The Informational Value of Legislative Records*

Brad L. LeVeck† and Stephanie A. Nail

Department of Political Science, University of California, Merced

†correspondence: bleveck@ucmerced.edu

Abstract

Existing research argues that parties use votes in Congress to create ideological brands for themselves. These brands furnish voters with information about how a party's legislators will vote across a number of issues. However, the literature lacks a summary measure that quantifies the amount of information that is generated by each party's activities in Congress. Here, we propose such a measure. Using this measure, we compare the information in parties' legislative records to the voting records of individual incumbents. We show that the records of parties and individual incumbents have become almost equally informative over the past four decades. This finding may help explain a number of phenomena in congressional elections, such as the recent increase in party centered voting and the recent decrease in the incumbency advantage.

*Prepared for the 2018 Western Political Science Association Meeting in San Francisco, CA. Presented on May 30, 2018.

Party labels serve as informational cues that voters can use to learn about their representatives in Congress (Dancey and Sheagley 2013; Ansolabehere and Jones 2010; Popkin 1994; Snyder and Ting 2003, Brady and Sniderman 1985). Theories of lawmaking argue that the information conveyed by these party labels is the result of purposeful action by congressional parties (Cox and McCubbins 2007; Snyder and Ting 2003). Largely through roll call votes, parties try to create an ideological "brand"; and, this brand furnishes voters with information about how candidates from the party will vote across a number of issues (Cox and McCubbins 2007; Snyder and Ting 2003; Woon and Pope 2008; Kim and LeVeck 2013).

Yet — despite the centrality of these arguments to research on lawmaking and congressional elections — there is no widely accepted measure that quantifies *the amount* of information that parties produce through their activities in Congress. That is, how much information about candidates' ideology is created by each party's legislative record?

Furthermore, there is no accepted method for comparing the information in party records to the information contained in other types of legislative records. For example, other research argues that voters react to the legislative records of individual incumbents — punishing incumbents who are ideologically out of step (Canes-Wrone, Brady, and Cogan 2002). Presumably this is because past votes in Congress provide voters with useful information about how an incumbent will vote in the future. However, we do not know precisely how much information an incumbent's record provides about their future votes. Nor do we know how much information is conveyed by an incumbent's individual record relative to information conveyed by their party's collective record.

Here, we study congressional records through the lens of information theory (Shannon 1948). Specifically, we use an information-theoretic measure, the Jensen-Shannon Divergence (JSD) (Lin 1991), to characterize the ideological information conveyed by legislative records. The JSD considers the uncertainty an observer might have when data can come from two or more distributions. For example, absent a party label, one might be quite uncertain about whether a legislator will cast a liberal or conservative vote on a particular bill. The JSD then quantifies how much uncertainty is reduced by adding labels (e.g. "Democrat" or "Republican"), such that an observer knows which distribution is generating the data.

After describing the JSD and how it can be applied to voting data, we present an illustrative application, and use the JSD to measure the informational value of party records regarding ideological voting on the left–right dimension (liberal-conservative). We then compare this to the information contained in individual legislators' records. We show that both types of records have conveyed increasing amounts of information since the 1970s. However, the informational value of party records has risen faster than the informational value of individual records. As a result, party records are now almost equally as informative regarding a candidate's ideology as the candidate's own record. We conclude by briefly discussing how this finding may shed light on a number of phenomena, such as increases in party centered voting (Bonica and Cox Forthcoming) and the decline of the incumbency advantage (Jacobson 2015).

Method

Here, we begin by briefly introducing two key concepts from information theory, which may be unfamiliar to many political scientists: *entropy* and *mutual information*. Readers who are already familiar with these concepts may wish to skip over these sections, and go directly to the section on the Jensen-Shannon Divergence.

Entropy

The JSD is based upon Shannon Entropy, which is defined by:

$$H(X) = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i) \tag{1}$$

$H(X)$ is a measure of uncertainty about a discrete random variable X. For example, X might be a binary random variable that represents whether a legislator casts a liberal or conservative vote on a particular bill. $H(X)$ is then maximized when there is an equal probability of each value x_i, which means that this distribution maximizes uncertainty. Therefore, continuing with the previous example, uncertainty about how the legislator will vote is highest when the legislator casts liberal and conservative votes with equal probability, in which case $H(X) = 1$. On the other hand, there will be no uncertainty if the legislator always casts liberal (or conservative) votes, in which case $H(X) = 0$.

The entropy of a random variable, $H(X)$, can also be interpreted as a measure of how much information is revealed by a data generating process. Under this interpretation, realizations of X convey more information if you are more uncertain prior to observing a given realization. For example, seeing a legislator cast a liberal vote will convey no new information if you already know that the legislator always takes the liberal side of an issue. However, it will convey quite a bit of information if you initially believe there is a 50/50 chance that the legislator will cast a liberal or conservative vote (i.e. this is the situation where entropy is maximized).

Mutual Information

Mutual information is defined by the equation

$$I(X,Y) = H(X) - H(X|Y) \tag{2}$$

where, $H(X|Y) = \sum_{j \in M} H(X|y_j)P(y_j)$. Because $H(X) \geq H(X|Y)$, mutual information is always positive, and is a measure how much the entropy of X is reduced if you know the realization of another variable, Y. For example, Y might represent the party of a particular legislator. If party affiliation is highly correlated with the ideology of a legislator's votes, then knowing Y (the party of a legislator) may substantially reduce one's uncertainty about X (whether the legislator takes the conservative or liberal side of a particular vote).

An alternative interpretation of $I(X,Y)$ is that it is a measure of the quantity of information Y provides about X. Under this interpretation, knowing Y will only provide you with new information about X if you are initially uncertain about X. To see this, note that if $H(X) = 0$ (i.e. there is zero uncertainty), then $I(X,Y) = 0$ as well. Furthermore, Y only provides information about X

to the extent X and Y are correlated. To continue the example above, if you are uncertain about a legislator's position, knowing their party affiliation will provide information to the extent that liberal or conservative votes are correlated with being a legislator from a particular party.

The Jensen-Shannon Divergence

The Jensen-Shannon divergence (Lin 1991) generalizes the concepts of entropy and mutual information to encompass situations where an observer knows that data is generated by one of n distributions. It then characterizes how much uncertainty is reduced if each of the n distributions are labeled, such that an observer knows exactly which distribution is generating a given set of data. In essence, the JSD is the mutual information between the labels and the aggregate data (Lin 1991).

$$JSD_{\pi_1,\cdots,\pi_n(P_1,\cdots,P_n)} = \overbrace{H(\sum_{i=1}^{n} \pi_i P_i)}^{\substack{\text{Uncertainty over a mixture} \\ \text{of } n \text{ unlabeled distributions}}} - \overbrace{\sum_{i=1}^{n} \pi_i H(P_i)}^{\substack{\text{Average uncertainty} \\ \text{of } n \text{ labeled distributions}}} \tag{3}$$

In equation 3 above, $\pi_1 \cdots \pi_n$ are the weights assigned to each distribution P_i. Usually these weights are simply $\pi_i = 1/n$ for all n distributions, but they could be adjusted to reflect the prior probability that data comes from a particular distribution. When entropy is defined using logarithms with base 2 (as in equation 1), the JSD is bounded between a minimum of 0 and a maximum of 1.

To briefly illustrate how the JSD might be used in the case of party records, let's assume that we are interested in guessing whether a legislator will vote yea or nay on a bill. Furthermore, assume that a legislator's vote (yea or nay) can be interpreted as taking the liberal or conservative side of the issue along a single left-right ideological dimension.[1] Given these assumptions, we could use equation 4 below to measure the information gained by knowing a legislator's party label. In this equation, *Dem* and *Rep* are probability distributions over a binary random variable that scores liberal votes as 0 and conservative votes as 1. An observer might estimate each of these distributions by using each party's legislative record in Congress. Therefore, consistent with the literature on

[1] The JSD does not require that we restrict ourselves to a single dimension. This is just to simplify the example.

partisan lawmaking, equation 4 implies that party labels are informative because they are linked to specific legislative records, which encode ideological brands (Cox and McCubbins 1993, 2007; Snyder and Ting 2003; Woon and Pope 2008). Also, equation 4 implies that the information in party labels depends jointly on the actions of both parties. Therefore, the party JSD measures the amount of information that is generally contained in parties' legislative records, rather than the information contained in any specific party's legislative record.

$$JSD_{\frac{1}{2}(Dem,Rep)} = \overbrace{H\left(\frac{1}{2}Dem + \frac{1}{2}Rep\right)}^{\substack{\text{Uncertainty given that a vote is} \\ \text{cast by a member of either party}}} - \overbrace{\left(\frac{1}{2}H(Dem) + \frac{1}{2}H(Rep)\right)}^{\substack{\text{Average uncertainty given that a vote is} \\ \text{cast by a member of a specific party}}} \quad (4)$$

The party JSD is calculated over the parties' entire legislative record because it is a measure of how much information is *produced* by the parties' legislative activities in Congress. It is not a measure of how much information is consumed by any particular voter. However, one could easily calculate the JSD on a subset of bills that voters are more likely to know about. For instance, researchers could calculate the party JSD for major legislation (Mayhew 1991). In this case, the party JSD would reflect the information that parties' legislative records provide regarding congressional votes on salient bills.

Application

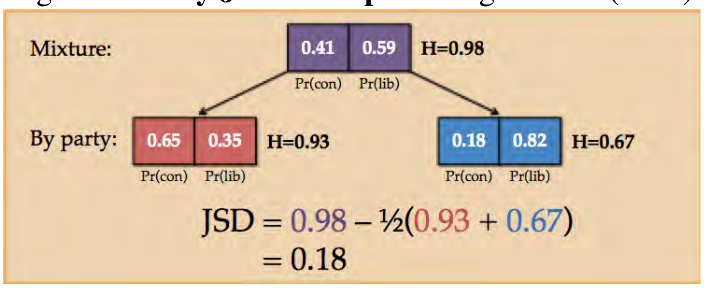
To illustrate how the JSD can be used in applied work, we use roll call votes from the 45-113th Congresses from 1878-2014.[2] For each bill, we code whether a yea vote is conservative or liberal using the following procedure: First, we take the median first dimension DW-NOMINATE score of the legislators who voted yea. Then we take the median first dimension DW-NOMINATE score for legislators who voted nay. If the median score of legislators who voted "yea" is greater than (i.e. more conservative than) the median ideology score of legislators who voted "nay," then a "yea"

[2]All data is obtained from http://voteview.com. Following Poole and Rosenthal (2007), we removed all consensus votes.

vote on the bill is classified as a conservative vote (1). Otherwise, it is coded liberal (0).

For each party, we then calculate the proportion of conservative votes cast in a given year, p_c, and use this as our estimate of the probability that a candidate from the party takes a conservative vote on any particular bill. We use $1 - p_c$ to estimate of the probability that party members take the liberal side of a vote. Using these estimated probability distributions, we calculate the party JSD according to equation 4 above. Figure 1 illustrates this procedure for the 101st congress in 1990.

Figure 1: **Party JSD Example:** Congress 101 (1990)



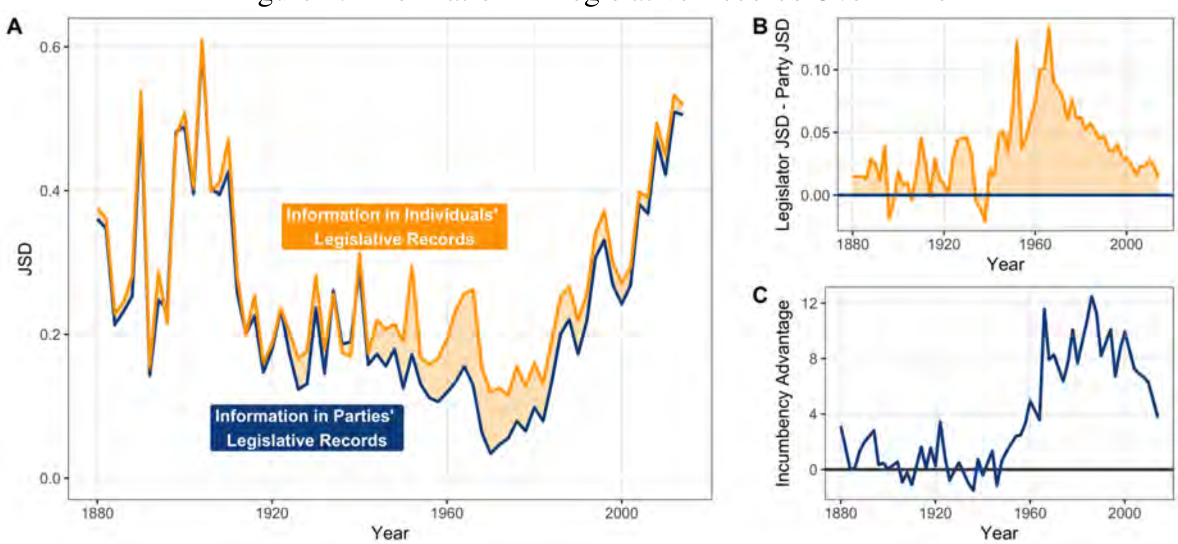
In addition to calculating a JSD for parties, we also calculate the JSD for legislators. We refer to this measure as the "legislator JSD". The procedure for measuring the legislator JSD is exactly the same as the procedure for measuring the party JSD, except we calculate the proportion of conservative votes cast by each individual legislator. We then calculate the JSD over each of the n legislators in a given congress using equation 3.

The results of this analysis can be seen below in Figure 2A, which shows the party JSD in blue and the individual legislator JSD in gold. Both the party and legislator JSD scores tend to travel in the same direction, and largely follow other measurements of polarization (Poole and Rosenthal 1997). This largely confirms the argument that, as parties have polarized, their legislative record has conveyed an increasing amount of information about the parties' ideological platforms (Grynaviski 2006, LeVeck and Kim 2013).

A new finding, however, is that after WWII there was an increasingly large gap between the information in legislators' and parties' records — with legislators' records conveying significantly more information about their their own ideology. This "legislator-party JSD gap" can be seen more easily in Figure 2B, which plots the difference between the legislator and party JSD over time. The gap between the party and legislator JSD widened from the late 40's until the mid 1960's, when it began to fall. Before 1948, the average difference between the legislator JSD and party JSD was 0.012. By 1966 it had reached more than 10 times that number, and was 0.134.

After 1966, the gap between the legislator and party JSD declined; and, by 2014 there was almost no difference between the two measures. Figure 2A shows that this narrowing gap between the legislator JSD and party JSD coincided with both the legislator and party JSD rising after 1966. However, the party JSD rose faster than the legislator JSD, eventually erasing any significant gap between the two measures (Figure 2B).

Figure 2: Information in Legislative Records Over Time



Discussion

Our findings above may help future research explain a number of important trends in congressional elections. Below, we note two of these trends and describe how they may be explained by temporal changes in the legislator-party JSD gap. However, we should be clear that our purpose is not to rigorously establish a causal relationship between the legislator-party JSD gap and specific trends in congressional elections. Such causal inference is well beyond the scope of this paper. Instead, we intend to illustrate how our measure could be useful to congressional election scholars.

Our hope then, is that future research will try to establish whether the relationships we describe below are, in fact, causal. For example, in a companion paper, we experimentally test whether increases in the party JSD over time cause voters to invest less effort in gathering information about individual incumbents (Nail 2018).

First, the diminishing gap between the legislator and party JSD may help explain the rise of party centered voting — where voters increasingly choose candidates on the basis of their party affiliation (Bartels 2000, Bonica and Cox forthcoming). It makes sense that voters would increasingly rely on the party labels of candidates, as these labels convey increasing information about a candidate's ideological voting, both in absolute and relative terms.

While previous work has recognized that increases in party-centered voting are likely related to party polarization (Kim and LeVeck 2013, Grynaviski 2006, Peskowitz 2017), we are the first to show that such polarization has coincided with a narrowing gap between the the ideological information in legislator and party records. This shrinking gap may have given voters an incentive to invest less effort in learning about their individual representatives in Congress — as learning about individual representatives' records is typically more costly than learning about the general activities of the parties (Popkin 1994, Brady and Sniderman 1985). If so, it would explain why voters know increasingly less information about their congressional representatives as individuals (Jacobson 2009), but know increasing amounts of information about what the parties in Congress stand for ideologically (Hetherington 2001, Smidt 2015).

Somewhat relatedly, our findings may help explain the rise and fall of the incumbency advantage, which was recently reported by Jacobson (2015). Panels B and C in Figure 2 show a temporal relationship between the legislator-party JSD gap (Figure 2B) and the rise and fall of the incumbency advantage (Figure 2C).[3] The rise and fall of the legislator-party JSD gap precedes the rise and fall of the incumbency advantage by about 10 years. This suggests that—if there is a causal relationship between the legislator-party JSD gap and the incumbency advantage—it is the legislator-party JSD gap driving the relationship.

[3] We use the well-known Gelman and King (1990) measure of the incumbency advantage.

Given what is known about the incumbency advantage, it makes some sense that its rise and fall would be related to the legislator-party JSD gap. One key advantage that incumbents have over non-incumbents is that they have a legislative record to run on (Jacobson 2009). However, this advantage would be diminished if people increasingly engaged in party-centered voting, and did not consider the legislative record of individual incumbents (Kim and LeVeck 2013, Bonica and Cox forthcoming, Jacobson 2015, LeVeck and Nail 2016, Peskowitz 2017). As we have argued above, shrinking the legislator-party JSD gap may lead to increases in party-centered voting. This, in turn, may shrink the incumbency advantage.

References

Ansolabehere, Stephen, and Philip Edward Jones. (2010).“Constituents' Responses to Congressional Roll-Call Voting." *American Journal of Political Science* 54(3): 583-597.

Ansolabehere, Stephen, James M. Snyder Jr, and Charles Stewart III. (2001). “Candidate positioning in US House elections." *American Journal of Political Science* 45(1): 136-159.

Bartels, Larry M. (2000). “Partisanship and Voting Behavior, 1952-2000.” *American Journal of Political Science* 44(1): 35-50.

Bonica, Adam, and Gary W. Cox. (2017). “Ideological Extremists in the US Congress: Out of Step but still in office." Forthcoming.

Brady, Henry E., and Paul M. Sniderman. (1985). “Attitude attribution: A group basis for political reasoning." *American Political Science Review* 79(4): 1061-1078.

Canes-Wrone, Brandice, David W. Brady, and John F. Cogan. (2002). “Out of step, out of office: Electoral accountability and House members' voting." *American Political Science Review* 96(1): 127-140.

Carroll, Royce, Jeff Lewis, James Lo, Nolan McCarty, Keith Poole, and Howard Rosenthal. “DW-NOMINATE Scores With Bootstrapped Standard Errors." DW-NOMINATE SCORES PAGE. N.p., n.d. Web. 15 Mar. 2016. <http://voteview.com/dwnomin.htm>.

Cox, Gary, and Mathew McCubbins. (2007). *Legislative Leviathan*. New York: Cambridge University Press.

Dancey, Logan, and Geoffrey Sheagley. (2013). “Heuristics behaving badly: party cues and voter knowledge." *American Journal of Political Science* 57(2): 312-325.

Gelman, Andrew, and Gary King. (1990). “Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34(4): 1142–1164.

Grynaviski, Jeffrey D. (2006). “A Bayesian learning model with applications to party identification." *Journal of Theoretical Politics* 18(3): 323-346.

Hetherington, Marc J. (2001). “Resurgent Mass Partisanship: The Role of Elite Polarization." *American Political Science Review* 95(3): 619–31.

Jacobson, Gary C. (2015). “It’s nothing personal: The decline of the incumbency advantage in US House elections." *Journal of Politics* 77(3): 861-873.

Kim, Henry A. and Brad L. LeVeck. (2013). “Money, Reputation, and Incumbency in U.S. House Elections, or Why Marginals Have Become More Expensive." 107(3): 492-504.

LeVeck, Brad L. and Stephanie A. Nail. (2016). “Evidence for a scale invariant relationship between the incumbency advantage and the nationalization of US House elections 1866-2014." *Research and Politics* 3(4): 1-4.

Lin, Jianhua. (1991). “Divergence measures based on the Shannon entropy." *IEEE Transactions on Information Theory* 37(1): 145-151.

Mayhew, David R. (1991). *Divided We Govern*. New Haven: Yale University Press.

Nail, Stephanie A. (2018). “How Much Will Voters Pay for a ‘Bit’ of Information?" Working paper.

Peskowitz, Zachary. (2017). “Ideological Signaling and Incumbency Advantage." *British Journal of Political Science*: 1-24.

Poole, Keith T, and Howard Rosenthal. (1997). *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.

Popkin, Samuel L. (1994). *The Reasoning Voter: Communication and Persuasion in Presidential Elections*. Chicago: University of Chicago: 49-71.

Shannon, Claude E. (1948). “A note on the concept of entropy." *Bell System Technology Journal* 27(3): 379-423.

Smidt, Corwin D. (2015). “Polarization and the Decline of the American Floating Voter." *American Journal of Political Science* 61(2): 365-381.

Snyder, James M., and Michael M. Ting. (2003). “Roll Calls, Party Labels, and Elections." *Political Analysis* 11(4): 419-444.

Woon, Jonathan, and Jeremy C Pope. (2008). “Made in Congress? Testing the Electoral Implications of Party Ideological Brand Names." *Journal of Politics* 70(3): 823-836.