Advance Access publication February 12, 2008

Political Analysis (2008) 16:303–323
doi:10.1093/pan/mpm035

Legislative Productivity of the U.S. Congress, 1789–2004

J. Tobin Grant
Department of Political Science, Mailcode 4501,
Southern Illinois University, Carbondale, IL 62901-4501
e-mail: grant@siu.edu (corresponding author)

Nathan J. Kelly
Department of Political Science, University of Tennessee,
1001 McClung Tower, Knoxville, TN 37996-0410
e-mail: nathan.j.kelly@gmail.com

We measure legislative productivity for the entire history of the U.S. Congress. Current measures of legislative productivity are problematic because they measure productivity for a limited number of decades and because they are based on different aspects of productivity. We provide a methodology for measuring (1) a Legislative Productivity Index (LPI) and (2) a Major Legislation Index (MLI). We use the W-CALC algorithm of Stimson (1999, *Public opinion in America: Moods, cycles, and swings*. 2nd ed. Boulder, CO: Westview Press) to combine information from previously used indicators of productivity into measures of the LPI and the MLI. We provide examinations of content, convergent, and construct validity. The construct validity model includes potential determinants of legislative productivity. We conclude that the LPI and the MLI are superior measures of productivity than other measures used in the literature.

1 Introduction

Since Mayhew's (1991) examination of the effects of divided government, there have been many studies of legislative productivity, which is implicitly defined as the amount of policy made by Congress. It is a macro-level variable in which high levels of productivity indicate many policy changes. Evaluations of legislative productivity require, first and foremost, valid measurement. Unfortunately, there is little consensus on what should be used as a valid measure of legislative productivity. Researchers define, redefine, and operationalize the concept of productivity, with a major point of contention being which legislation should be counted toward production (e.g., Mayhew 1991, 1993; Kelly 1993; Krehbiel 1998; Binder 1999; Coleman 1999; Howell et al. 2000; Clinton and Lapinski 2006). Rather than drawing us closer to understanding legislative productivity, the use of alternative operationalizations has resulted in divergent theoretical claims (Coleman 1999; Howell et al. 2000).

Author's note: A previous version of this article was presented at the 2006 meeting of the Midwest Political Science Association. For suggestions and criticisms, we thank James Avery, Jason Barabas, Janet Box-Steffensmeier, Sean Ehrlich, Chuck Finocchiaro, Jason Gainous, Mark Hurwitz, Jennifer Jerit, Gregg Johnson, Jana Morgan, Scott McClurg, and four anonymous reviewers.

This measurement problem is accentuated by the limited number of Congresses measured. Nearly all the research on legislative productivity has measures that are limited to Congresses in the second half of the 20th century. With each Congress lasting 2 years, this results in a small number of observations. Thus, it is of little surprise that even subtle changes in conceptual definition result in different theoretical conclusions. It is the goal of this paper to provide a valid measure of legislative productivity that will move the study of legislative productivity beyond its current state of semantic confusion and toward stronger models of Congressional policy making. The remainder of the paper is organized in three sections. In the first section, we discuss previous efforts to measure legislative productivity. Current measures fall into two broad categories—those that focus on important policy enactments and those that examine general lawmaking. We go on to argue that both general and important lawmaking contribute to policy production and should both be considered in a complete measure of legislative productivity. Also in this first part of the paper, we develop a multiple indicator approach to the problem that overcomes previous measurement issues and examines general and important lawmaking both separately and as a combined measure of legislative productivity. In the second section, we discuss a variety of validity tests that show the appropriateness and usefulness of combining single indicators of legislative production into indexes of productivity. We find that a measure combining both important and general lawmaking is the most appropriate strategy for measuring overall legislative production. Finally, in the third section, we conclude by summarizing our results and briefly discussing some further implications of our work.

2 Measurement of Legislative Productivity

Which Congresses have created the most policy? What are the causes of higher or lower levels of productivity? How has productivity changed over the course of Congressional history? To answer these and other important questions, we must first agree on a valid measure of productivity. Unfortunately, there is no consensus on the measurement of productivity. Rather, there are many measures of productivity, each of which is based on very similar, but not identical, criteria. Making the measurement even more difficult is that each measurement procedure has strengths, but none is broad enough to encapsulate all pertinent aspects of productivity. In addition, we lack measurements for most of the 109 Congresses. The longest time period any measure spans is the aggregations of Clinton and Lapinski (2006) (58 Congresses). Most cover Congresses since the 1950s (25–30 Congresses). We need valid measures of productivity that allow us to gauge productivity over the entire history of the U.S. Congress.

Problems of conceptual clarity even exist with the most recent advances in measurement that focus on the importance of *individual statutes*. Clinton and Lapinski (2006) provide the most rigorous measurement of statute-level significance. Using an item-response model, the Clinton-Lapinski (CL) model estimates the significance of each statute passed by Congress from 1887 through 1994. Legislative productivity, however, is a macro-level variable. It is an aggregation of the individual policy changes enacted by each Congress. How to move from a rigorous measurement of statute-level significance to aggregate productivity requires, as they rightly note, arbitrary measurement decisions.[1]

[1]As Clinton and Lapinski (2006) explain: "One difficulty raised by our measure is determining how to aggregate and summarize the statute-level estimates .... Since we estimate the significance for every public law, the appropriate measure for us is more ambiguous. Consequently, we follow Baumgartner and Jones (2002, 2004) and construct a measure focusing on statutes whose significance rank exceeds an exogenous threshold. *Since the threshold is arbitrary*—it is unclear whether analyzing the top 500 (of 37,766) enacted statutes provides a better measure of accomplishment than analyzing the top 2000—we examine the consequences of several thresholds" (245, emphasis added).

Though the CL model provides more precise measurements for the importance of individual statutes, there are no criteria that provide guidance as to what level of importance should be used to count the productivity of a Congress—the choice of counting how many of the 500 most important laws were passed by a Congress is as justifiable as the count of the 3500 most important laws.[2] The aggregation problem occurs because the model measures the significance of individual statutes, not the productivity of a Congress.

We offer a different solution to the problem of measuring legislative productivity. Unlike the CL methodology, we do not focus on estimating the importance of individual statutes. Rather, our goal is to focus on the *macro-level variable*. It is at the macro level that most—if not all—of the empirical and theoretical analysis has focused. This solution is rooted in classic measurement theory. Underlying every indicator of legislative productivity are aspects of a latent concept. Some indicators do a better job of capturing the underlying concept than others. Thus, the best way to ensure valid and reliable measurement is to make use of multiple indicators of an underlying concept. Because we cannot directly observe legislative productivity, we do best to think of it as a latent variable that is manifest in various indicators.[3] Rather than favoring one indicator, we consider each of the current operationalizations of legislative productivity as providing information on the level of policy changes by Congress. When legislative productivity increases, we should observe an increase in the various indicators of legislative productivity. If indicators each tap the same latent concept, then the various definitions of legislative productivity should have common movement across time.

2.1 *Indicators of Productivity*

Previous indicators of legislative productivity vary in their operationalizations, but they follow a similar measurement process. An aspect of productivity is operationalized. Statutes are then selected based on the operationalization. Once the statutes are selected, productivity is measured by counting the number of statutes for each Congress. In this section, we discuss indicators that are available based on various operationalizations previously used. We organize indicators based on the aspect of productivity operationalized and the time period covered by the measurement.

One indicator of productivity is the number of public laws enacted by a Congress. This indicator is available for the entire history of the Congress.[4] During the first 50 years of the U.S. Congress, the average number of statutes enacted by each Congress was 119. By the 1870s, the number of laws was averaging more than twice this level. During the 20th century, Congresses have passed an average of over 635 laws, with some enacting over 900 laws. During recent Congresses (1995–2003), the average has been 421 statutes. There are two advantages of this measure of productivity. First, unlike every other measure, it is available for every Congress, and it will be available for every Congress into the future. Second, it provides a straightforward measure of the quantity of laws enacted. Procedures to differentiate broad policy change from trivial pieces of legislation are not necessary.

[2]For their figures, Clinton and Lapinski (2006) sum the estimated significance of each of the top statutes, but since the estimated significance does not vary among the top statutes, there is little difference between a sum of the number of laws passing an arbitrary threshold and a sum of the estimated significance of each of the statutes (see 245–246).

[3]Clinton and Lapinski (2006) use multiple indicators of individual statute significance. We, by contrast, use multiple indicators of macro-level legislative production.

[4]We exclude private laws because they apply to a private individual or entity and thus do not create *public* policy. Data for 1789–1968 are from the U.S. Census Bureau (1975). More recent data are from Stanley and Niemi (2007) and the Library of Congress (http://thomas.loc.gov).

Whereas this latter property of public laws yields a measure that reduces the need for researcher decisions, it is a measure that has generally not been used in the legislative productivity literature. Scholars interested in legislative productivity have rightly noted that simply counting laws without accounting for their content is likely to produce measurement error when attempting to measure policy production. Not all laws are created equal in their contribution to policy change. Still, there should be no doubt that lawmaking in general is an aspect of policy production. Thus, it measures one conceptual aspect of legislative productivity.

Because not all laws contribute to major policy change, most scholars have operationalized indicators that capture the production of important legislation selected based on contemporaneous evaluations by the media. The measures that fall under this category evaluate productivity during the postwar period. Each differs in their operationalization but hold in common the use of contemporaneous evaluations to evaluate how many major pieces of legislation were enacted. One indicator is Mayhew's (2005) list of laws from "Sweep 1," updated through 2002. This measure classifies a law as important if it was viewed by journalists and other informed observers as important legislation at the time it was enacted. A second is compiled by Howell et al. (2000). Their "A1" category is a list of laws defined as important based on news accounts and the length of coverage in *Congressional Quarterly* (CQ). A third indicator defines an important law as one which ranks in the top 10% of CQ's length of coverage (Howell et al. 2000). A fourth also focuses exclusively on CQ coverage. Baumgartner and Jones (2003) rank the top 500 "most important" laws based on their lines of coverage in CQ. Using recent updates to their data on CQ coverage, we identify the 500 most covered statutes through the 107th Congress.[5] Each of these four indicators provides measures of legislative productivity of important legislation by those witnessing and evaluating the legislative process at the time of enactment.

A second approach to operationalizing important policy changes for the postwar period is to draw from retrospective evaluations by policy experts. We include two for this time period. The first is Mayhew's (2005) "Sweep 2" laws. These are laws that policy experts identified as being important. Because of the lag between enactment and retrospective evaluation, this indicator is not scored past the late 1980s. The second indicator is the production of what Kelly (1993) defines as innovative legislation. These are laws that were identified both by contemporaneous media (Sweep 1) and by policy experts (Sweep 2).

The most obvious limitations of both of these approaches are the limited time periods they cover. At best, studies that analyze significant legislative production use data from the postwar period. By ignoring the first three-quarters of American history, previous research on legislative productivity ignores the development of lawmaking in Congress. To extend the measure of productivity for the previous Congresses requires new measures of productivity.

We have found several indicators, largely ignored by scholars, that gauge legislative productivity for Congresses prior to the 1940s. Measures of important legislation enacted prior to 1945 have not been widely used in political science, but we have found three candidates for inclusion in our measure. First, Christianson (1996) identifies laws during each Congress (1789–1994) that were "major enactments," and we utilize this count of important laws as an additional indicator of legislative productivity. Second, we include a series constructed by Stathis (2003) that "documents the Congress's most momentous productivity in determining national politics." A third indicator of historic legislation that

[5]Data available from the Policy Agendas Project, http://www.policyagendas.org/.

extends to the early years of the United States is borrowed from Castel and Gibson (1975), who replicate CQ's list of "key votes" prior to 1946. Included in our variable are items from this list of key votes that eventually became law.[6] Taken together, these historical measures help fill in the 155 years of history not analyzed in existing models of lawmaking and allow us to provide a more complete understanding of American macro politics.

These latter three series that count the number of important statutes enacted during a Congress overcome one of the problems with the most commonly used measures of legislative productivity. These measures include a much longer time span. More generally, the benefit of focusing on major legislative enactments is that such measures account for an aspect of content in the laws that they include—importance. Making important policy change is qualitatively different than passing more mundane laws. The problem with these measures, however, is that general lawmaking is completely ignored. Even the most arcane and mundane statutes contribute to policy production. To focus on either general or important lawmaking while excluding the other likely ignores a portion of policy production. We treat the issue of measuring productivity based on important policy production versus general policymaking versus a combination of the two as an empirical question. We examine three measures that account for the three possible strategies and conduct validation exercises to determine which is most sensible. One measure is simply the number of public laws considered separately. The second measure is an index that combines measures of important legislation. The third is an index that combines indicators of both important legislation and general lawmaking. In the next section, we discuss the creation of the two indexes and then move in Section 3 to testing the validity of each of the measures.

2.2 *Using W-CALC to Estimate Productivity*

Combining the indicators discussed above into indexes of legislative productivity is complicated by two problems. First, each series is on a different scale. Even for the series that report a number indicating an amount of important legislation, the fact that each series uses a different definition of what constitutes productivity makes direct comparison of movement over time between series technically inappropriate. The lack of direct comparability is particularly obvious for the measure that examines the total number of public laws passed. Second, there is discontinuity in the dates that each series examines. At one extreme, the count of all public laws is available for the entire series (108 Congresses). At the other, two series have data only for 20 Congresses. The result is a set of overlapping series, with all but one having gaps of time with no data.

We overcome these problems and estimate two macro indexes for each U.S. Congress from 1789 to 2004: Legislative Productivity Index (LPI), which measures legislative productivity including the policy impact of ordinary legislation, and Major Legislation Index (MLI), which measures the production of major legislation only. We make use of a method developed by Stimson (1999) that estimates a single measure of a latent concept using

[6]Two measures are not included. Clinton and Lapinski (2006) estimate levels of productivity (which they call achievement), but they have not made public their estimates or the data used to create them. Smock (1999) conducted an examination of landmark documents from the U.S. Congress. Several pieces of legislation that were enacted as law are included in Smock's analysis. The pieces of legislation mentioned in Smock's work can be extracted to create an indicator of historic legislation that extends back to the first Congress. A potential problem with this measure, however, is that although it includes many historic laws, it "does not focus on laws or attempt to present all possible candidates for the status of landmark legislation from Congress" (xii). Indeed, when added to our measure of legislative productivity, the measure from Smock (1999) does not perform particularly well as an indicator of the underlying concept. The correlation of the Smock series with the combined measure is mere 0.03, and excluding the Smock series increases the variance in the combined measure explained by the constituent series from 41% to 47%. Therefore, we exclude the Smock series.

multiple time series of indicators, each of which may be measured on a different scale and may be available for different periods of time. Stimson originally employed the technique to generate a measure of mass preferences, the Public Mood, using multiple series of issue preferences collected in national surveys. The W-CALC algorithm, however, is not limited to this one application (c.f. Stimson 1999).[7] We use this method to overcome the same underlying problems of scaling and coverage, but in our case to generate indexes measuring the legislative productivity concept.[8] Stimson's algorithm provides a way to accomplish this. Perhaps most importantly, the procedure allows us, after the creation of the combined measure, to take a step back and assess whether the constituent series actually tap a single latent concept as we have hypothesized.

The first step in the algorithm is converting each series to a consistent scale by expressing every datum—the value each series takes on at a particular point in time—as a ratio to a fixed reference point.[9] The process begins by setting the index at time t to an arbitrary value of 100. This gives us a way to place previous values of all the indicators on a comparable metric. For example, a series counting the number of landmark statutes (Stathis) has a value of 17 in the 107th Congress. We will call this period t. At $t-1$ (106th Congress), the series has a value of 12. We can convert the series to the metric in which the index is 100 at time t by the following operation:

$$\text{Stathis}_{t-k} = 100 \times \left(\frac{\text{Stathis}_{t-k}}{\text{Stathis}_t}\right). \tag{1}$$

Following the procedure, the series is rescaled such that its value in the 107th Congress is 100, and its value in the 106th is about 70.6. The same can be done for all other series that have a value for t and $t-1$. This rescales all the available series such that their value at t is 100 and their value at $t-1$ is a ratio to t. This yields multiple series with the same value for t and on a comparable metric for $t-1$ such that a central tendency for $t-1$ can be computed. This central tendency is legislative productivity.

Values for time points prior to $t-1$ can also be expressed as a ratio to t, but the procedure to this point would leave us with an index utilizing only indicators with a value for the final time point. We can go beyond this because we now have an essential new piece of information—an estimate for productivity in times prior to the final period. Thus, the procedure described above can be repeated recursively, starting at $t-1$ and proceeding to $t-2$ and so on back in time. Instead of using 100 as an arbitrary starting point for the rescaling procedure, we can utilize the estimated value of the index as the base metric.

[7]Our use of Stimson's algorithm does not, of necessity, imply a measure with ideological direction. Stimson's mood measure has ideological direction, but our measure of legislative productivity has no inherent ideological content. In fact, the use of this method does not require or imply a latent construct with ideological content. We simply apply Stimson's method to a similar kind of statistical problem.

[8]Analyzing legislative productivity over the entire history of the United States is simply not possible without making use of a strategy to combine these multiple measures available during various times and measuring various specific manifestations of policymaking. The only series that is available over the entire time period is the generic public laws measure. This variable, although related to legislative production, is not a full manifestation of the concept since it systematically devalues important enactments. This is similar to the situation confronted by Stimson. Although there are public opinion surveys asking the same question identically over the full period that Stimson analyzed (e.g., questions on liberal-conservative self-identification), these individual items do not fully capture the public mood, since it is a combination of attitudes toward a variety of issues. Likewise, legislative production is not fully captured by the one series—public laws—that is available over the entire history of the U.S. Congress. Thus, it becomes essential to combine these multiple measures in order to analyze legislative productivity over the full period.

[9]Our discussion of Stimson's algorithm borrows heavily from the description in Appendix 1 of Stimson (1999, 133–137).

Some indicators available at $t - 1$ will have already been considered because they were measured in the final period whereas others are new. The old and the new values are rescaled using the new information about the summary index at $t - 1$. This process of backwards recursion continues until every available data point from every available indicator has been rolled into the index. The algorithm works out a common metric and then averages all the data available at a given time:

$$\text{Index}_t = \frac{\sum_{i=1}^{n} \sum_{j=1}^{t} \frac{\text{Indicator}_j}{\text{Indicator}_i} \times \text{Metric}_b}{n}, \qquad (2)$$

where:
$i=1,n$ is all available indicators for period t,
$j=1,t$ is all available dyadic comparisons for indicator i,
b is the base period for the recursive metric generation, and
Metric_b is the value of the metric for period b (this is an arbitrary value that is initially set at 100 in the algorithm).

Starting with the last time period and working back may influence the values computed for the summary index. The problem is that the individual indicators are not all available over the same period of time. By starting at the end, as opposed to the beginning of the period under analysis, some series enter into the calculation more frequently and, thus, exert more weight on the final index. In order to ensure that working back in time does not unduly bias the index, the Stimson algorithm repeats the process described above, but starting at the beginning and moving to the end of the period under analysis. Averaging the index produced by backward recursion with that produced by forward recursion generates the final index.[10]

2.3 *Estimates*

We estimate two indexes. The first is the LPI. This index measures a concept of legislative production defined by the creation of policy change through the enactment of statutes. Policy change includes both ordinary and extraordinary laws. The LPI includes each of the indicators of major legislative change and the total number of statutes enacted. The total number of public laws provides a general indication of how much gets done in Congress and distinguishes our measure from those used in studies of legislative productivity, which have only examined important lawmaking. Despite the difference in magnitude between the number of important bills and the overall amount of legislation passed, movement in both should be similar over time if they both measure an aspect of policy production, as we posit that they do.

The second index is the MLI. This index is an estimate of legislative production of major policy changes. As we have noted above, most scholars have used indicators of major (alternatively landmark, important, or significant) legislation. The MLI includes each of these indicators, but it excludes the total number of statutes from the estimation of the index. As we discuss in more detail below, there is empirical evidence that suggests that the LPI is a better measure of productivity. However, there is a conceptual distinction, which we discuss below, that necessitates the estimation of a separate index for the production of major legislation.

Our measures of both indexes are plotted in Fig. 1 (a list of index values are available in the Appendix). In addition to the indexes, we include a plot of the total number of statutes

[10]The final measure is also transformed to match a "weighted average of the means and standard deviations of the included indicators. Weighting is by the final estimates of commonality, so that each item contributes to the scale metric in the same proportion as its contribution to the scale" (Stimson 1999, 137).

enacted. To visually compare movement of the series over time, we standardize each series. Both the LPI and the MLI follow similar patterns over time. The two series are correlated at 0.94. The differences in the series can be attributed to the inclusion of ordinary legislation. As the broader measure of productivity, the LPI is generally lower than the MLI during times when the production of ordinary legislation is decreased and higher during times when ordinary legislation is greater. The MLI is not influenced by the production of such legislation. Both the MLI and the LPI show less volatility than the total number of statutes enacted.

Productivity is very high at the start of the Congress. The first Congress had the highest level of productivity until the New Deal. The major policies that were enacted during this first Congress continue to have a direct influence on policy today. The first Congress set policies that formed the basic structure of the new government and set precedents for how the new nation would react to problems in the polity. Following these initial levels of productivity, the LPI and the MLI are low, particularly during the 1820s and 1870s. Many of the issues during this time dealt with small changes to patent laws, immigration, or the military. The major enactments often revolved around the expansion of U.S. territory. The one high point is at the start of the Civil War, a time in which supporters of the Union were able to pass significant policy changes in response to the Civil War. Following the Civil War, both the LPI and the MLI continue to increase steadily through the remainder of the 19th century.

During the 20th century, there were two time periods with exceptionally high levels of productivity. The responses to the Great Depression through the creation of the New Deal and other programs led to an unprecedented level of policy change during the 73rd Congress. The next major increase in productivity occurred after the 1964 election with the Great Society programs. The landslide victory of Johnson and the Democrats resulted in the creation of the Great Society programs, which were in many ways as expansive as the New Deal. Indeed, according to our measure, this was the time of greatest productivity in U.S. history. The most noteworthy developments of this time period were civil rights legislation and the creation of Medicaid and Medicare through the Social Security Act. Since the 1960s, productivity has waned. The MLI shows a spike in major legislation during the early 1980s; the LPI estimates that this spike is more muted. Both indexes, however, estimate that during the 1980s and 1990s, productivity has declined. Recent Congresses are producing major legislation at levels as low as during the 1950s. Overall productivity (the LPI) is currently at a level not seen since the 1920s.

The estimates plotted in Fig. 1 provide measurements of productivity prior to the 1870s, a time period previously unmeasured in the literature. Furthermore, it shows two alternative measures. The MLI draws information from the many indicators of major legislative accomplishment. The LPI broadens the concept of productivity to include the passage of the thousands of ordinary enactments not included by those indicators. Together, the MLI and the LPI provide the first glimpse of productivity for the entire history of Congress, adding to our description of Congress and opening the door for further study of Congress and lawmaking.

3 Validation

Validation is the process of evaluating how well a measure corresponds to its concept. Adcock and Collier (2001) discuss three types of validation: content validation (evaluating how well the measure includes aspects of the concept), convergent-divergent validation (evaluating how similar/different the measure is to other concepts), and construct validation (evaluating how well the measure relates to other concepts). We have discussed the

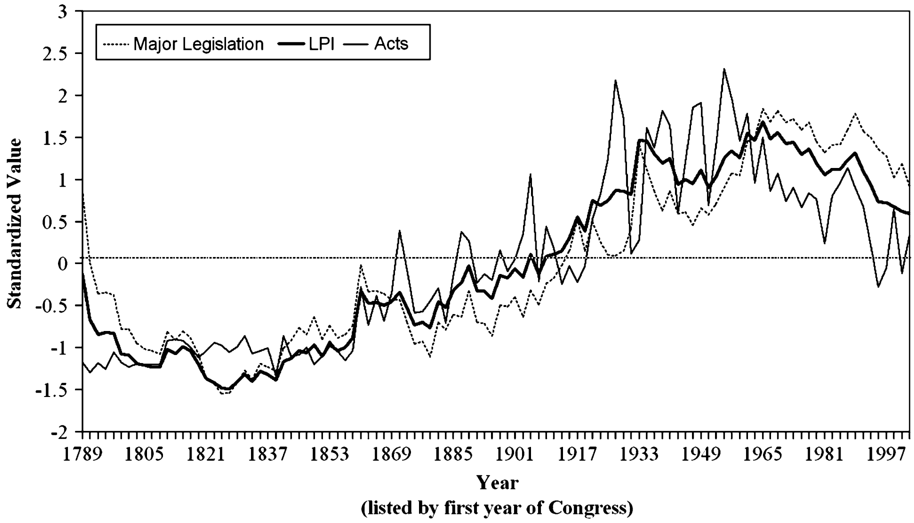

Fig. 1 LPI, MLI, and All Public Acts, 1789–2004.

content of the measure above. In this section, we examine the relationship of the LPI and the MLI with other measures and causal models of legislative productivity. In addition to addressing the issue of validity, we also seek to examine whether combining these multiple measures of policymaking provides an improvement over utilizing just one measure—public acts—the only other alternative for examining policy production through the full history of the United States. We will show that our indexes are valid measures of legislative policy production and are more theoretically and empirically advantageous than using only a single indicator of general lawmaking.

3.1 *Convergent-Divergent Validation*

Validation involves both convergent and divergent validations. Convergent validation is the process of demonstrating the similarity between a measure and other measures, often previously used measures of the same (or highly similar) concept. Divergent validity is used to differentiate measures. This is a critical part of validation because the explication of new measures involves a catch-22: the new measure must be similar enough to previous measures (convergent validity), but it should not be too similar to other measures. Part of divergent validity is demonstrating these differences. Divergent validity also includes comparisons to measures of concepts that are theoretically distinct from the new concept. Such differences provide evidence that the new measure is tapping the right concept and is not spuriously measuring a different concept.

To examine convergent validity, we first examine how well the constituent series fit together and perform as indicators of the underlying legislative productivity concept. In Figs 2 and 3, we plot the constituent series of the LPI and the MLI along with the indexes themselves, with all series transformed to 3 year moving averages of standardized values. The lighter lines represent the components of the indexes, whereas the heavy lines indicate the LPI (in Fig. 2) or the MLI (in Fig. 3). The first important aspect of the figures is that they graphically show how the indicators of legislative production move together over time. What we see is that these indicators appear to share a great deal of variance. Although the

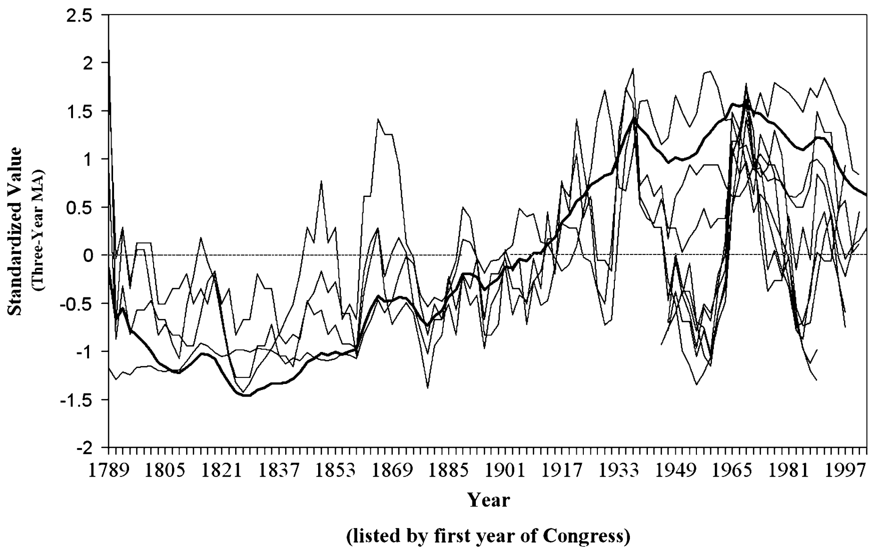

Fig. 2 Constituent indicators and the LPI (dark line is the LPI).

indicators vary somewhat in their degree of fluctuation from one period to the next (even after smoothing via a 3-year moving average), they generally seem to move in similar directions at similar points in time. In fact, many of the series overlap closely enough that it is impossible to draw this chart in a way in which the individual components can be individually identified. Most of the lines simply track with each other well enough that it is difficult for the eye to differentiate one line from the next. There are a few exceptions to this general rule, however. The most notable divergence from this pattern is present for the LPI during the 1940s to 1960s. The top line during this period is the public laws series, and its path diverges somewhat from the indicators of important lawmaking. This deviation is more remarkable in its level than in its movement over time. From 1941 to 1963, the public laws measure maintains a level substantially higher than the other indicators of legislative production. However, both the Stathis and Christianson series follow the same general trajectory during this period—little change in production between the beginning and end of the period with fluctuations up and down occurring at similar points in time. The cluster of series indicating the lowest levels of production are all series focusing on important legislation based on some form of print media coverage. During this period in which the indicators of legislative productivity diverge from one another, the LPI does what we would expect it to do—it takes account of both general and important lawmaking by taking on values lower than all public laws and higher than the series that count only stringently defined important legislation. The series begin to move in concert once again toward the end of the 1960s, with public laws again at a higher level but with fairly consistent movement over time across all the indicators.

We continue our convergent validation by reporting the correlation of the individual series with the LPI and the MLI. Table 1 corroborates the visual evidence seen in Fig. 2. All the individual indicators perform reasonably well, with the total number of acts and the Stathis series correlated most strongly with the summary measure. The contemporaneous evaluations by the media contribute to the index but do so with less weight than the other

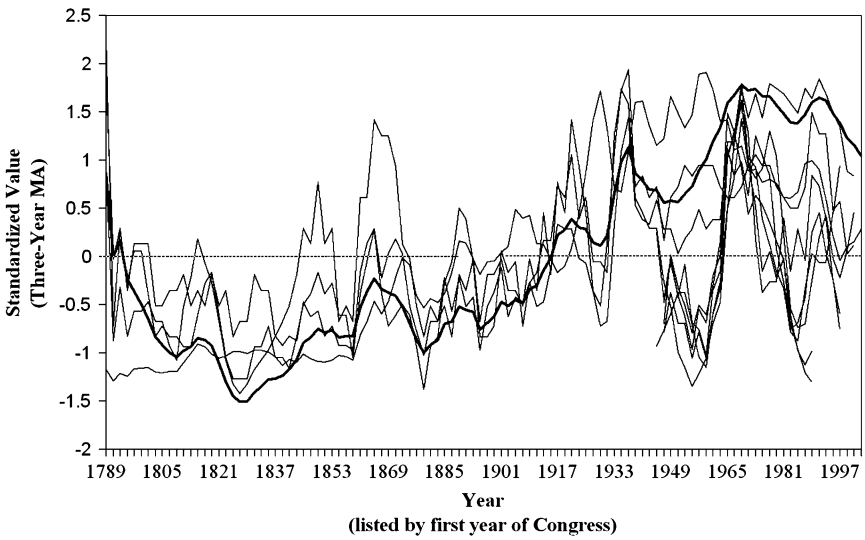

Fig. 3 Constituent indicators and the MLI (dark line is the MLI).

series. The LPI is clearly more than just an indicator of total laws since several measures of important legislation are also highly correlated with the summary measure and the correlation between the LPI and public acts is clearly not perfect.[11]

The relationships between the composite indicators and the MLI show how the LPI and the MLI are different from each other. The LPI has a stronger relationship to the total number of acts (Acts), which was used to estimate the LPI but not used in the MLI. The LPI also has a stronger relationship with some of the more stringent measures of important legislation (Mayhew's Sweep 2, Kelly, and Howell et al.'s more restrictive indicator of media evaluation). The MLI, however, is more strongly related to the other measures of contemporaneous evaluations, including Mayhew's Sweep 1 and measures using CQ. The MLI is also more strongly related to Castel's attempt to extend CQ's key votes for the rest of the history of Congress.

One possible critique of the LPI is that it is highly correlated with the number of statutes enacted: with such a high correlation, why bother with the LPI? Theoretically, they are different measures, but empirically, they may be correlated to such a degree that some may consider using the number of acts as a proxy measure. We would caution against this for both theoretical and empirical reasons. Table 1 provides two pieces of evidence which suggest that total number of statutes is not a substitute for the LPI. First, the MLI, which was estimated without any information on the total number of statutes, is also highly correlated with this variable. Second, although Acts is highly correlated with measures

[11]This is in stark contrast to the measure created by Clinton and Lapinski, which seems empirically to be little more than a measure of important policy production back in time to the late 1800s. Evidence of this is provided by Clinton and Lapinski (2006), who report a 0.94 correlation of their measure with the existing measures created by Mayhew (1991). Our measure differs enough from existing measures to provide an empirical improvement while maintaining enough overlap with the existing measures to show that our measure taps an underlying concept. Simply stated, the LPI measures both general lawmaking and important lawmaking combined rather than just one or the other. Importantly, our measure also covers approximately 100 more years of history than the CL measure.

Table 1 Correlations between indicators and productivity

			Correlation		
Indicator[a]	*Criteria*	*Years*[b]	*LPI*[c]	*MLI*[c]	*Acts*
Acts	Total number of public laws	1789–2004	0.87	0.73	1.00
Mayhew (Sweep 2)	Retrospective judgments	1947–1986	0.78	0.60	0.01
Stathis	Retrospective judgments	1789–2002	0.77	0.87	0.55
Kelly	Contemporaneous and retrospective	1947–1986	0.76	0.54	0.10
Christianson	Retrospective judgments	1789–1994	0.72	0.76	0.56
Howell et al. (A1)	Media evaluations	1945–1994	0.62	0.47	0.04
Howell et al. (CQ)	Media evaluations	1945–1994	0.43	0.54	−0.30
Baumgartner and Jones	Media evaluations	1949–2004	0.43	0.73	−0.17
Castel	Laws with key votes	1789–1946	0.38	0.55	0.19
Mayhew (Sweep 1)	Contemporaneous evaluations	1947–2002	0.29	0.54	−0.23

[a]Acts was used as part of the estimation of the LPI; it was not used in the estimation of the MLI. All other indicators were used to estimate both the LPI and the MLI.
[b]Years is a general description for the Congresses. Congresses also include some time in the year following the election. For example, the 108th Congress began in 2003 and continued until the start of the 109th in 2005.
[c]Correlation between the LPI and the MLI is 0.94.

of latent productivity, it is not correlated with the series used to estimate this latent variable. For time series variables, the correlations are weak; only two show a modest positive correlation. The rest are slightly positive or negative. Acts provides information necessary to estimate productivity, but tests of its relationship to other indicators suggest that it would be a poor substitute for the LPI.

Table 2 presents further tests of divergent validity. First, we report the correlation of legislative productivity with Binder's (2003) five measures of legislative gridlock. Binder's measures examine how much of the legislative agenda was not accomplished in each Congress (since 1947). Gridlock is conceptually distinct from legislative productivity. Indeed, we should expect to see a negative correlation between the two. This is exactly what we observe, with the correlation between the LPI and measures of gridlock ranging from –0.44 to –0.63. The correlations between the MLI and gridlock further show the empirical difference between the LPI and the MLI. Our index of major legislation, unlike our index of general productivity, is completely unrelated to gridlock. Notably, the LPI is an improvement over a focus only on general lawmaking. The LPI correlates more strongly with Binder's measures than does the public laws series, and the differences here are appreciable for most of Binder's measures.

Second, we continue our divergent validation by examining the relationships between productivity and activity. These are two distinct concepts. Activity is the amount of time and resources spent during legislating. Productivity is the amount of policy changes resulting from this activity. To measure activity, we use the total number of roll call votes in the House, Senate, and the chambers together. As shown in Table 1, activity is weakly related to productivity. In fact, activity may be negatively correlated. This is particularly the case when we examine the correlation between activity and general lawmaking, which clearly demonstrates that activity and productivity are distinct concepts.

Finally, we conclude this portion of the validation process by demonstrating the dissimilarity between our indexes and the aggregated measure of significance estimated by Clinton and Lapinski (2006). Our purpose is to highlight the reasons why the LPI and the MLI are more valid measures of productivity than those provided by the CL model. We

Table 2 Correlations between divergent measures

	Years[a]	Correlation		
		LPI	*MLI*	*Acts*
Gridlock measures (Binder 2003)				
Gridlock 1	1947–2000	−0.62	−0.04	−0.53
Gridlock 2	1947–2000	−0.63	0.02	−0.51
Gridlock 3	1947–2000	−0.55	0.06	−0.48
Gridlock 4	1947–2000	−0.56	−0.03	−0.43
Gridlock 5	1947–2000	−0.44	0.02	−0.32
Activity measures				
Number of roll call votes in the House	1789–2004	0.01	0.16	−0.48
Number of roll call votes in the Senate	1789–2004	0.31	0.39	−0.58
Sum of roll call votes in the House and the Senate	1789–2004	0.18	0.31	0.04
ln of sum of roll call votes in the House and the Senate	1789–2004	0.19	0.29	−0.72

[a]Years is a general description for the Congresses. Congresses also include some time in the year following the election. For example, the 108th Congress began in 2003 and continued until the start of the 109th in 2005.

conducted our comparison of the measures with caution. The CL data are not available.[12] To obtain the measures presented in the figures of Clinton and Lapinski (2006), we copied the figures onto a computer graphics program and then measured the pixel height of the plots presented in the article. Clinton and Lapinski report that their two series are correlated at 0.87; our estimates of their two series are correlated at 0.84. Thus, we are confident that our estimates provide sufficient measures for the purposes of comparing our indexes with their two aggregated measures.

Figure 4 presents the LPI, the MLI, and the two CL series. The CL series are the sum of the estimated significance of the 500 and 3500 most significant enactments, respectively. The MLI and the LPI are highly correlated during this time period ($r = 0.94$). Both the LPI and the MLI are more highly correlated with the 3500 threshold aggregation (0.77 and 0.75, respectively) than they are with the 500 threshold (0.47 and 0.43, respectively). The divergences that result in these low correlations reveal the difference in the variables being measured by our indexes and the CL aggregations. During this time period (1887–1994), the Congresses with the highest LPI and MLI scores are found in the 1960s (1965–66, 1969–70, 1961–62, and 1967–68) and at the start of the New Deal (1933–34). The time periods with the lowest productivity are estimated to be those of the 1800s. For the CL series, the estimates reveal a different pattern. At both thresholds, the most productive Congresses include the 65th Congress (1917–18) and the 77th Congress (1941–42). According to the 500 threshold aggregation, the 1880s (1881–1890) and the 1920s (1921–1930) each had *more accomplishments* than either the 1950s (1951–1960) or the 1980s (1981–1990). Indeed, the 500 threshold aggregation finds the fourth least accomplished Congress to be the 101st (1987–1988). The 3500 threshold series also presents some interesting results. The 65th Congress (1917–1918) is considered the sixth most accomplished during this time period. The 1920s are estimated to be a more productive than the 1950s or the 1980s. Together, we find both series (1) to be distinct from our indexes and (2) to have less face validity than either the LPI or the MLI.

[12]We have requested the data presented in the article. The authors responded that the data should be made available on their Web sites soon. At this time, the aggregations are not available. The need to compare our measurements to their aggregations requires that use our approximations based on their published figures.

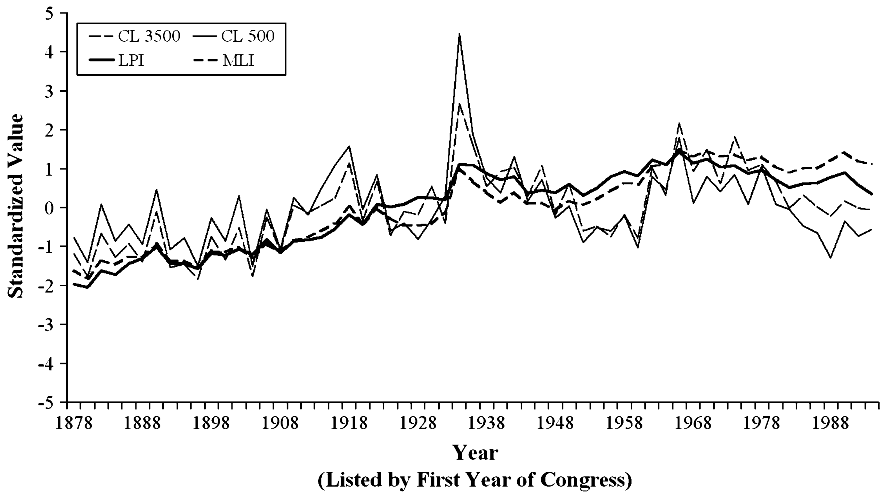

Fig. 4 Comparison of LPI and CL aggregations at 500 and 3500 thresholds.

In summary, our convergent-divergent validation reveals that the LPI and the MLI are similar enough to other series to suggest that they are valid measurements. Yet, they are distinct enough from these other measures to indicate that the LPI and the MLI are measuring unique macro-level variables. This is encouraging given that both the LPI and the MLI are the only measures, other than the count of all laws, that are available for the entire history of Congress.

3.2 *Construct Validity*

Further validation of our indexes may be accomplished by conducting construct validation tests, that is, by evaluating explanations of legislative productivity. We model legislative productivity using a lagged dependent variable (LDV) model. This model is based on previous findings from models of productivity from the postwar period. Because we have lacked measures of productivity for longer time periods, we must be cautious in estimating models. Our goal is to examine the relationship between our measures of productivity and variables that might, according to previous research, determine levels of productivity. We assess the relationship of the LPI, the MLI, and the total number of statutes with measures of partisanship and societal demands.

The separation of powers created by the U.S. Constitution has implications for law-making, a point not lost on those who have conducted work on legislative production in the past. Since Mayhew's (1991, 2005) seminal work on divided government and legislative productivity, there has been no shortage of work examining and reexamining the question of whether divided government influences the policymaking process. Divided government has a negative impact on lawmaking when only the most highly important laws are considered, but with even a slight relaxation of the criteria for importance, the effects of divided government deteriorate rapidly (Howell et al. 2000). We identify any Congress in which the same party controls the House, Senate, and the executive branch as being unified. All others are instances of divided government. In our models, we code divided government as one. Unified governments received a value of zero.

In addition, we examine changes from unified to divided government. Congresses in which a party gains unified control should produce more policy changes than in the previous Congress in which power was divided. Likewise, Congresses that change from unified to divided should be less productive. Our models thus include change in divided government (1 = change from unified to divided, 0 = same type of party regime as previous Congress, and −1 = change from divided to unified control).

There is also the potential for productivity to be depressed by increased conflict between parties within Congress. We operationalize partisanship as the difference in preferences between the mean DW-NOMINATE score for the two largest parties in each chamber.[13] We then take the average of the differences for the House and Senate. The larger this measure of partisanship, the less productive Congress should be.

Productivity should be determined by more than partisanship; it should also be influenced by changes in societal demands, as measured by the public's political preferences. When citizens ask for more (or less) government action, more (or less) legislative production should follow. To tap the general ideological preferences of the mass public, we make use of Stimson's (1999) measure of the Public Mood. Mood is a summary measure of public preferences for more or less government action in a variety of policy domains. When Public Mood liberalism increases, we expect to see increases in productivity.

Due to its reliance on mass opinion surveys, the measure of Mood is limited to the past five decades. To measure societal demands for the entire time period, we use a measure of technological advances. Technological advances create new economies and behaviors that were not previously in existence. The invention of the steam locomotive creates a demand for new types of transportation policies. The industrial revolution creates new dilemmas for consumer and labor protections. New forms of communication alter the way information is gathered and used. The explosion of computer-related inventions changes worker productivity, communication, and the meaning of privacy. Our measure of technological innovations is the number of patents for inventions that were approved during each Congress.[14] Other indicators of societal demand or technology are not available reliably prior to the beginning of the 20th century. Census counts of population are decennial; measures of gross domestic product and other economic measures were not reliably measured prior to the 1930s. Furthermore, patents provide a more reasonable indicator of general technological advancement than more specific measures such as length of railroad track, number of telephone subscribers, or miles of highway. To compensate for the diminishing effects of new technology, we include the natural log of this number in our model.

We estimate LDV models using ordinary least squares. This model provides a conservative estimation of construct validity—the inclusion of the lagged variable is necessary but has the potential of reducing the explanatory power of the other variables.[15] We regress each measure of productivity on divided government, changes in unified/divided government, differences between parties, and societal demands (patents or policy mood). Table 3 presents our estimates for the entire history of Congress. Table 4 presents models for 1950

[13]Data on DW-NOMINATE and roll call votes taken from DW-NOMINATE files, http://voteview.com/. See Poole and Rosenthal (1997), for more information on scores.

[14]Data on the number of patents comes from the Historical Statistics of the United States and the U.S. Patent Office Patent Technology Monitoring Division table "U.S. Patent Activity Annual Years 1790 to Present", http://www1.uspto.gov/go/taf/h_counts.htm, September 29, 2004 and March 27, 2006.

[15]On the other hand, if the underlying metrics of the series are changing over time, inclusion of the LDV might bias the construct validity tests in our favor. Inclusion of the LDV means that only short-term changes are being explained in these models. We might observe less construct validity if the explanatory variables in these models were forced to explain long-term changes. In order to account for this possibility, we also estimated these models using regression with Newey-West standard errors. This strategy accounts for autocorrelation without the inclusion of a LDV. The substantive conclusions using this approach are unchanged, which suggests that the underlying metric of the indexes is stable over time.

Table 3 Models of productivity (1st–108th Congresses)

Independent variables	*LPI*	*MLI*	*Public Laws*
Divided government$_t$	0.84 (2.22)	0.36 (0.27)	–0.11 (0.07)
Δ Divided government$_t$	–5.59* (1.86)	–0.80* (0.23)	–0.02 (0.06)
Party differences$_t$	–4.79 (6.72)	–1.10 (0.78)	–0.42* (0.21)
Patents issued (ln)$_t$	2.23* (0.649)	0.24* (0.06)	0.14* (0.03)
LDV	0.91* (0.03)	0.88* (0.03)	0.50* (0.09)
Constant	–9.55* (5.38)	–0.37 (0.67)	1.86* (0.37)
Adjusted R^2	0.98	0.96	0.87
N	107	107	107

Note. Each model estimated using ordinary least squares. Time period in years covers 1789 through 2004. "Public Laws" is the ln of the total number of statutes enacted. Model regressing total number of statutes provides nearly identical inferences as the ln model.
*$p < 0.05$, two-tailed.

through 2004; these models include mood rather than patents as a measure of societal demand. Together, the models suggest that (1) the measures display construct validity but (2) each measure is the result of different processes.

The LPI displays the strongest construct validity. The results of both models demonstrate that the LPI is determined by both partisanship and societal demands. For the model estimated using data from every Congress, the effect of partisanship is manifest in the change in divided government variable. Congresses, where there is a transition from divided to unified government, are more productive; Congresses that change from unified control to divided are less productive. In more recent Congresses, productivity is driven by party polarization. Congress is more productive when parties hold similar positions than when parties are more distant from each other. These findings hold despite the inclusion of the lagged variable (which is a strong predictor in each model).

The models in Tables 3 and 4 also demonstrate how the LPI is similar and different from the MLI. Using data from the entire set of Congresses, the MLI is driven by a process similar to the one determining the LPI; both the LPI and the MLI are determined by changes in divided/unified government and societal demand, as measured by technological innovation. Using data (including policy mood) since 1950, models of the MLI and the LPI

Table 4 Models of productivity (81st–108th Congresses)

Independent variables	*LPI*	*MLI*	*Public Laws*
Divided government$_t$	3.50 (3.80)	0.33 (0.44)	0.13 (0.07)
Δ Divided government$_t$	–2.54 (3.37)	–0.19 (0.39)	–0.08 (0.07)
Party differences$_t$	–56.86* (21.17)	–3.58 (21.17)	–2.69* (0.49)
Policy mood$_t$	0.82* (0.36)	0.07 (0.04)	0.03* (0.01)
LDV	0.73* (0.12)	0.88* (0.09)	–0.22* (0.18)
Constant	25.39 (35.66)	–0.05 (3.50)	7.48* (1.25)
Adjusted R^2	0.83	0.77	0.72
N	27	27	27

Note. Each model estimated using ordinary least squares. Time period in years covers 1949 through 2004. "Public Laws" is the ln of the total number of statutes enacted. Model regressing total number of statutes provides nearly identical inferences as the ln model.
*$p < 0.05$, two-tailed.

Table 5 Models of productivity (45th–103rd Congresses)

Independent variables	*LPI*	*MLI*	*CL 500*	*CL 3500*
Divided government$_t$	–0.05 (0.08)	0.11 (0.07)	–0.34 (0.31)	–0.05 (0.08)
Δ Divided government$_t$	–0.12 (0.08)	–0.21* (0.06)	–0.46 (0.26)	–0.12 (0.08)
Party differences$_t$	–2.05* (0.06)	–1.35* (0.38)	–1.51 (1.06)	–2.40* (0.93)
Patents issued (ln)$_t$	0.40* (0.15)	0.56* (0.14)	0.14 (0.31)	0.62* (0.26)
Lagged dependent variable	0.54* (0.13)	0.56* (0.09)	0.32* (0.13)	0.21 (0.14)
Constant	–2.72 (1.42)	–5.43* (1.50)	–0.43 (3.92)	–5.23 (3.17)
Adjusted R^2	0.95	0.95	0.27	0.53
N	58	58	58	58

Note. Each model estimated using ordinary least squares. Time period in years covers 1877 through 1994. See text for measurement details. Each dependent variable is the same as presented in Fig. 4, which uses the series standardized over the time period.
*$p < 0.05$, two-tailed.

display differences. The MLI during this time period is driven simply by its lag; the remaining differences between Congresses are not due to partisanship or societal demands.

We also estimate models of legislative productivity as measured by the total number of laws enacted. The divergent validation results presented above suggest that the LPI, which includes ordinary legislation, should not be substituted with a count of enactments. Reporting this distinction is important because the public laws series is the only other series available over the full history of Congress and because the public laws series is used (along with other indicators) to estimate the LPI. The results of our models of the public laws series demonstrate the process determining the LPI, and the total number laws are similar but there are important differences. In Table 3, our model of public laws shows that both the LPI and public laws are determined by partisanship, societal demands, and the lag; yet, there are important differences in how these are manifest. First, the number of public laws is not determined by divided government. Rather, it is driven by party differences. Second, the magnitude of the lag is much lower for the public laws model, suggesting a longer memory for the LPI and less variability over time. The model in Table 4 displays this second difference in starker contrasts. Although both the LPI and the total number of public laws have each, since 1950, been determined by party differences and policy mood, the effect of the lag is very different. The LPI is strongly determined by its previous values; the same is not the case for public laws. Each Congress's number of laws enacted is determined by only partisanship and societal demands; there is no influence of the previous Congress on productivity of ordinary enactments.

Finally, we conclude our construct validation by estimating models of productivity as measured by the LPI, the MLI, and the two CL series (500 and 3500 thresholds). Because the models use data for Congresses prior to 1950, we use technological innovations (patents) rather than public mood; the models have the same specifications as those in Table 5. For both the LPI and the MLI, the results are consistent with the received wisdom on Congress and productivity. Productivity is determined by partisanship, particularly the policy differences between the parties in Congress. For the MLI, a change from unified to divided government results in a reduction in the production of major legislation. Both measures are also determined by societal demands (patents) and by their previous productivity (lag variable). The CL measures have different relationships with these variables, relationships that suggest that they do not have strong construct validity. The 500 statute threshold series is most problematic. Other than the effect of a weak lag, none of the variables influences this measure of productivity. The lower threshold aggregation (3500 statutes) is better, but the model has a weak fit. For the models of the LPI and the MLI, the percent of variance explained is at a level one would expect from time series models estimated with a LDV (95% for each model). The percentage for the CL models fare much worse, with the 500 threshold model explaining only a quarter of the variance and the 3500 threshold explaining about half. These are unusually poor levels of fit for LDV models, due in large part to the weak connection between the CL measures and their lag. That legislative productivity at time *t* is not strongly connected to lagged levels of productivity seems unlikely. These results taken as a whole indicate weak construct validity of the CL aggregations.

Together, our construct validation results in several conclusions. First, the LPI, the MLI, and the total count of enactments each display construct validity. Given the weakness of our theories of productivity, it is difficult to make firm conclusions about construct validity, but the results indicate that all three are driven by processes that include partisanship and societal demands. Second, the LPI has the strongest construct validity with variables that we would expect to drive legislative productivity. Using data spanning the entire history of Congress or data from recent Congresses with alternative measures, we see the expected relationship between partisanship, societal demands, and productivity. Third, the LPI, the MLI, and public acts are distinct. These three measures may be highly correlated with each other, but each of the three measures is driven by a different process. Finally, models of alternative measures estimated by Clinton and Lapinski (2006) fail to have the same level of construct validity as either the LPI or the MLI. The measure based on the top 500 laws is particularly problematic; other than a weak lagged effect, this aggregation measure appears to be unrelated to either partisanship or societal demands. Both models of the CL aggregations have very weak explanatory power, despite the inclusion of the LDV. Together, we find the results of our construct validation to strongly suggest that our measures, particularly the LPI, are valid measures of productivity.

4 Conclusion

We have presented a methodology for measuring legislative productivity for the entire history of the U.S. Congress. We have argued that our measure has content, convergent, and construct validity. Specifically with regard to convergent and construct validity, we have presented data showing that the LPI and the MLI are correlated with related concepts and that causal models produce results consistent with theoretical expectations. Our results support the conclusion that the LPI is a valid measure of legislative production for the entire history of Congress, and the MLI is a valid measure of the production of major legislation. The measurement of these indexes will allow future research to assess differences in policymaking over time and to find historically generalizable explanations for policy production. This is not possible based on existing measures of legislative productivity.

The measurement methodology used to construct our indexes is not limited to the measurement of productivity in the U.S. Congress. There are many concepts in political science that have competing measures of the same latent concept. As with productivity, these alternative indicators tap different aspects of the concept. They may or may not be available for the same time periods. Using a similar approach to the one we have taken, other scholars should be able to use information from competing measures to estimate a single measure of the underlying concept for a longer time period than would be possible if only one indicator was chosen. The result will be more theoretically grounded and more empirically powerful measures of political phenomena.

Appendix

Table A1 LPI by Congress and first year of Congress

Congress	*Year*	*LPI*	*Congress*	*Year*	*LPI*	*Congress*	*Year*	*LPI*
1	1789	82.1	37	1861	70.7	73	1933	173.9
2	1791	51.2	38	1863	62.1	74	1935	173.3
3	1793	40.9	39	1865	62.6	75	1937	164.4
4	1795	42.3	40	1867	60.8	76	1939	158.0
5	1797	41.8	41	1869	63.0	77	1941	161.1
6	1799	27.4	42	1871	69.6	78	1943	143.6
7	1801	26.9	43	1873	59.2	79	1945	147.0
8	1803	21.1	44	1875	47.2	80	1947	144.2
9	1805	20.0	45	1877	49.5	81	1949	153.2
10	1807	18.6	46	1879	45.7	82	1951	141.4
11	1809	18.5	47	1881	63.3	83	1953	148.7
12	1811	30.7	48	1883	59.4	84	1955	161.3
13	1813	28.1	49	1885	70.9	85	1957	166.1
14	1815	32.5	50	1887	76.1	86	1959	161.8
15	1817	29.7	51	1889	88.1	87	1961	178.5
16	1819	20.3	52	1891	70.5	88	1963	173.8
17	1821	10.6	53	1893	70.6	89	1965	186.6
18	1823	8.1	54	1895	65.4	90	1967	174.9
19	1825	4.2	55	1897	81.3	91	1969	178.8
20	1827	3.9	56	1899	79.5	92	1971	171.3
21	1829	8.6	57	1901	86.2	93	1973	172.3
22	1831	14.2	58	1903	79.8	94	1975	164.1
23	1833	8.9	59	1905	95.9	95	1977	167.9
24	1835	15.8	60	1907	82.4	96	1979	157.8
25	1837	13.7	61	1909	94.3	97	1981	149.9
26	1839	10.0	62	1911	95.8	98	1983	153.8
27	1841	22.6	63	1913	98.0	99	1985	154.0
28	1843	24.8	64	1915	107.0	100	1987	160.5
29	1845	30.0	65	1917	121.3	101	1989	164.8
30	1847	28.5	66	1919	111.8	102	1991	152.5
31	1849	34.1	67	1921	132.5	103	1993	143.3
32	1851	26.3	68	1923	129.3	104	1995	131.7
33	1853	34.2	69	1925	132.3	105	1997	131.0
34	1855	29.7	70	1927	139.6	106	1999	128.3
35	1857	32.0	71	1929	138.8	107	2001	124.8
36	1859	38.6	72	1931	136.6	108	2003	123.7

Table A2 MLI by Congress and first year of Congress

Congress	*Year*	*MLI*	*Congress*	*Year*	*MLI*	*Congress*	*Year*	*MLI*
1	1789	15.4	37	1861	10.9	73	1933	18.2
2	1791	11.0	38	1863	9.3	74	1935	16.7
3	1793	9.1	39	1865	9.3	75	1937	15.4
4	1795	9.2	40	1867	9.2	76	1939	14.2
5	1797	9.1	41	1869	8.7	77	1941	15.4
6	1799	7.0	42	1871	8.8	78	1943	14.0
7	1801	7.0	43	1873	7.4	79	1945	14.1
8	1803	6.1	44	1875	6.1	80	1947	13.3
9	1805	5.8	45	1877	6.3	81	1949	14.3
10	1807	5.7	46	1879	5.3	82	1951	13.9
11	1809	5.5	47	1881	7.4	83	1953	14.6
12	1811	6.8	48	1883	7.0	84	1955	15.5
13	1813	6.4	49	1885	7.9	85	1957	16.4
14	1815	6.9	50	1887	7.8	86	1959	16.3
15	1817	6.5	51	1889	9.3	87	1961	18.4
16	1819	5.5	52	1891	7.4	88	1963	18.6
17	1821	4.0	53	1893	7.4	89	1965	20.3
18	1823	3.7	54	1895	6.6	90	1967	19.5
19	1825	3.1	55	1897	8.5	91	1969	20.2
20	1827	3.2	56	1899	8.4	92	1971	19.5
21	1829	3.8	57	1901	9.0	93	1973	19.7
22	1831	4.5	58	1903	7.7	94	1975	19.1
23	1833	4.0	59	1905	9.4	95	1977	19.5
24	1835	4.9	60	1907	8.5	96	1979	18.4
25	1837	4.7	61	1909	9.8	97	1981	17.7
26	1839	4.5	62	1911	10.1	98	1983	18.1
27	1841	5.9	63	1913	10.9	99	1985	18.2
28	1843	6.3	64	1915	11.7	100	1987	19.1
29	1845	7.1	65	1917	13.7	101	1989	20.0
30	1847	6.7	66	1919	11.7	102	1991	19.0
31	1849	7.7	67	1921	13.5	103	1993	18.6
32	1851	6.4	68	1923	12.3	104	1995	17.9
33	1853	7.2	69	1925	11.5	105	1997	17.5
34	1855	6.5	70	1927	11.4	106	1999	16.2
35	1857	6.7	71	1929	11.7	107	2001	17.0
36	1859	7.2	72	1931	13.1	108	2003	15.7

References

Adcock, Robert, and David Collier. 2001. Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review* 95:529–46.

Baumgartner, Frank R., and Bryan D. Jones. 2002. *Policy dynamics.* Chicago, IL: Chicago University Press.

Baumgartner, Frank R., and Bryan D. Jones. 2003. Representation and agenda setting. Paper presented at the 2003 annual meeting of the American Political Science Association. Philadelphia, PA.

Baumgartner, Frank R., and Bryan D. Jones. 2004. The evolution of American government: Information, attention, and policy punctuations. Unpublished manuscript. Pennsylvania State University.

Binder, Sarah. 1999. The dynamics of legislative gridlock, 1947–1996. *American Political Science Review* 93:519–33.

Binder, Sarah. 2003. *Stalemate: The causes and consequences of legislative gridlock.* Washington, DC: Brookings Institution Press.

Castel, Albert, and Scott L. Gibson. 1975. *The Yeas and Nays: Key Congressional Decisions, 1774–1945.* Kalamazoo, MI: New Issues Press, Institute of Public Affairs, Western Michigan University.

Christianson, Stephen G. 1996. *Facts about the Congress.* New York: The H.W. Wilson Company.

Clinton, Joshua D., and John S. Lapinski. 2006. Measuring legislative accomplishment, 1877–1994. *American Journal of Political Science* 50:232–49.

Coleman, John J. 1999. Unified government, divided government, and party responsiveness. *American Political Science Review* 93:821–35.

Howell, William E., E. Scott Adler, Charles Cameron, and Charles Riemann. 2000. Divided government and the legislative productivity of Congress. *Legislative Studies Quarterly* 25:285–312.

Kelly, Sean. 1993. Divided we govern? A reassessment. *Polity* 25:475–84.

Krehbiel, Keith. 1998. *Pivotal politics: A theory of U.S. lawmaking.* Chicago, IL: University of Chicago Press.

Mayhew, David R. 1991. *Divided we govern: Party control, lawmaking and investigations, 1946–1990.* 1st ed. New Haven, CT: Yale University Press.

Mayhew, David R. 1993. Let's stick with the longer list. *Polity* 25:485–8.

Mayhew, David R. 2005. *Divided we govern: Party control, lawmaking and investigations, 1946–2002.* 2nd ed. New Haven, CT: Yale University Press.

Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A political-economic history of roll call voting.* New York: Oxford University Press.

Smock, Raymond. 1999. *Landmark documents on the U.S. Congress.* Washington, DC: CQ Press.

Stanley, Harold W., and Richard G. Niemi. 2007. *Vital statistics on American politics 2005–2006.* Washington, DC: CQ Press.

Stathis, Stephen. 2003. *Landmark legislation: 1774–2002.* Washington, DC: CQ Press.

Stimson, James A. 1999. *Public opinion in America: Moods, cycles, and swings.* 2nd ed. Boulder, CO: Westview Press.

U.S. Department of Commerce, Bureau of the Census. 1975. *Historical statistics of the United States, Colonial Times to 1970.* Washington, DC: U.S. Government Printing Office.