WILLIAM HOWELL
University of Wisconsin
SCOTT ADLER
University of Colorado
CHARLES CAMERON
Columbia University
CHARLES RIEMANN
University of Connecticut

Divided Government and the Legislative Productivity of Congress, 1945–94

This paper contributes to the literature on divided government and legislative productivity. We begin by reexamining Mayhew's data on landmark enactments. We show that Mayhew's claim that divided government does not affect legislative productivity is a consequence of aggregating time series that exhibit different behavior. We then extend Mayhew's analysis by broadening the concept of significance and creating a new four-category measure that encompasses *all* 17,663 public laws enacted in the period of 1945–94. Using appropriate time-series techniques, we demonstrate that periods of divided government depress the production of landmark legislation by about 30%, at least when productivity is measured on the basis of contemporaneous perceptions of legislative significance. Divided government, however, has no substantive effect on the production of important, albeit not landmark, legislation and actually has a positive effect on the passage of trivial laws.

Introduction

What determines the legislative productivity of Congress? A long tradition in American political science identifies political parties as the glue that holds together the institutions so carefully separated by the founders (Fiorina 1980; Ford 1898; Key 1964; Schattschneider 1942; Sundquist 1988; Wilson 1961). In this view, high productivity is associated with unitary party government directed by a vigorous president (recall the New Freedom, New Deal, and Great Society Congresses). Conversely, divided government is thought to lead to gridlock, paralysis, and legislative slumps.

In the wake of David Mayhew's *Divided We Govern* (1991), the effects of divided government have attracted new attention. In his innovative and influential work, Mayhew inventoried landmark legislation enacted between 1947 and 1990. Examining patterns for each Congress, he concluded that periods of unified government do not correlate with surges in legislative productivity and that periods of divided government do not necessarily lead to legislative gridlock. Thus, he reasoned, "divided we govern."

Mayhew's analysis, contradicting the received wisdom of several generations of political scientists, has stimulated much new work. Alternative measures of legislative productivity and more refined models of lawmaking have emerged (Binder 1999; Edwards, Barrett, and Peake 1997; Jones 1994; Krehbiel 1996; Stimson, MacKuen, and Erikson 1995). Additionally, authors have challenged Mayhew's methods of calculating important legislation as well as his findings concerning divided government (Coleman 1999; Kelly 1993).

In keeping with this line of critique, our paper contributes to the new literature on divided government and, more broadly, legislative productivity. First, we reexamine Mayhew's data on landmark enactments. We show that Mayhew's data are not stationary. Without accounting for this feature of the data, we cannot know what effect divided government has on legislative productivity. Taking into account the nonstationarity of the data, we show that the no-effect finding is a consequence of aggregating time series that have different behavior. After making appropriate econometric adjustments, we find that Mayhew's revamped "Sweep One" series—measuring contemporaneous perceptions of landmark legislation—indicates about a 30% reduction in legislative productivity during periods of divided government. In contrast, laws identified as landmark only by retrospective evaluations show a significant increase during periods of divided government. We suggest that the former series is more appropriate for studying the politics of congressional productivity.

We then extend Mayhew's analysis by broadening the concept of significance and shifting from the use of a dichotomous variable to one that differentiates between multiple levels of legislative importance. We create new measures to encompass *all* 17,663 public laws enacted in the period of 1945–94, not just landmark enactments. In doing this, we pay particular attention to methodological problems in scoring legislative significance and provide several independent checks on the validity of the measures. The most important class, the Group A series, corresponds closely to Mayhew's series. Our Group B series addresses the legislative productivity of Congress in nonlandmark but nonetheless highly consequential legislation. The Group C and D series (ordinary and minor enactments, respectively) then allow a comprehensive study of legislative productivity. Using a variety of measures, we show that the depressing effect of divided government attenuates as one moves from the most significant to less significant legislation and actually reverses at the lowest levels of significance. We conclude by discussing the applications for this new measure in more theoretically motivated studies of legislative productivity.

The Divided Government Effect in the Mayhew Data

Construction of the Mayhew Series

To test whether periods of unified government are more productive than periods of divided government, Mayhew compiled two lists of landmark enactments. The first, Sweep One, attempted to tap into enactments' contemporaneous *political* significance. Accordingly, the procedure relied on reportage in the annual roundup stories in the *New York Times* and *Washington Post*. These stories summarize the most important legislative accomplishments of the session in the opinion of the newspapers' editors and Washington correspondents. In order for a law to make it onto the Sweep One list, the roundup authors had to declare it an outstanding legislative accomplishment, not merely of that session, but of any session.

The second list, "Sweep Two," attempted to capture enactments' retrospective *policy* significance. To compile the Sweep Two list, Mayhew culled policy histories to identify landmark enactments in different policy arenas. Because of the necessary lag associated with retrospective judgments, Mayhew ended Sweep Two with the 99th Congress (1985–86). For the 100th–103d Congresses, Mayhew relied exclusively on the Sweep One methodology.

To test the divided government hypothesis, Mayhew employed the union of Sweep One and Two laws. The use of the union raises issues concerning the aggregation of data with different behavior. After discussing the appropriate statistical techniques to analyze these time series, we highlight problems caused by aggregation.

Nonstationarity of the Mayhew Data

Problems of nonstationarity often plague time-series data. If a series is stationary, then the mean, variance, and autocorrelations are time invariant. Ordinary Least Squares (OLS) regressions or standard

Autoregressive Moving-Average (ARMA) models are then appropriate. But when a time series is not stationary, t and F-tests generated from OLS regressions may be seriously misleading.

Nonstationary time-series data can be either trend or difference stationary. Trend stationary data can be made stationary by removing the deterministic trend; difference stationary data can be made stationary by taking differences. To check for nonstationarity in the Mayhew data, we apply a battery of augmented Dickey-Fuller tests (Enders 1995, 256–58).

The diagnostics indicate that neither the Sweep One series (the contemporaneously politically significant laws), nor the Sweep Two series (the retrospectively policy-significant laws), or even the union of the two series, are stationary.[1] DeBoef and Granato (1997) show that it is difficult to determine with absolute certainty whether or not a finite series such as this is stationary. However, they also affirm that the correct treatment of near-integrated series is to regard them as if they are integrated. Further analysis reveals these series as all polynomial trend stationary.[2]

Since the data are nonstationary, a t-test for the difference in means between periods of unified versus divided government is inappropriate because it will largely reflect the distribution of periods of divided government relative patterns in the overall time series, rather than the true effect of divided government. The results of multiple-regression analyses of the data will also be meaningless unless the regressions correct for the nonstationarity of the data so that the residuals are made stationary (Enders 1995, 216–20; Granger and Newbold 1974). To date, no discussion of the productivity of Congress has recognized the problems created by the nonstationarity of the Mayhew series. However, without accounting for this feature of the data, we simply cannot know what effect divided government has on legislative productivity.

We reanalyze Mayhew's data using appropriate methods for time-series data. As independent variables, we include an appropriately fitted polynomial in time along with a dummy for divided government.[3] This procedure renders the data trend stationary and thereby enables us to examine the direct effect of divided government. After ensuring that the residuals of this model are stationary, we run Poisson regressions (which are more appropriate for event-counts) and report our results in Table 1. OLS regressions generate virtually identical results, though they do not fit the data as well.

The first column of Table 1 replicates Mayhew's no-effect finding using the union of the contemporaneously and retrospectively judged

TABLE 1
Mayhew Data Reexamined
(standard errors in parentheses)

Variables	S1 ∪ S2	S1	S2-only
Unified Government	0.03	0.25**	–0.63**
	(0.13)	(0.14)	(0.29)
Time	0.21**	0.11**	0.32**
	(0.05)	(0.04)	(0.12)
Time^2	–0.01**	-0.01**	–0.01**
	(0.002)	(0.00)	(0.01)
Constant	1.59**	1.55**	–0.08
		(0.26)	(0.26)
(0.62)			
n	20	24	20
Residual Deviance	24.55	19.82	27.04

Note: Results generated from Poisson regressions.

*significant at the .1 level; **significant at the .05 level.

significant legislation (Sweep One and Sweep Two data) for the years in which both are available.[4] Because the polynomial in time de-trends the data—i.e., the residuals from the regression are made stationary, as indicated by a Dickey-Fuller test—the indicated t-statistics are nonspurious. Unified government does not approach statistical significance when the dependent variable is the union of the Sweep One and Sweep Two data.

Aggregation Effects

To investigate aggregation effects, we separate the dependent variable into two discrete groups that together constitute the entirety of Mayhew's time series: the Sweep One laws (most of which are included in the Sweep Two list as well) and the Sweep Two-only laws. The latter group consists of laws that were not afforded landmark status by contemporaneous observers but ultimately reached that level of significance with the passage of time, at least in the retrospective judgments of policy historians. Notable examples include the Motor Vehicle Air Pollution Control Act of 1966, the outlawing of the Communist Party in 1954, and the Gun Control Act of 1968. These two series are plotted in the lower panel of Figure 1 (as an aid in visualizing the data, dotted lines show the fit from a locally weighted regression). Since the Sweep Two data ends in 1986, the Sweep Two-only series is somewhat shorter than the Sweep One series.

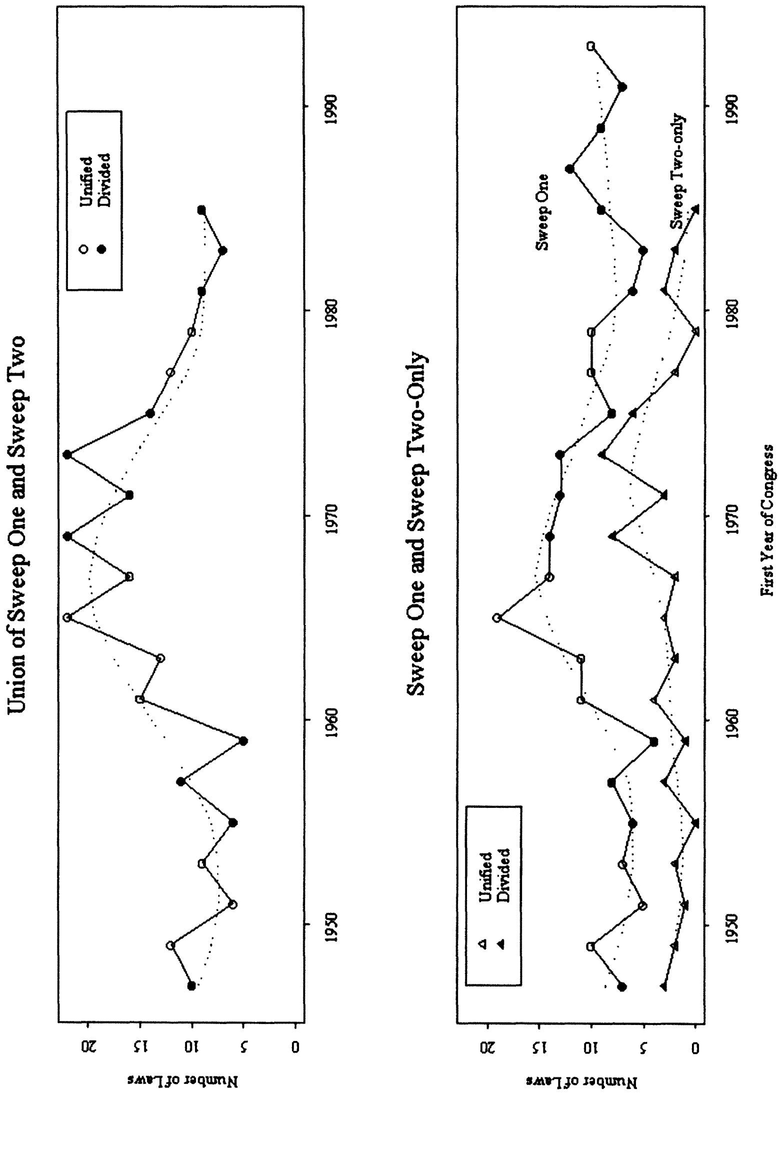
Columns 2 and 3 present the results for the Sweep One and Sweep Two-only groups. Again, the polynomials in time de-trend the data. As shown, unified government significantly increases the production of Sweep One laws, by about three laws per Congress. In contrast, however, unified government significantly *decreases* the production of Sweep Two-only laws, a reverse effect, by about two laws per Congress. Robust regressions (not shown) generate the same results, suggesting they are not the consequence of a few outliers.

The proximal explanation for the no-effect finding is now clear: when the Sweep One and Sweep Two data are added together, the positive effects of divided government in the retrospectively judged significant (Sweep Two only) laws cancel out the negative effect in the contemporaneously judged significant (Sweep One) laws. The no-effect finding results from aggregating two rather different series.

Is it plausible that divided government stimulates the production of laws judged to be extremely important *only in retrospect*? Perhaps the furor created by divided government systematically obscures the perceptions of contemporaneous observers. But at least equally plausible is the possibility of selection bias in the policy histories. Most of the histories Mayhew relied upon are of relatively recent vintage—only two of the retrospectives were published before 1975 and none before 1968. While they survey longer periods, their focus on modern policy naturally leads them to dwell on the more recent past, when *divided government is more common*. Consequently, the Sweep Two-only laws may display a spurious surge during divided government. If this selection bias were at work in the data, one might expect to see the reverse effect only in the latter part of the series—and this is in fact the case.[5]

Conclusions based on such short series must be tentative. In fact, the Sweep Two-only reverse effect may simply be spurious, a consequence of a relative handful of observations in an abbreviated series. The reverse effect, indeed, depends upon large law counts in the 91st and 93d Congresses, both of which were characterized by divided government. If the Sweep Two-only (retrospectively significant) effect is artificial, it should vanish when longer series become available and

FIGURE 1
The Mayhew Series



perhaps in studies of particular policy arenas. We defer such studies to the future.

Even were better data from retrospective policy histories available, we believe congressional scholars should focus on Sweep One-type data. Such data, when well constructed, will reflect the viewpoint of the actors who actually created the politics of a particular Congress and whose behavior is the subject of theories of congressional politics. In contrast, the retrospective data, when they are at variance with data based upon contemporaneous judgments, reflect the impact of subsequent events and the possibly different sensibilities of later chroniclers. The contemporaneous data, therefore, seem more appropriate for positive studies of the historical politics of Congress.

New Measures of Legislative Productivity

Data

Mayhew's data on highly significant legislation paint only a partial portrait of the legislative output of the post-WWII Congresses. In this section, we complete the portrait by presenting law counts across the entire range of legislative significance. We create a four-category measure of legislative significance that allows for a comprehensive account of legislative productivity ranging from "average" or "typical" laws to those enactments of extraordinary significance. To create such a measure, we rely on source reporting from three corps of elite observers: the Capitol Hill reporters of the *Washington Post*, those of the *New York Times*, and the writers and editors of the *Congressional Quarterly* (*CQ Almanac*). We focus exclusively on mentions of public laws, excluding constitutional amendments, resolutions, treaties, nominations, and the like. We do so because making treaties (for instance) involves a radically different constitutional "game" from enacting laws and probably requires separate analysis.

Each mentioned public law from each source was positively identified and cross-classified with citations in other sources. The underlying data for the study, therefore, are not simple counts for each year or Congress but cross-classified public laws.

To collect mentioned public laws in the two newspapers, we coded three kinds of stories in each year between 1945 and 1994: the roundup story summarizing the legislative session, evaluative editorials, which usually enumerate the accomplishments and shortcomings of the session, and the "closing" story announcing the end of the legislative session. The latter often contains reviews of the session's major enactments. They also contain mentions of less significant legislation enacted at the end of the session, but the cross-validation requirement discussed below screens out such bogus mentions.[6] The stories were identified by scanning each page of the newspapers in the weeks surrounding the closing date of the session.[7]

From 1948 to 1987, the authors of the *CQ Almanac*, our third reporting source, compiled their own list of noteworthy enactments, included in a summary section along with capsule descriptions of the legislation. Since the *Almanac*'s summary *(CQ Summary)* was culled in an obvious way from stories in the body of the *Almanac*, constructing a comparable summary for the three years prior to 1948 and for the years after 1987 proved to be straightforward (see note 12). The mentions in the summary were coded using the same rules employed for the newspapers.

We make extensive use of Mayhew's Sweep One laws, discussed earlier. Mayhew's data cover the 80th–103d Congresses (1947–94); we compiled a similar list for the 79th Congress, using the same sources and method.[8]

In all, 2,458 separate public laws were mentioned in the sources.[9] (A model for source reporting and the issues surrounding source consistency are treated in the Appendix.) In the period of 1945–94, Congress enacted 17,663 public laws, so fewer than 14% were mentioned in the sources. In an average year, the *CQ Summary* reported about three times as many enactments as either of the newspapers, which typically reported about the same number of stories (median *CQ Summary* = 43, median *Washington Post* = median *New York Times* = 15). Both newspapers report about four of Mayhew's non-Sweep One public laws for every Sweep One law they mention (median Sweep One = 4).[10]

Classifying Enactments by Legislative Significance

Given the analysis reported in the Appendix, we define four significance classes:

1. Group A—Landmark enactments: Mayhew's Sweep One public laws[11]
2. Group B—Major enactments: all other public laws mentioned in either the *New York Times* or *Washington Post* and greater than or equal to six pages in coverage in the *CQ Almanac*
3. Group C—Ordinary enactments: all other public laws mentioned in the *CQ Summary*[12]
4. Group D—Minor enactments: all remaining public laws, including commemorative legislation

TABLE 2
Measures of Enactments by Legislative Significance,
79th–103d Congresses

Congress	Year	Group A	Group B	Group C	Group D	A-Sum	B-Sum	A1
79	1945	7	8	79	639	72.5	84.5	10
80	1947	6	3	51	846	75	26.4	3
81	1949	9	11	83	818	117.7	112.2	13
82	1951	5	6	65	518	52.9	75.6	4
83	1953	6	6	51	718	56.9	55	8
84	1955	6	7	37	978	55	54.2	4
85	1957	8	2	71	855	62.5	13.5	3
86	1959	4	10	73	713	66	83.2	4
87	1961	10	15	83	777	154.8	182.1	14
88	1963	10	11	67	578	137.4	112	6
89	1965	19	18	146	627	271.5	178	22
90	1967	14	21	96	509	181.7	240	17
91	1969	13	22	84	577	179.8	261.2	14
92	1971	10	15	62	520	127.2	199.7	10
93	1973	13	15	133	488	155.9	172.5	10
94	1975	8	21	61	497	73.3	235	4
95	1977	13	18	49	553	141	204.5	13
96	1979	10	8	69	526	92.1	101.7	6
97	1981	6	11	52	404	117.4	180.1	7
98	1983	5	5	83	530	88.2	56.9	2
99	1985	9	6	72	577	116.4	166.7	7
100	1987	11	12	53	637	139	226.6	11
101	1989	8	16	56	570	197	207.9	12
102	1991	6	16	51	517	78.5	184.7	10
103	1993	10	9	43	403	83.3	84.2	5
Total		226	292	1770	15375	2893	3498.4	219

The threshold for Group B has been set high enough to ensure comparable data over time, while the definitions of Groups C and D do not hinge on censored data.

To avoid possible bias caused by huge omnibus bills, we also calculate a supplementary measure for Groups A and B based on the sum of the pages in *CQ* devoted to describing the laws rather than the count of laws. These two measures are referred to as Group A-sum and Group B-sum. Additionally, we define an alternative to Mayhew's Sweep One laws based on a page-length threshold. Group A1 consists of public laws other than appropriations mentioned in *all* three sources and greater than or equal to eight pages in length in *CQ*.[13] This group is about the same size as Mayhew's Sweep One (Group A) but contains a somewhat different set of laws. It seems likely that this very restrictive category taps into significance at the highest levels. We employ this category to test the robustness of Mayhew's findings about legislation at the highest levels of significance. Table 2 displays the aggregate data by Congress from 1945 to 1994.

Are the Significance Classes Actually Different?

Do Groups A–D actually tap into different levels of legislative significance? If not, no further analysis is warranted. We take three different approaches to this question.

Table 3 addresses face validity. The table provides descriptions of five laws chosen at random from Groups A–C.[14] The laws in the Group A sample are indeed towering oaks within the legislative forest. Two—the War Powers Act and the Reagan Tax Cut—are arguably sequoias. The Group B sample contains laws of genuine noteworthiness in their day but no oaks and certainly no sequoias. The Group C sample is clearly composed of ordinary legislation.

Mayhew's Sweep Two laws, those enactments judged most consequential in retrospective policy histories, provide a potential second check on whether our measures capture some notion of legislative significance. We argue that policy significance is often an important component of perceived political significance. If so, and if the categorization sorts laws appropriately, then the Sweep Two laws (those judged significant in retrospective policy histories) should load disproportionately into the upper groups. As shown in Table 4, almost two-thirds of the Group A laws are also Sweep Two laws.[15] This percentage sharply declines as one moves from landmark enactments to minor laws, in which Sweep Two laws exist only as a trace element. About 85% of all Sweep Two laws are found in Groups A and B. (Group C contains about the same number of Sweep Two laws as Group B, but Group C is almost six times the size of Group B.)

The average length of *CQ* coverage in the four groups affords another check on their validity. By construction, Group B contains no laws with fewer than six pages of coverage in the *CQ Almanac*. No such restriction is imposed on the Group A laws. If there were little difference in the actual significance levels in the two groups, the average length of coverage in Group A should be shorter than that in Group B. In fact, the median length of Group A stories is 10.0 pages,

TABLE 3
Sample Laws from Groups A, B, and C

Public Law	Year	Description	*CQ* Length
Group A: Landmark Legislation			
85-686	1958	Trade Agreements Extension Act of 1958: strengthened presidential authority to make agreements with other countries, extending that authority for four years	11
89-563	1966	Federal Traffic Safety Act: Nader-inspired safety standards	10
93-148	1973	War Powers Act: asserted congressional authority over war-making	11.5
97-34	1981	Economic Recovery Act of 1981: The "Reagan Revolution's" tax component, which created enormous continuing deficits	14
103-322	1994	Omnibus Crime Act of 1994: $30 billion crime bill with assault weapon ban	21.5
Group B: Major Legislation			
87-61	1961	Highway Act of 1961: increased highway use taxes to ensure completion of 41,000 miles of the interstate highway system	9
90-575	1968	Extension of all major higher education acts including National Defense Education Act (NDEA), Higher Education Act of 1965, and National Vocational Student Loan Insurance Act of 1965	8.75
94-164	1975	Extension of tax cuts, after previous presidential veto of similar legislation	11
98-215	1983	Aid to the Nicaraguan contras, for 1983	9.5
103-325	1994	Established community development banks to aid economic development in distressed areas	7
Group C: Ordinary Legislation			
84-864	1956	Raised the borrowing power of the Commodity Credit Corporation (CCC) for farm price supports (second increase in 84th)	6.5
89-298	1965	Authorized $2 billion for 140 rivers and harbors projects of the Army Corps of Engineers, only one of which was controversial	6.5
93-182	1973	Established daylight savings time for 1974–76	3.25
98-161	1983	Increased the debt limit (second of three increases that year)	2.25
103-403	1994	Reauthorized the Small Business Administration	2.25

TABLE 4
Distribution of Sweep Two Laws in Groups A–D

Significance Group	Sweep Two Laws in Significance Group (Column 1)	Public Laws in Significance Group (Column 2)	Percentage of Group in Sweep Two ((1)/(2) x 100)	Percentage of Sweep Two in Group ((1)/200)
Group A	146	226	64.6	73.0
Group B	23	292	7.9	11.5
Group C	27	1733	1.6	13.5
Group D	4	15412	0.0	2.0
Total	200	17663	1.1	100.0

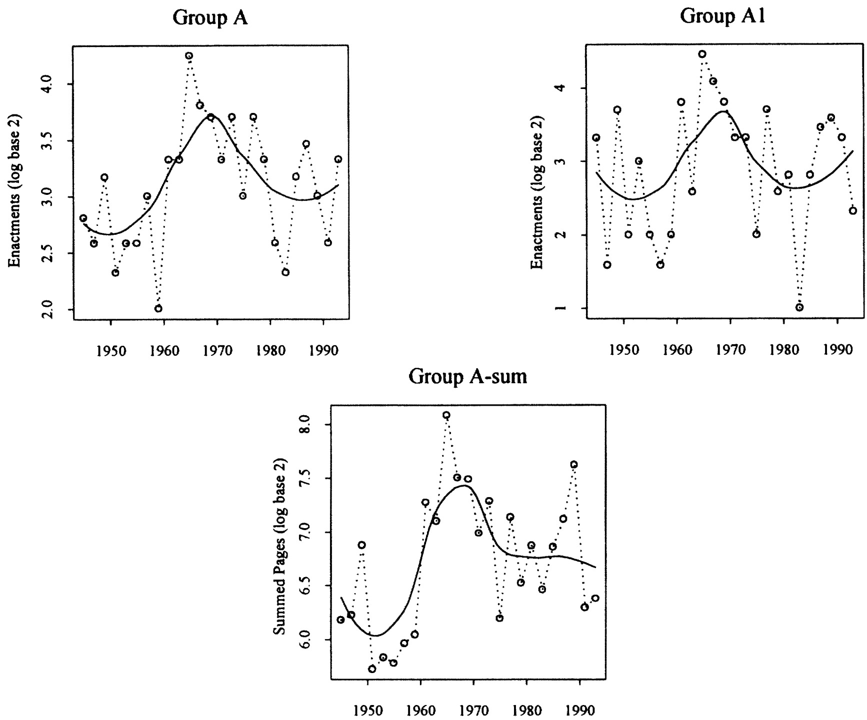
while that of Group B stories is 9.0 pages (means of 12.8 and 12.0, respectively); the upper value of the interquartile range for the Group A stories is 16.0, while that of Group B is 14.2 pages. Group A stories tend to be somewhat longer than Group B stories despite the length requirement applied to Group B.[16] A random sample of laws in Groups C and D yields estimates of mean lengths of 3.4 pages and 0.3 pages, respectively.

To summarize, a variety of independent checks confirm that the four categories tap into legislative significance and tend to sort the enactments appropriately. These tests show that the sources have a common and consistent perception of "legislative significance" and, therefore, that the criteria we use to judge the importance of different bills is meaningful.

Trends in the Data

Figure 2 examines all three measures of Group A legislation over time, along with fits from extremely flexible, nonparametric, locally weighted regression models. These models help uncover structure in the data.[17] All three versions reveal an exceptionally prominent feature in the data, the "bulge." A wave of enactments of landmark legislation swells from a trough in the early- to mid-1950s, crests around 1970, and then declines from that point through the mid-1980s. The actual peak in the data occurs in all cases in the 89th Congress, the Great Society Congress. In the three measures involving counts, the two troughs on either side of the bulge are approximately the same height, so the series appears roughly symmetrical. However, the A-sum category, which takes into account the increasing use of omnibus

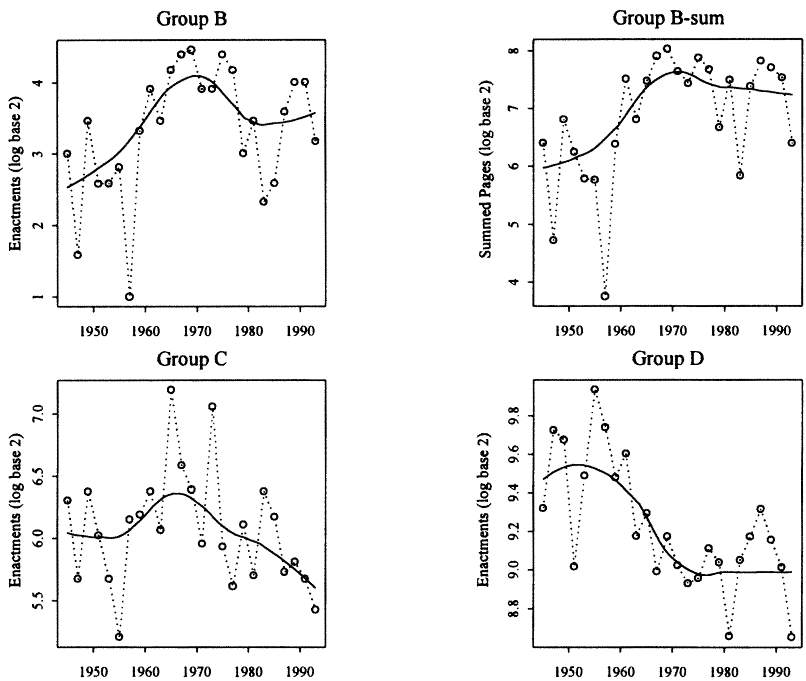
FIGURE 2
Three Measures of Landmark Legislation, by Congress



measures in the later part of the series, displays a different pattern. In the post-bulge era, enactments do not fall to their previous level—there is a pronounced "ratchet effect" (Higgs 1987).

Figure 3 presents a similar analysis for the other significance categories. The bulge in enactments has a dramatic impact on these categories. First, the bulge reappears in Groups B and C but not in Group D. In other words, the wave of congressional activism lifted all legislative boats except the most minor. Second, the ratchet reappears in Group B (major legislation). In Group B-sum, the measure sensitive to omnibus legislation, the ratchet is so pronounced that the series resembles a step function: production begins at a uniformly low level, ramps up to a much higher level, and stays there. In contrast, there is no ratchet in Group C—levels fall farther after the bulge than before. The pattern in Group D is also a step function but exactly opposite to that in Group B-sum: production begins at a high level, ramps down at

FIGURE 3
Measures of Other Enactments, by Congress



the same time that Group B-sum ramps up, and then stays at a low level of production.

The following picture emerges from the data. A huge surge of legislative activism began in the Eisenhower years and peaked with the Great Society but continued well into the Nixon administration (as noted by Mayhew). This surge was so pervasive that it temporarily boosted enactments of all significance levels except the most minor. After the Nixon years, the wave of landmark enactments subsided, but production of major legislation continued at levels unknown in the first half of the post-war period. We conjecture that managing the new programs and vastly increased state responsibilities created during the bulge required ongoing enactments of major legislation, hence, the ratchet and step effects. Even once the forces that created the bulge weakened, the continuing activity in major legislation crowded out ordinary and especially minor legislation. These have now approached modern lows.

Divided Government Revisited

Given defensible measures of legislative productivity that span multiple levels of significance, we can reconsider the effect of divided government on legislative productivity. As before, we test all measures for stationarity. All are trend stationary as shown by Dickey-Fuller tests, though different functional forms are required to de-trend different series: the effect of time on Group D legislation, for example, differs markedly than that for Groups A and B. We run Poisson regressions after making sure that the residuals of each are stationary, and hence the *t*-statistics are nonspurious. Once more, OLS regressions generate virtually identical results, but because Poissons are more appropriate, we omit the OLS findings (except for Group D, in which the numbers of laws are sufficiently large to warrant using OLS).

Table 5 presents our results. For Group A legislation, divided government is statistically significant and appropriately signed. Again, we find that during periods of divided government, Congress produces significantly fewer landmark bills than during periods of unified government. For Group A1, the size of the coefficient even increases.

How large is the divided government effect on the production of Group A legislation? The predicted magnitude of the effect depends on the baseline value of production determined by the polynomial in time and ranges from a low of 1.6 laws in the 79th Congress (1945–46) to a maximum of 3.7 laws in the 93d Congress (1973–74). However, the percentage effect is constant in the model. It is easily calculated as a 28% reduction in output by the move from unified to divided government and, conversely, a 39% increase in output by the move from divided to unified.

For Groups B, B-sum, and C, divided government never approaches statistical significance. Divided government, it seems, has little or no impact on Congress's ability to pass legislation in these significance realms. While these groups contain major enactments, they do not consist of the historic and landmark laws represented in the Group A series.

In Group D, the least significant laws, unified government exerts a statistically significant negative effect. That is, divided government actually increases the number of relatively unimportant laws passed by Congress, by about 97 laws per session. This can be seen as a kind of substitution effect. Perhaps the depressing effect of divided government on landmark legislation frees resources or impels representatives to pass commemorative and minor legislation for which they can claim credit in their reelection bids.

TABLE 5
Divided Government and the New Measures
(standard errors in parentheses)

Variables	Group A	Group A1	Group A-Sum	Group B	Group B-Sum	Group C	Group D
Unified	0.281**	0.470**	23.057	0.150	5.878	0.086	–97.42**
	(0.111)	(0.219)	(18.554)	(0.171)	(24.811)	(0.120)	(43.55)
Time	0.111**	0.093	13.040**	0.165**	20.214**	0.059*	19.68
	(0.032)	(0.058)	(4.721)	(0.050)	6.279	(0.031)	(26.00)
$Time^2$	–0.004**	–0.003	–0.452**	–0.006**	–0.626**	–0.003**	–4.31
	(0.001)	(0.002)	(0.189)	(0.002)	(0.251)	(0.001)	(2.56)
$Time^3$	—	—	—	—	—	—	0.13**
							(0.07)
Constant	1.476**	1.451**	33.965	1.479**	19.211	4.014**	797.41**
	(0.202)	(0.368)	(27.746)	(0.323)	(37.308)	(0.184)	(75.45)
n	25	25	25	25	25	25	25
Residual Deviance	16.15	51.35	44842.85	46.32	74920.11	127.27	(0.590)[a]

Note: Results generated from Robust Poisson regressions, except for Group D, for which raw counts from an OLS regression are reported. Therefore, [a] indicates adjusted R^2, not residual deviance.

*significant at the .10 level; **significant at the .05 level.

A finding worth highlighting concerns Group A-sum, the group constructed to study the effect of omnibus legislation (see Column 7 in Table 2). The sign on unified government takes the (by now) expected positive direction, indicating a decrease of 28 pages in *CQ*'s total coverage of A-level laws under divided government. (Recall that Group A-sum is measured in pages of *CQ* coverage not law counts.) Since the average A-level law received 12.9 pages of coverage in *CQ*, this coefficient suggests a drop per Congress of about two A-level laws due to divided government. However, the coefficient does not reach conventional levels of statistical significance. Perhaps, then, divided government depresses the production of landmark laws simply because of the increased use of omnibus legislation after 1980. If this were so, however, the divided government effect should vanish if we truncate the analysis at 1980. This intuition does not bear out: a regression similar to those in Table 1 run only for the pre-1980 period continues to find a detectable effect from unified government ($t = 1.73$). The estimated effect is somewhat smaller, about 2.6 laws rather than the 3.1 estimated for the entire series. The divided government effect, nonetheless, cannot be explained away by the rise of omnibus legislation.

Regardless of one's evaluation of the substantive significance of the divided government effect, the following is worth noting: the surge of enactments that gives rise to the nonstationarity in the data is a much larger effect than that of divided government. Actual A-level enactments rose from about 5–7 per Congress in the late 1940s and early 1950s to about 10–14 per Congress in the early 1970s. The model in Column 1 of Table 5 predicts a rise from 5.8 enactments in the 79th Congress to 13.3 in the 93d (given unified government), an increase of 7.5 laws—twice the largest estimated divided government effect of 3.7 laws.

Conclusion

This paper presents an extensive analysis of how divided government affects legislative productivity. Using appropriate time-series techniques, we demonstrate that periods of divided government depress the production of landmark legislation by about 30%, at least when productivity is measured on the basis of contemporaneous perceptions of important legislation. This depressing effect is not due to increased use of omnibus legislation after 1980. The divided government effect attenuates for less significant enactments (Groups B and C) and reverses for the least-significant laws (Group D).

In addition, divided government appears to increase the production of legislation deemed very important solely in retrospective reviews, a finding we conjecture is due to selection bias in recent policy histories. Mayhew's no-effect finding on divided government and legislative productivity is a consequence of aggregating two quite different series, one based on contemporaneous perceptions, with a divided effect, and one based on retrospective evaluations, with a reverse effect. We suggest that measures based on contemporaneous perspectives are more meaningful for studying the politics of divided government.

We perceive three main avenues for future research. First and foremost, more attention must be given to the theoretical foundations of legislative productivity. With the exception of Krehbiel (1998), studies of legislative productivity have yet to link well-developed theory, appropriate independent variables, and the new data on legislative outputs. The macro-oriented measures of legislative productivity we present offer a new means of testing microtheories of Congress.

Second, future work on legislative productivity might do well to focus less on the politics of divided government and more on explaining the most prominent feature of the post-WWII legislative activity, the bulge (presented in Figures 2 and 3). The bulge accounts for far more variance than divided government in the production of Group A legislation. Moreover, unlike the divided government effect, the time-dependent bulge was not restricted to the top tier of significant enactments. It also accounts for boosted output in more routine laws and the dramatically depressed production of commemorative and other less important enactments. Further research on the legislative productivity of Congress should focus on explaining this dramatic event in the history of the post-war Congresses.

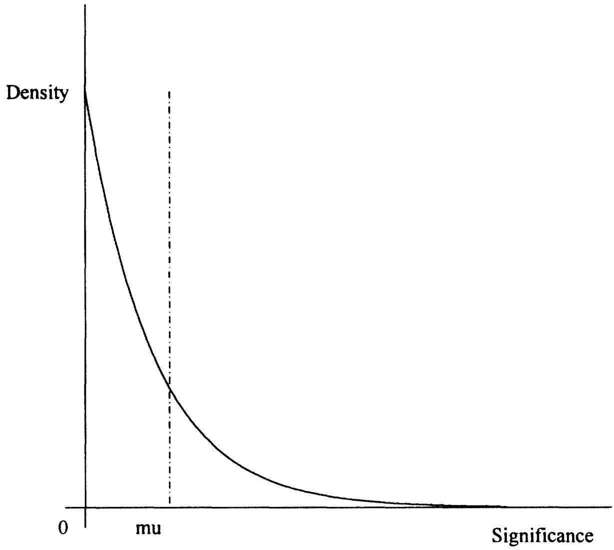
Finally, because we rely primarily upon law counts, we can only infer how divided government affects the content of bills. It may be, for instance, that during periods of divided government, what would otherwise be extremely innovative, and therefore landmark, legislation is compromised down into the lower echelons of legislative significance. Modifying the output measures to reflect bill content in more detail—for example, by including ideological valence or policy composition—seems a promising direction for research.

William G. Howell will be Assistant Professor of Political Science, University of Wisconsin, Madison, Wisconsin 53706 beginning in the fall of 2000. E. Scott Adler is Assistant Professor of Political Science, University of Colorado, Boulder, Colorado 80309. Charles Cameron is Associate Professor of Political Science, Columbia University, New York, NY 10027. Charles Riemann is Visiting Assistant Professor of Political Science, University of Connecticut, Storrs, Connecticut 06269.

APPENDIX
A Model of Source Reporting and Testing for Source Consistency

Interpreting the source data requires a model of source reporting. Assume the significance of any enacted bill is clearly discernible to a knowledgeable observer within the congressional community. Also assume there are a great many bills of very little significance and only a few of much significance. Such would be the case if bills were distributed by significance according to an exponential distribution as shown in Figure A1. Then imagine an observer establishing a reporting threshold μ (mu) such that the observer reports all enactments with at least μ amount of significance but does not report public laws with less significance. That is to say, the observer reports the right-hand tail of the distribution, as shown in Figure A1. We take this to be a stylized representation of the legislative roundup coverage in the *New York Times, Washington Post*, and *CQ Almanac*.

FIGURE A1
Threshold Reporting Model



In reality, Type I and Type II errors intrude on actual source reporting: the observer may fail to report some public laws with significance greater than μ and may also report some with significance less than μ. So an empirical distribution of bills reported by a source would probably fail to show a completely crisp cut-off at μ. However, if the departures from the stylized model are relatively few, the distribution will retain the hypothesized shape.

If the data are generated by threshold reporting, it may be possible to use the source mentions to score legislative significance. But if not, they are poor candidates for this effort. Fortunately, the model suggests a simple but demanding test. Suppose there are several observers who view the significance of bills similarly and employ threshold reporting, but they use different reporting thresholds. The observer with the highest threshold will report the existence of a particular set of public laws. The observer with the next highest threshold will report a larger number of bills, including *all* the bills reported by the first observer, plus others. A third observer with an even lower threshold will report all the bills mentioned by the first two observers, plus additional bills of even less significance. In short, the enactments reported by the observers will be *nested* in one another. Conversely, suppose each observer simply reports bills drawn at random from the universe of bills. If the number drawn is small relative to the universe (in this case 17,663), it is very unlikely that the bills will display this nesting property.

The *Washington Post* and *New York Times* generally use the same model of legislative significance. For example, of the 893 public laws mentioned in the source stories in the *New York Times*, 572 or 64% were also mentioned in the source stories in the *Washington Post.* Simple chi-square tests for independence reject the hypothesis that the mentions in the sources are drawn independently from one another ($p = 0.00$ in all three cases). But this is only a minimal and easily satisfied requirement. Figure A2 provides a much more informative look at the nesting properties of the source data, showing the percentage of mentions in the *smaller* of the two sources that are also mentioned in the *larger* of the two sources for each of the years from 1945 to 1994. The yearly counts are not arranged in chronological order, so we may easily identify low and high scoring years.

If the sources are nested, these percentages should be high. The top two panels provide strong evidence that the *Washington Post* and *New York Times* mentions are indeed nested in the *CQ Summary* mentions, with the exception of perhaps one year in either case. The *New York Times* and *Washington Post* offer a particularly demanding test since they typically report similar but fairly small numbers of stories (in most years ranging between about 8 and 32 mentions). Thus, even a moderate degree of divergence in what they report will dramatically degrade the nesting. But, as shown in the bottom panel of Figure A2, in all years but one, more than two-thirds of the *New York Times* and *Washington Post* mentions are nested in each other. In 30 of the 50 years, more than 80% of the mentions are nested within each other, and in 20% of the years the mentions nest *perfectly*. All together, the pool of mentions in the smaller source is 690, that in the larger is 1,017, and the number of mentions in both sources is 572; so 83% of the mentions in the smaller pool are nested within the larger pool. This degree of nesting strongly suggests that the authors of the roundups identify the same characteristic in the public laws, i.e., legislative significance, and report on the laws accordingly.

The Problem of Source Consistency

Even if the source mentions are generated by the threshold-reporting model, censoring can degrade the data's comparability across years. Consider a single observer who reports on the distribution of bills in two different years. Assume the distribution is identical in the two years, but the observer uses different thresholds, e.g., $\mu_1 > \mu_2$. In the year in which the reporter uses the higher threshold, the production of significant public laws will appear to slump. But the decline will be spurious, reflecting only the censoring of public laws with significance levels between μ_1 and μ_2. More generally, suppose two different sources report across several years using different thresholds from one another and different thresholds in different years. To avoid the censoring problem, one should enforce as the threshold a value of μ greater than or equal to the *maximum* of the *minimum* μ*s*.

Censoring seems most likely to affect the newspaper mentions. Idiosyncratic factors, such as the occurrence of a particularly important story on the day a session ends, may limit the space available for a roundup. More importantly, shrinking column inches may lead to rising reporting thresholds over time. An examination of the number of mentions in the *New York Times* and *Washington Post* supports this concern. In the *New York Times*, almost all years before 1955 lie above the mean number of mentions; almost all years after 1979 lie below the mean. Simple examination of articles and, indeed, the entire paper suggests shrinking coverage. In the *Washington Post*, the data show an increase in mentions up to the middle 1960s, followed by a decline through the late 1970s. Most years after 1975 lie below the mean, suggesting shrinking space and (consequently) rising thresholds.

FIGURE A2
Nesting of Mentions of Public Laws

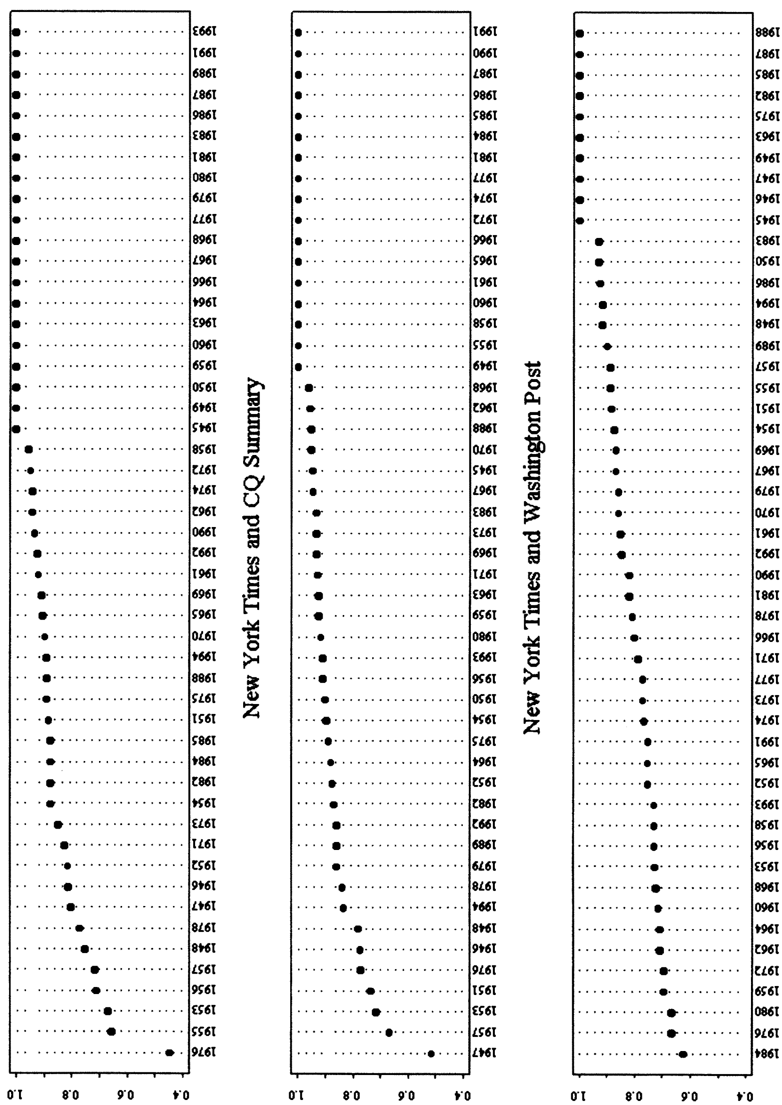
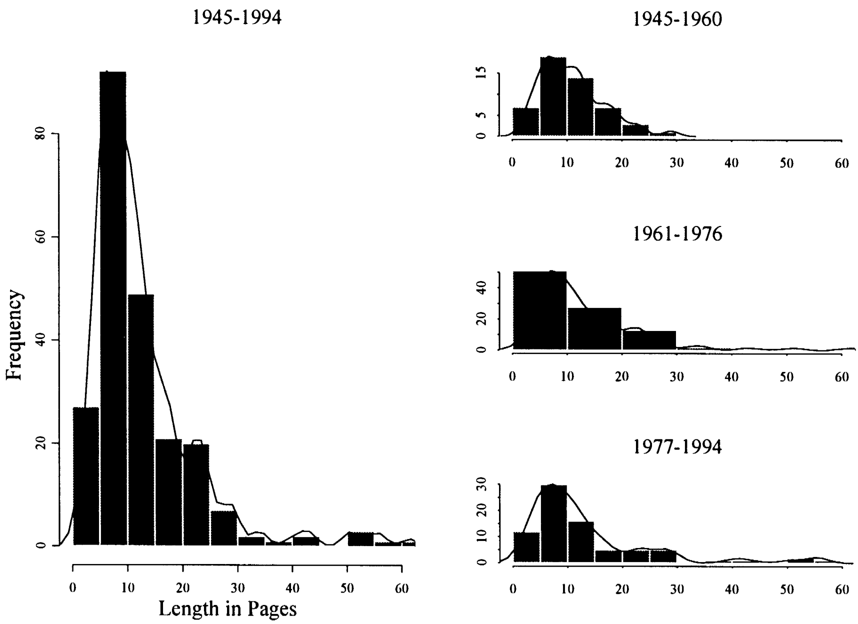


FIGURE A3
Distribution of Sweep One Stories



To solve the censoring problem, we use the *CQ Almanac*. Every public law mentioned in the newspapers was discussed in the body of the *CQ Almanac*, if only in summary listings of public laws. The length of *CQ*'s coverage is strongly related to apparent legislative significance, e.g., laws on the order of Medicare receive careful descriptions and detailed legislative histories while minor laws are banished to terse summary listings. The length of *CQ*'s discussion of an enactment is therefore a plausible proxy for significance.

The writers and editors of *CQ Almanac* may have changed their reporting standards over a 50-year period.[18] Perhaps, then, two public laws with comparable significance might be discussed at different lengths at different times. We use Mayhew's Sweep One data (again, which measures significant legislation based upon the contemporaneous judgments of political generalists) to make sure that similar bills are reported at similar length over time. The reporting threshold for the Sweep One data is so high that these data are almost certainly uncensored, i.e., the number of laws in each session that meet this extraordinarily demanding threshold is so small that even shrunken roundups surely report all the Sweep One enactments (subject to occasional Type II errors).[19]

Figure A3 summarizes this analysis. The histogram in the leftmost panel shows the distribution of Sweep One stories by page length in the *CQ Almanac* over the entire period. Superimposed on the histogram is a nonparametric estimate of the empirical distribution (a kernel smoother, with smoothing parameter set to show considerable detail in the distribution). As expected, the distribution strongly suggests an exponential distribution with a few short outliers, either laws mistakenly identified as Sweep One caliber, or Sweep One laws covered in an unusually terse fashion in the *CQ Almanac*. As can be seen, the revealed cutoff is approximately six pages. The three panels on the right repeat this analysis for the early, middle, and late years of the time series.[20] The empirical distribution is remarkably similar across the periods, in all cases suggesting a cutoff of about six pages. No adjustments to the *Congressional Quarterly* page lengths seem necessary.[21]

A similar analysis for all the public laws reported in the newspapers yields for each year an estimated *CQ* page length below which the larger of the two newspaper sources did not report public laws. In the early years, in which the voluminous roundup articles might mention as many as 50 public laws, the apparent reporting threshold was as low as one to two pages in *CQ* coverage. This threshold rose to as high as four pages during the 1960s, when Congress enacted many notable laws but the space afforded the newspaper roundups was shrinking. The shrunken roundups of the late 1980s employ a cutoff almost as high as the cutoff for the Sweep One laws. The shrinking coverage in the *New York Times* and *Washington Post* roundups requires imposing a rather high threshold on the newspaper mentions if the data are to be comparable over time.

NOTES

Cameron gratefully thanks the Hoover Institution, where the bulk of his work was completed. For helpful suggestions, we thank seminar participants at Stanford and the University of California at Berkeley, especially John Ferejohn. We thank Sharyn O'Halloran for sharing data on landmark bills. We particularly thank Harry Paarsch, whose insights repeatedly altered our understanding of the data, and David Mayhew for his comments, suggestions, and encouragement. Jennifer Ghandi, Brad Joseph, Patrice Johnnson, Jason Lynch, Rebecca Miller, Anita Ramen, and Eric Reinhardt provided excellent research assistance. NSF grant SES-9223396 partially supported data collection.

1. The nonstationarity of the Mayhew data is readily confirmed through formal statistical tests, i.e., a battery of augmented Dickey-Fuller tests. To test for the existence of a unit root, we estimate $\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \Sigma B_i \Delta y_{t-I} + \varepsilon_t$, where y = Group A legislation, t = time, Δy_t is a first differenced term, and y_{t-I} is a lagged variable (for i = 1, 2, generating first and second differences on the lagged variables). The values of γ and the intercept, when compared to Dickey-Fuller critical values, provide preliminary evidence about the existence of a unit root. If neither are significantly different from zero, then we reestimate the model with fewer regressors and reexamine the significance of the intercept and lagged-variable coefficient. If necessary, this process is repeated once more, with one less regressor. For a full discussion on these tests, see Doldado, Jenkinson, and Sosvilla-Rivero 1990, and Enders 1995, 256–58. The Dickey-Fuller test is biased against the null that a series is stationary. However, because the battery of tests we run for all series suggests nonstationarity, and after making appropriate adjustments, the series are made stationary, we are confident about these findings.

2. This basic feature of the data can be seen in the top panel of Figure 1, which graphs the union of the Sweep One and Sweep Two laws for the Congresses in which both series are available (1947–86). The counts are shown as small circles connected with lines. Superimposed on the data is a dashed line that represents a simple, locally weighted regression, an aid in visualizing the patterns in the data. The fit from a locally weighted regression is span = 3/4, degree = 2. As noted by Mayhew, this measure of landmark enactments steadily rises through the late 1950s, peaks in the Congresses of the late 1960s and early 1970s, and steadily declines thereafter. Clearly the mean is not time invariant.

3. Alternatively, we could have first regressed the Sweep One and Two series on a polynomial in time and then utilized the residuals from the de-trended data. Virtually identical results are generated using this method.

4. The regressions in Table 7.2 of Mayhew 1991 append the union of Sweep One and Sweep Two for 1947–86 to the Sweep One data for 1987–90. However, the results for this somewhat longer series are quite similar to those reported in the text for the union of Sweep One and Sweep Two.

5. In Poisson regressions of the type reported above, the *t* statistic on unified is –2.3 for the last half of the series but only 1.3 for the first half.

6. As discussed below, we require each story to meet a minimum threshold of length in *CQ*. Large but otherwise unnoteworthy appropriations passed in the last moments of a Congress may be "nominated" by a closing story and slip through the "confirmation" screen. This possibility should be borne in mind when evaluating appropriations passed on the last day or so of a Congress.

7. Achieving a positive identification of every enactment mentioned in the newspaper stories was a laborious task. First, we searched the newspaper stories for mentions at least three separate times. Second, each mentioned enactment was positively identified by at least two coders (among whom the senior author was always included), who used *CQ Almanac* and occasionally the *Congressional Index* and the *Digest of General and Public Bills*. Disagreement among coders was broken by a third and, in some cases, an additional fourth coder. Each of the mentions in the newspapers was also coded for the length of the public law's coverage in the body of the *CQ Almanac*. The length was recorded in pages, rounded to the nearest quarter page.

8. The paperback edition of Mayhew 1991 contains a list of Sweep One laws for the 102d Congress; Mayhew 1995 provides a list for the 103d Congress.

9. Counts of enactments suffer from a potentially serious shortcoming: mammoth omnibus bills count as only one enactment. Beginning in the late 1970s, some of the most important enactments are huge omnibus bills, e.g., budget reconciliation bills, deficit reduction packages, or continuing resolutions. To offset this problem, we provide a supplement to simple counts in the higher significance categories: sums of pages of coverage in the *CQ Almanac*. In this scheme, a budget reconciliation act discussed in 40 pages in *CQ* counts for more than an enactment discussed in only 6 or 8 pages.

10. Yearly counts of the mentions are skewed toward large values and display monotone spread (the variance increases with the mean). This pattern is typical in count data but hinders analysis since outliers distort the mean, residuals from fits are unlikely to be normally distributed, and comparisons across sources must take into account not only the differences in means but the changing variance of the data. Fortunately, a logarithmic transformation pulls in the outliers, significantly normalizes

the distributions, and stabilizes the variances so that monotone spread vanishes. Accordingly, we use log counts in the presentation of these data unless otherwise indicated. To avoid cumbersome fractional values of ten, we employ log base two (Cleveland 1993). Raw counts are easily recovered, e.g., a count of "5" indicates $2^5 = 32$.

11. Our counts differ very slightly from Mayhew's since we exclude constitutional amendments, add data for 1945 and 1946, and require a very strict correspondence between each item and a single public law. Otherwise the lists are the same.

12. A statistical analysis of the text of *CQ Almanacs* reveals that the summaries include authorizations greater than or equal to 1.5 pages and all appropriations outside the appropriation section of the *Almanac*. We used this rule to hypothesize the probable contents of *CQ Summaries* before 1947 and after 1991.

13. We exclude appropriations to maintain comparability with the Sweep One laws; Mayhew also excluded appropriations.

14. The enactments in each group were sorted by year of enactment and public law number; every *n*th law was then drawn, where *n* equals the number of enactments in the group divided by five.

15. The few Sweep Two laws that are not included on the Sweep One list were analyzed in section 2 and, we found, generated the no-effect finding.

16. If the Group A laws are truncated with respect to length in the same way as the Group B laws, the mean length of the truncated Group A laws is 14.5 pages.

17. On locally weighted regression see *inter alia* Cleveland, Grosse, and Shyu 1993, Cleveland 1993, and Hastie and Tibshirani 1990. The models shown were created using the loess function in S-Plus. The models were selected through comparison of alternatives using analysis of deviance tables (McCullagh and Nelder 1989). The alternative models gradually reduced the span and increased the degree of the model. A more local model and a locally quadratic model were selected over a less local and locally linear model only if the former clearly outperformed the latter ($p > .10$). Examination of normal quintile plots of the residuals indicated whether least squares was satisfactory or whether robust methods were needed. In Figure 5, the spans = 3/4 for the first three models and 2/3 for the last. All are locally quadratic (degree = 2), and none employs the bisquare alternative to least squares. In Figure 6, the spans = 1/2 for the first three models and 7/8 in the last. Bisquare was used for all models except the last, and degree = 1 for all models except the last. None of the fits shown are particularly sensitive to changes in any of these parameters.

18. The principal "eras" are: 1) 1945–47 (quarterly volumes); 2) 1948–54 (annual volumes, three columns per page); 3) 1955–65 (two columns per page); 4) 1966–67 (different typeface); 5) 1968–present (slightly different typeface).

19. Mayhew 1991 believes that the standard employed for Sweep One laws is constant over time, and, based on our own reading of the roundups, we agree.

20. In fact, we undertook this analysis for each of the periods in which *CQ* used the same page layout and in smaller increments in each of the lengthier eras. The results change little, although in the last five years of the series there appears to be an unusually large number of Sweep One stories under six pages in length (e.g., the Family Leave Act, the Brady Bill, and reparations to Asian Americans interred during WWII).

21. It may seem surprising that *CQ*'s coverage of major laws, as measured in page length, is so consistent over time, but perhaps one explanation lies in the economics of publishing. The editors of *CQ* may have an approximate idea of how long

the entire *Almanac* can be on average, with some variation allowed from year to year. At any rate, simple inspection of a complete set of *CQ Almanacs* reveals that they have changed relatively little in page length over the 50-year period.

REFERENCES

Binder, Sarah. 1999. "The Dynamics of Legislative Gridlock, 1947–96." *American Political Science Review* 93:519–33.

Cleveland, William S. 1993. *Visualizing Data*. Summit, New Jersey: Hobart Press.

Cleveland, William S., Eric Grosse, and William M. Shyu. 1993. "Local Regression Models." In *Statistical Models in S*, ed. John M. Chambers and Trevor J. Hastie. New York: Chapman & Hall.

Coleman, John J. 1999. "Unified Government, Divided Government, and Party Responsiveness." *American Political Science Review* 93:821–36.

DeBoef, Suzanna, and Jim Granato. 1997. "Near-Integrated Data and the Analysis of Political Relationships." *American Journal of Political Science* 41:619–40.

Doldado, Juan, Tim Jenkinson, and Simon Sosvilla-Rivero. 1990. "Cointegration and Unit Roots." *Journal of Economic Surveys* 4:249–73.

Edwards, George C. III, Andres Barrett, and Jeffrey Peake. 1997. "The Legislative Impact of Divided Government." *American Journal of Political Science* 41:545–63.

Enders, Walter. 1995. *Applied Econometric Time Series*. New York: John Wiley & Sons.

Fiorina, Morris. 1980. "The Decline of Collective Responsibility in American Politics." *Daedalus* 109(3):25–45.

Ford, Henry Jones. 1898. *The Rise and Growth of American Politics*. New York: Macmillan.

Granger, Clive, and P. Newbold. 1974. "Spurious Regressions in Econometrics." *Journal of Econometrics* 2:111–20.

Hall, Richard. 1996. *Participation in Congress*. New Haven, CT: Yale University Press.

Hastie, Trevor J., and Robert J. Tibshirani. 1990. *Generalized Additive Models*. London: Chapman and Hall.

Higgs, Robert. 1987. *Crisis and Leviathan: Critical Episodes in the Growth of American Government*. New York: Oxford University Press.

Jones, Charles. 1994. *The Presidency in a Separated System*. Washington, DC: Brookings Institution.

Key, V.O., Jr. 1964. *Politics, Parties, and Pressure Groups*. New York: Crowell.

Kelly, Sean. 1993. "Divided We Govern? A Reassessment." *Polity* 25:475–84.

Krehbiel, Keith. 1996. "Institutional and Partisan Sources of Gridlock, A Theory of Divided and Unified Government." *Journal of Theoretical Politics* 8:7–40.

Krehbiel, Keith. 1998. *Pivotal Politics: A Theory of U.S. Lawmaking*. Chicago: University of Chicago Press.

Mayhew, David R. 1991. *Divided We Govern: Party Control, Lawmaking, and Investigations, 1946–1990*. New Haven: Yale University Press.

Mayhew, David R. 1995. "Clinton, the 103d Congress, and Unified Party Control: What are the Lessons?" Presented at a conference honoring Stanley Kelley, Jr., Princeton University, October 27–28, 1995.

Mayhew, David R. 1996. "Presidential Elections and Policy Change: How Much of a Connection Is There?" In *American Presidential Elections: Process, Policy and Political Change*, ed. Harvey Schantz. New York: State University of New York Press.

McCullagh, P., and J.A. Nelder. 1989. *Generalized Linear Models,* 2d ed. London: Chapman and Hall.

Schattschneider, E. E. 1942. *Party Government*. New York: Rinehart.

Stimson, James, Michael MacKuen, and Robert Erikson. 1995. "Dynamic Representation." *American Political Science Review* 89:543–65.

Sundquist, James L. 1988. "Needed: A Political Theory for the New Era of Coalition Government in the United States." *Political Science Quarterly* 103:613–35.

Wilson, Woodrow. [1908] 1961. *Constitutional Government in the United States*. Reprint. New York: Columbia University Press.