification indicators that capture the extent to which a law is included into the United States Code (USC) and its future modifications. We hypothesize that higher quality laws will be included in more sections of the USC, and will undergo fewer

[31] Event study plots for betweenness centrality are similar, and are available upon request.

Figure 6: Event study – degree centrality



(a) At the introduction (b) After 10 congresses

Notes: These plots show the outputs of the DiD estimation according to model 2. The 95th congress works as the benchmark of the analysis. The specification includes year, sponsor-district fixed effects and the following set of controls: political experience, party, gender and age of the sponsor, nature of the law, subject of the law and a measure of ideology of the sponsor. The picture includes 5% and 10% confidence intervals.

amendments.

Table 6 presents the main results of the difference-in-difference analysis. In panel A, all of the dependent variables are in absolute values, while in panel B they are scaled by the number of sections of the USC in which the laws appear. Column 1 shows that the cosponsorship reform led to more wide-ranging laws, included in more sections of the USC. The effect is statistically significant and fairly large, as it represents about 50% of the dependent variable mean.[32] At the same time, the number of amendments increases, both in total and for fixed time windows (columns 2-4). This is not because laws originating in the House post-reform are included in larger sections of the code (column 5), which are more likely to undergo revisions. On the other hand, when we scale the dependent variable by the number of USC sections in which the laws appear, the effect on number of amendments completely vanishes (Panel B). This suggests that the positive impact in Panel A is completely driven by the fact that laws after the reform enter into more sections and therefore undergo amendments more often.

Figure 7 presents the event-study plot for the number of USC sections in which the law

[32]Furthermore, this result also emerges if we use as dependent variable the number of sections of the USC scaled by the number of words and characters, in logarithm. These results are not shown and are available upon request.

is included. There is no evidence of differential pre-trends, while there is sharp jump in coefficients in the post-reform period. Overall, the results in this section suggest that laws approved in the House after the reintroduction of co-sponsorship are more relevant for the legal system, as they are included into more sections of the USC. Once we adequately take into account the number of sections of the code in which the laws appear, there is no evidence that the number of amendments increases. These results, combined, support the idea that the reform raises the average quality of laws.

Table 6: Codification measures

Dep. variable:	N. sections USC	N. amendments total	N. amendments within 5	N. amendments within 10 congresses	Av. size sections USC congresses
	(1)	(2)	(3)	(4)	(5)
Panel A: **Absolute values**					
Diff-in-Diff	5.782**	17.50*	8.839*	13.93*	-1.754
	(2.647)	(10.12)	(4.611)	(7.160)	(7.513)
N	4682	4699	4699	4699	4682
Dep. var. mean	11.37	40.35	17.95	28.61	53.74
Panel B: **over N. USC**					
Diff-in-Diff	-	-0.376	-0.0315	-0.176	-
		(0.459)	(0.180)	(0.303)	
N		2893	2893	2893	
Dep. var. mean		4.19	1.80	2.95	
Year & Sponsor-district FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

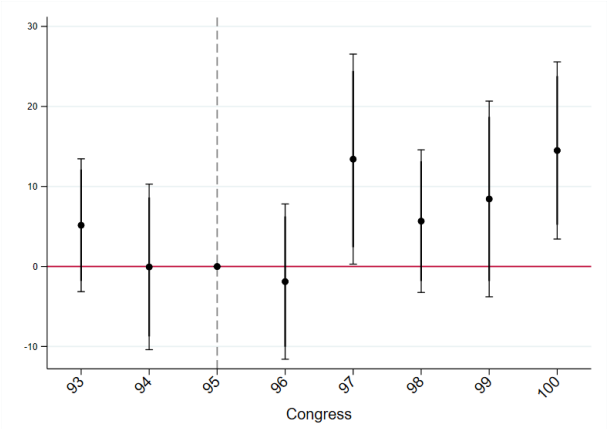
Notes: The specification includes year and sponsor-district fixed effects and it controls for party and political experience of the sponsor (number laws sponsored), the nature of the law (public or private law), the subject of the law, the gender and age of the sponsor and the measure of ideology of the sponsor. Robust standard errors, clustered at the sponsor-state-year level, are in parentheses: *: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6 Mechanisms

In the previous section we have established that removal of the limit on the number of cosponsors in the 96$^{\text{th}}$ Congress led to a significant improvement in the quality of laws originating in the House. In the current section, we aim to shed light on the main mechanisms driving these results.

The model in Section 3 suggests that the reform in the 96$^{\text{th}}$ Congress created higher

Figure 7: Event study – codification measures



Notes: These plots show the outputs of the DiD estimation according to model 2. The 95th congress works as the benchmark of the analysis. The specification includes year, sponsor-district fixed effects and the following set of controls: political experience, party, gender and age of the sponsor, nature of the law, subject of the law and a measure of ideology of the sponsor. The picture includes 5% and 10% confidence intervals.

average quality of laws because of three potential channels.

1. *Reduction of the congestion effect.* The limit on the number of cosponsors created incentive for legislators to draft too many bills, reducing the amount of time that committees can devote to improving the quality of each bill (the extra quality input, captured by the parameter E in the model). Removal of the cap on the number of cosponsors eliminates this congestion effect, resulting in higher quality laws.

2. *Reduction in the amount of frivolous bills.* With limits on the ability to cosponsor, legislators have an incentive to draft bills even if they are of of low quality, given that the floor player cannot act as a perfect filter of quality. This could per-se result in laws of lower quality. Elimination of the cosponsorship cap reduces the amount of frivolous bills.

3. *Better screening.* If the ability of the floor player to screen low quality bills depends on the amount of bills that are brought to the floor, then the elimination of the cosponsorship limit and the resulting decrease in the number of bills may also lead to higher

quality of approved laws.[33]

The first two channels have different predictions for the distribution of quality of approved laws. The congestion hypothesis implies that the lower quality of laws in the pre-reform period derives primarily from the inability of Congress members to adequately revise and improve the initial drafts of proposed legislation. Therefore, elimination of the cosponsorship cap will result in a larger number of high quality laws that have undergone a rigorous revision process. In other words, the reduction of the congestion effect will affect the right tail of law quality.

On the other hand, the frivolous bills hypothesis (coupled with imperfect screening by the floor player) implies that in the pre-reform regime there were many bills of low quality that became law. Elimination of the cosponsorship cap will therefore affect primarily the left tail of the distribution of approved laws.

We can therefore use quantile regressions to disentangle between these two channels. Reduction of the congestion effect will result in an increase in the upper quantiles of the various measures of quality. On the other hand, the reduction in the number of frivolous bills will primarily affect the lower quantiles of the quality indicators. The empirical specification is similar to that of equation 1, but now we use quantile regression instead of OLS:[34]

$$Q_\tau(y_{ict}) = \beta_{0,\tau} + \beta_{1,\tau} H_c + \beta_{2,\tau} H_c \cdot Post_t + \gamma_t + \delta_{d(i,c,t)} + \zeta' X_{ict} + \epsilon_{ict}, \quad \forall \tau \in \{0.3, 0.8\}$$

We estimate this model for the 30^{th} and the the 80^{th} quantile. We expect that the congestion effect channel will affect primarily the 80th quantile, while the frivolous bill channel will affect the 30th quantile. We estimate the model for a variety of different quality measures: the number of in-citations (in total, on average, and within 5 and 10 congresses), degree

[33] In the model, we have assumed that the floor player's screening ability z is constant and does not depend on the number of drafted bills, but this can be easily modified. With constant screening ability, the elimination of the cosponsosrhip limit does not necessarily affect the quality of approved laws through the screening channel, even if the other two channels are still in place.

[34] Compared to model 1, the only difference we bring to the specification, to allow model converge, is that we include simplified fixed effects that always capture the state of the main sponsor.

centrality (overall, and within 5 and 10 congresses), and the number of sections of the USC in which a law appears.

The top panel of Table 7 shows the difference-in-difference coefficients for the regressions modeling the 80$^{\text{th}}$ quantile of the outcome variables. There is some evidence that the cosponsorship reform affected the right tail of the number of in-citations (columns 1 to 4) and the codification measures (column 8), while there is no statistically significant effect on degree centrality, regardless of how we measure it (columns 5 to 7).[35,36] The bottom panel of the Table shows the coefficients for the regressions modeling the 30$^{\text{th}}$ quantile. We find positive but statistically insignificant coefficients for the citation and codification measures, and negative and insignificant coefficients for the centrality measures.

We conclude based on this analysis that the main channel through which legislative quality improves is the reduction of the congestion effect. Removal of the cosponsorship cap allows Congress members to devote more time and effort to improving the quality of proposed legislation, which results in a higher quality of laws that are eventually approved. On the other hand, there is little support for the hypothesis that a large number of laws in the era of limited cosponsorship were pure frivolous bills with no merit.

Table 7: Mechanisms: congestion and frivolous bills hypotheses

Dep. variable:	In-link total	In-link average (over total)	In-link within 5 congresses	In-link within 10 congresses	Degree centr. at the introduction	Degree centr. after 5 congresses	Degree centr. after 10 congresses	N. sections USC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: **80th percentile**								
Diff-in-Diff	15.24	0.463**	5.432	12.99**	-0.230	0.0449	0.0665	2.978**
	(10.01)	(0.220)	(5.200)	(6.234)	(0.231)	(0.142)	(0.108)	(1.417)
N	4682	4682	4682	4682	4682	4682	4682	4682
Panel B: **30th percentile**								
Diff-in-Diff	0.265	0.286	0.389	0.407	-0.00989	-0.00140	-0.00131	0.0310
	(0.671)	(0.423)	(0.564)	(0.496)	(0.0196)	(0.0169)	(0.0139)	(0.114)
N	4682	4682	4682	4682	4682	4682	4682	4682
Year & Sponsor-state FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The specification includes year and sponsor-state fixed effects and it controls for party and political experience of the sponsor (number laws sponsored), the nature of the law (public or private law), the subject of the law, the gender and age of the sponsor and the measure of ideology of the sponsor. Robust standard errors are in parentheses: *: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[35]Similar results emerge if we use different percentile thresholds for this analysis, such as the 90$^{\text{th}}$ or the 75$^{\text{th}}$ percentiles.

[36]We find similar results if we adopt a simpler approach, comparing the best laws (top 20% in terms of quality distribution) before and after the reform in the House, and run a t-test on the two average values.

7 Concluding remarks

This paper provides new theory and new evidence on the importance of signalling incentives of politicians for the quality of laws. Using the cosponsorship reform in the 96th Congress as a quasi-natural experiment, we have discovered that removing incentives to bill proliferation has dramatically improved the quality of laws, in terms of citation, network centrality and codification measures. Moreover, the main mechanism seems to be the removal of a significant congestion effect in the legislative revision process.

In future research it could be useful to explore the effects of other reforms that presumably affect legislative incentives of politicians. The greater opportunities of politicians to signal positions, ideology, loyalty etc using new social media and targeted communication may reduce the need to do so in the legislative process. On the other hand, the polarization effects of debates on social media may affect downwards the likelihood of bipartisan legislation. Hence the net effect of the new technologies on legislative incentives is not obvious, and hence ultimately an important empirical question.

References

Adler, E Scott and Wilkerson, John D (2013). *Congress and the politics of problem solving*: Cambridge University Press.

Benson, Christopher L and Magee, Christopher L (2015). "Quantitative determination of technological improvement from patent data," *PloS one*, 10 (4), e0121635.

Boella, Guido, Di Caro, Luigi, Violato, Andrea, and Robaldo, Livio (2014). "Exploiting networks in Law," in *International Conference on Language Resources and Evaluation (lrec)*, 1654–1658.

Bommarito, Michael J. and Katz, Daniel M. (2010). "A mathematical approach to the study of the United States Code," *Physica A: Statistical Mechanics and its Applications*, 389 (19), 4195–4200.

Bornmann, Lutz, Haunschild, Robin, and Hug, Sven E (2018). "Visualizing the context of citations referencing papers published by Eugene Garfield: A new type of keyword co-occurrence analysis," *Scientometrics*, 114 (2), 427–437.

Box-Steffensmeier, Janet M, Christenson, Dino P, and Craig, Alison W (2019). "Cue-taking in congress: interest group signals from dear colleague letters," *American Journal of Political Science*, 63 (1), 163–180.

Campbell, James E (1982). "Cosponsoring legislation in the US Congress," *Legislative Studies Quarterly*, 415–422.

Cooper, Joseph and Young, Cheryl D (1989). "Bill introduction in the nineteenth century: A study of institutional change," *Legislative Studies Quarterly*, 67–105.

Foarta, Dana and Morelli, Massimo (2021). "Complexity and the reform process."

Fowler, James H (2006). "Legislative cosponsorship networks in the US House and Senate," *Social networks*, 28 (4), 454–465.

Fowler, James H and Jeon, Sangick (2008). "The authority of Supreme Court precedent," *Social networks*, 30 (1), 16–30.

Fowler, James H, Johnson, Timothy R, Spriggs, James F, Jeon, Sangick, and Wahlbeck, Paul J (2007). "Network analysis and the law: Measuring the legal importance of precedents at the US Supreme Court," *Political Analysis*, 15 (3), 324–346.

Gagliarducci, Stefano and Paserman, M Daniele (2022). "Gender Differences in Cooperative Environments? Evidence from The U.S. Congress," *The Economic Journal*, 132 (641), 218–257, 10.1093/ej/ueab069.

Gratton, Gabriele, Guiso, Luigi, Michelacci, Claudio, and Morelli, Massimo (2021). "From Weber to Kafka: Political Instability and the Overproduction of Laws," *American Economic Review*, 111 (9), 2964–3003.

Jackson, Matthew O (2008). "Average distance, diameter, and clustering in social networks with homophily," in *International Workshop on Internet and Network Economics*, 4–11, Springer.

Jaffe, Adam B and De Rassenfosse, Gaétan (2019). "Patent citation data in social science research: Overview and best practices," *Research handbook on the economics of intellectual property law*.

Jaffe, Adam B and Trajtenberg, Manuel (2002). *Patents, citations, and innovations: A window on the knowledge economy*: MIT press.

Katz, Daniel Martin and Bommarito, Michael James (2014). "Measuring the complexity of the law: the United States Code," *Artificial intelligence and law*, 22 (4), 337–374.

Kessler, Daniel and Krehbiel, Keith (1996). "Dynamics of cosponsorship," *American Political Science Review*, 90 (3), 555–566.

Kincaid, J Peter, Fishburne Jr, Robert P, Rogers, Richard L, and Chissom, Brad S (1975). "Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel,"Technical report, Naval Technical Training Command Millington TN Research Branch.

Koger, Gregory (2003). "Position taking and cosponsorship in the US House," *Legislative Studies Quarterly*, 28 (2), 225–246.

Koniaris, Marios, Anagnostopoulos, Ioannis, and Vassiliou, Yannis (2018). "Network Analysis in the Legal Domain: A complex model for European Union legal sources," *Journal of Complex Networks*, 6 (2), 243–268.

Krehbiel, Keith (1995). "Cosponsors and Wafflers from A to Z," *American Journal of Political Science*, 906–923.

Laband, David N, Seals, Richard A, and Wilbrandt, Eric J (2015). "On the importance of inequality in politics: duplicate bills and bill co-sponsorship in the US House of Representatives," *Economics of Governance*, 16 (4), 353–378.

Lewis, Jeffrey B, Poole, Keith, Rosenthal, Howard, Boche, Adam, Rudkin, Aaron, and Sonnet, Luke (2021). "Voteview: Congressional roll-call votes database," *See https://voteview.com/*.

Lukach, Ruslan and Lukach, Maryna (2007). "Ranking USPTO patent documents by importance using random surfer method (pagerank)," *Available at SSRN 996595*.

Lyte, Alexander, Slater, David, and Michel, Shaun (2015). "Network Measures of the United States Code,"Technical report, MITRE CORP MCLEAN VA.

Matthews, Donald R (1960). *U.S. Senators and Their World*: New York: Random House.

Mayhew, David R (2004). *Congress: The electoral connection*: Yale university press.

Mazzega, Pierre, Bourcier, Danièle, and Boulet, Romain (2009). "The network of French legal codes," in *Proceedings of the 12th international conference on artificial intelligence and law*, 236–237.

Narin, Francis et al. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity*: Computer Horizons Cherry Hill, NJ.

Narin, Francis (1994). "Patent bibliometrics," *Scientometrics*, 30 (1), 147–155.

Oleszek, Mark J (2018). "Sponsorship and cosponsorship of Senate bills," *Congressional Research Service*, 98–279.

Schweitzer, Frank, Fagiolo, Giorgio, Sornette, Didier, Vega-Redondo, Fernando, Vespignani, Alessandro, and White, Douglas R (2009). "Economic networks: The new challenges," *science*, 325 (5939), 422–425.

Thomas, Scott and Grofman, Bernard (1993). "The effects of congressional rules about bill cosponsorship on duplicate bills: Changing incentives for credit claiming," *Public Choice*, 75 (1), 93–98.

Volden, Craig and Wiseman, Alan E (2014). *Legislative effectiveness in the United States congress: The lawmakers*: Cambridge university press.

Waltl, Bernhard and Matthes, Florian (2014). "Towards measures of complexity: Applying structural and linguistic metrics to German laws," in *Legal knowledge and information systems*, 153–162: IOS Press.

Whalen, Ryan (2016). "Legal networks: The promises and challenges of legal network analysis," *Mich. St. L. Rev.*, 539.

Wilson, Rick K and Young, Cheryl D (1993). "Cosponsorship of Legislation in the US Senate," in *annual meeting of the Midwest Political Science Association, Chicago*.

——— (1997). "Cosponsorship in the US Congress," *Legislative Studies Quarterly*, 25–43.

Wilsona, Neda Sakhaeea Mark C and Zakerib, Golbon (2016). "New Zealand legislation network," in *Legal Knowledge and Information Systems: JURIX 2016: The Twenty-Ninth Annual Conference*, 294, 199, IOS Press.

Yu, Tinghua and Ash, Elliott (2021). "Polarization and political selection."

A Appendix 1: Additional figures

Figure A.1: Frequency of topics discussed by laws.

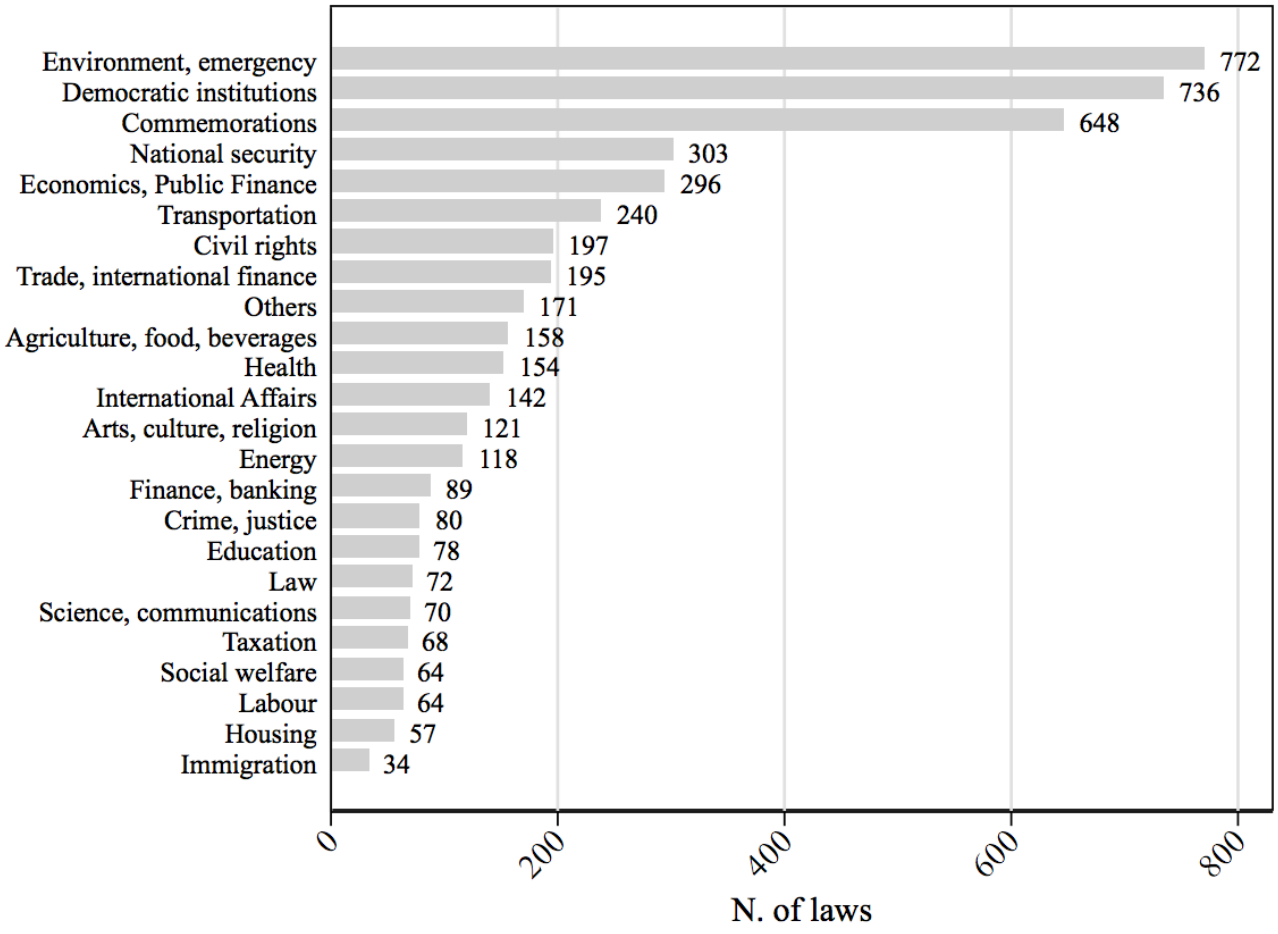
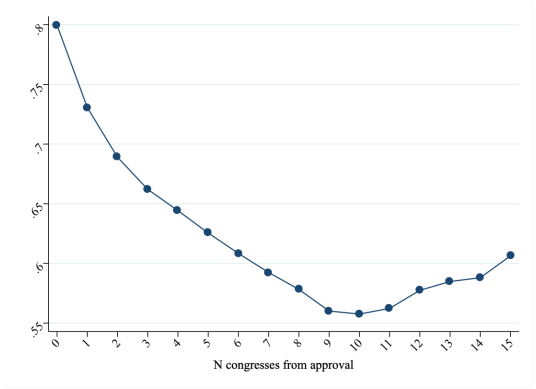
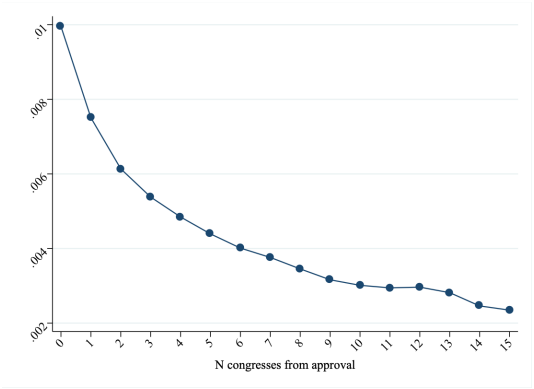


Figure A.2: Centrality of laws in the network.



(a) Degree centrality



(b) Betweeness centrality

B Appendix 2: Additional tables

Table B.1: Descriptive statistics - Dataset

	Average value	Standard dev.	Min	Max
Panel A: Bills				
House	.751	.432	0	1
Senate	.249	.432	0	1
Num. subjects	6.07	7.10	0	66
Num. actions	3.67	7.12	1	396
Num. cosponsors	6.71	22.56	0	367
Introduced only	.913	.281	0	1
Passed one chamber	.085	.279	0	1
Passed two chambers	.052	.222	0	1
Became law	.048	.214	0	1
Num. duplicates bills	3.37	8.36	1	132
N	117,456			
Panel B: Laws				
House	.600	.489	0	1
Senate	.399	.489	0	1
Num. subjects	10.04	13.47	0	66
Num. actions	19.92	20.06	5	355
Num. cosponsors	20.01	47.70	0	301
Num. days approval	211.94	173.56	0	735
Gender sponsor (female)	.032	.177	0	1
Age sponsor	57.39	10.65	30	88
Num. committees	2.08	1.20	0	27
N	4,934			

Notes: The variables *introduced only*, *passed one chamber*, *passed two chambers* and *became law* are dummy. *Num. words not DC* counts the number of words that belong to the Dale-Chall list. *Num. words poly-syllables* counts the number of words with more than three syllables.

Table B.2: Descriptive statistics - Quality measures of laws

	Average value	Standard dev.	Min	Max
Panel A: Text measures of laws				
Nun. pages	8.17	30.30	1	879
Nun. sentences	166.71	611.77	3	12,916
Nun. words	2,096.98	8,309.62	37	211,111
Nun. tokens	337.78	583.98	32	10,937
Av. length words	6.84	.383	5.09	8.12
Av. length sentences	118.07	48.60	28.87	690
Num. words not-DC	1,373.53	5,537.80	24	144,390
Num. words poly-syllables	541.57	2,076.94	3	44,771
N	4,934			
Panel B: citations measures				
Out-citations to USC	14.40	51.51	0	1,632
Out-citations to public laws	.813	4.72	0	200
Out-citations to US Statutes at Large	2.03	7.61	0	148
Out-citations to public laws (total)	70.30	206.57	0	2,982
Av. N. sentences (into cited laws)	457.52	783.17	0	11,704
Av. N. words (into cited laws)	6,032.42	10,506.61	0	154,748
Av. N. token (into cited laws)	588.50	707.23	0	7,405
Av. N. words not-DC (into cited laws)	3,941.32	6,885.16	0	99,100
Av. N. words poly-syllables (into cited laws)	1,565.24	2,677.21	0	40,113
Av. N. out-citations (into cited laws)	150.07	227.01	0	2,544
In-citations (total)	56.78	113.36	0	1,587
In-citations within 5 con.	20.97	41.48	0	532
In-citations within 10 con.	33.97	67.13	0	899
In-citations within 15 con.	46.91	93.59	0	1,372
In-citations within 20 con.	55.36	110.38	0	1,587
N	4,934			
Panel C: centrality measures				
Degree centrality (at the introduction)	.990	2.39	0	34.76
Betweenness centrality (at the introduction)	.018	.123	0	3.18
N	4,934			
Panel D: codification measures				
N. sections USC (total)	11.37	47.38	0	1,172
N. amendments (total)	40.35	199.15	0	5,393
N. amendments within 5 con.	17.95	91.24	0	2,578
N. amendments within 10 con.	28.61	140.97	0	4,017
N. amendments within 15 con.	36.33	179.61	0	4,975
N. amendments within 20 con.	39.94	197.00	0	5,391
Av. size sections USC	53.74	101.16	0	1,714
N	4,934			

Notes: *out-citations to USC/public laws/US Statutes at Large* measures, respectively the total number of citations in the texts of laws to sections of the USC, to public laws and to parts of the US Statutes at Large. *out-citations (total)* measures the total number of citations expressed in terms of slip law type. The values of the centrality measures are reported as in the congress of introduction of the law. *N. sections USC* measures the number of sections of the USC where the law has been included.