

# How Much Will Voters Pay for a “Bit” of Information?\*

Stephanie A. Nail

Postdoctoral Research Fellow

Stanford University

September 7, 2020

## Abstract

Over the past four decades, technology has decreased the cost of acquiring information. Despite this pattern, there is not a clear indication that people have become more knowledgeable about their representatives in government on many dimensions. In fact, there is even some evidence that people are now less knowledgeable about some facts – such as the names of their members of Congress (Jacobson 2015). I argue that cheaper information has not increased voters’ knowledge about individual candidates because voters possess an even cheaper and increasingly informative cue: party id. As parties have become more ideologically distinct, voters have been increasingly able to guess how any given representative voted on a salient bill. Therefore, individuals should be less likely to seek out specific information about what individual legislators do in Congress. I test this hypothesis using a decision-theoretic experiment. In the experiment, participants try to guess how a candidate voted on a particular bill for a monetary reward, and may pay to acquire an informative signal before guessing. This is analogous to investing effort in learning facts about a candidate’s record, which is costly. I find that participants that have the party label available to them are indeed less willing to pay for an informative signal when it becomes easier to guess a candidate’s vote based on their party id. As shown in treatments mimicking the informational conditions in 1970 through 2008, individuals’ willingness to pay for more information when they have the party label has decreased by 30%.

**Keywords:** Information theory, decision-theoretic experiment, willingness to pay

---

\*Draft Job Market Paper. Please do not cite without author’s permission.

# 1 Introduction

Over the past four decades, access to information has become much cheaper. Through internet search engines such as Google, individuals can ask questions and receive simultaneous answers. Despite this pattern, however, there is not a clear indication that people have more *direct* knowledge about how their representatives have voted in Congress<sup>1</sup>. In fact, there is evidence that people are now less likely to know specific facts about their Congress members than in the past (Jacobson 2015). As just one example, voters are less likely to know their representative's name than they were in the past (Jacobson 2015).

In this paper, I argue that cheaper information has not increased voters' knowledge about individual candidates because voters possess an even cheaper and increasingly informative cue: party id. Party labels (or party cues) can be used as a heuristic or information shortcut in political decision-making. Increasingly complex political environments have encouraged the use of heuristics such as the party label in political decision making. Especially in recent years, individuals have turned to the party label as an easy way to make relatively well-informed decisions at little cost (Dancey and Sheagley 2013). Given that an individual has basic knowledge of what each party's platform consists of, the party label furnishes information about candidates or members of Congress that is easy to come by and does not incur a high cost to the individual.

We know from existing research that parties can use their activities in Congress to create information (Cox and McCubbins 2005). This information is then disseminated through the party label, which is used by individuals to learn about what the parties are doing in Congress and to make political decisions. Over the years, as the parties have become more and more ideologically distinct, individuals have been increasingly able to guess how any

---

<sup>1</sup>Direct knowledge is knowledge about an individual representative, such as their issue positions, voting history, or even their name. In contrast, *indirect* knowledge is knowledge inferred about individual representatives using other information – such as guessing how a representative voted based on their party label.

given representative voted on a salient bill. If this is the case, individual voters should be less likely to seek out specific information about what their legislators have done in Congress. Specifically, when the party label becomes more informative, making a well-informed decision becomes easier; therefore, individuals should be less willing to pay the cost of seeking out information on individual legislators' voting records.

This leads us to think that the information level of the party label changes the way that people perceive parties, and therefore, how they act and vote. However, it is increasingly difficult to know if changing the information level of the party label has a causal effect because we cannot go back through history and observational data is not sufficient. We can, however, use an experiment to mimic the different informational conditions individuals were exposed to with respect to the informativeness of the party label over the last few decades.

Building on this idea, I experimentally test the effect of varying the level of information conveyed by party brands. In particular, I test whether increasing levels of information conveyed by party brands has caused voters to invest less in learning about their representatives' individual records. I test this conjecture using a decision-theoretic experiment varying the amount of information that a candidate's party label conveys about how that candidate voted<sup>2</sup>. In the experiment, participants try to guess how a candidate voted on a particular bill for a monetary reward, and may pay to receive an informative signal before guessing. This payment is analogous to investing effort in learning facts about a candidate's record, which is costly. I hypothesize that participants with the party label available to them will be less willing to pay for an informative signal when it is easier to guess a candidate's vote based on their party id. As I use different informational conditions to mimic the information environment that individuals were exposed to in different years (1970, 1980, etc.), I expect

---

<sup>2</sup>This paper builds on a previous paper (as well as a poster that was presented at PolMeth 2017), and investigates the effect of varying the level of information conveyed by party brands. In "The Informational Value of Party Labels and Legislator Voting Records," we show that the information conveyed by a party's record has increased over time. In the current paper, I test whether this has caused voters to invest less in learning about their representatives' individual records.

that as the party label becomes more informative, individuals will be less willing to pay the cost of seeking out individual-level information on a candidate.

I find that, in treatments mimicking the informational conditions in 1970-2008, individuals' willingness to pay for more information when they have the party label has decreased by 30%. I find evidence that individuals have become more party-centric in their knowledge and evaluation of candidates and that a large part of this party-centric shift is driven by the information that is contained in party records. If parties continue voting in blocks and in ideologically homogeneous ways, the party label is increasingly effective. One drawback however, is that voters who use the party label in this way will hold their representatives accountable on the basis of their party, not their individual voting records. One important implication that can be drawn from the findings is that individuals are less willing to seek out more information when they have the party label and it is informative. This could minimize one key part of the incumbency advantage – having a record to run on. Furthermore, the results speak to issues of citizen competence and the use of heuristics. If individuals are incompetent, meaning that they are using heuristics because they are easier and more cost-effective, they should be immune to the varying informativeness of party cues. In contrast, if individuals are competent, they should use party labels as a heuristic more when it is more informative. By looking at the same heuristic, party labels, and varying how informative they are, I show that individuals are more likely to use party labels when they are informative.

## **2 Party Label Informativeness and Willingness to Pay**

Early work on voting behavior in the 1950s and 60s found that voters often knew very little about where candidates or their parties stood on a host of salient issues. Individuals were found to have neither a clear set of beliefs nor understand politics at an acceptable

level (Converse 1964; Campbell et al. 1960). This may be fundamentally detrimental for democratic accountability. If individuals are not aware of where parties stand on issues or how their representatives vote in Congress, they are not able to determine if their interests are being represented and are unable to hold their representatives accountable for their actions.

However, voters' knowledge about politics has changed substantially over the last 6 decades. For example, voters are now much more likely to know where the parties and their candidates stand on a host of issues (Hetherington 2001; Levendusky 2010). In fact, voters who pay very little attention to politics now know as much about the differences between the parties' positions as voters who paid a lot of attention to politics in the 1970s (Smidt 2017). Relatedly, work by Dancey and Sheagley (2013) finds that many voters can guess how their senator voted on a number of salient bills. Importantly, however, these changes in voter knowledge are largely being driven by what voters know about the parties, rather than an increase in voters' direct knowledge of candidates' own votes. For example, Dancey and Sheagley (2013) show that while citizens can often guess how their senator voted, they almost always get it wrong in the cases where their representative's vote diverges from the majority of the representative's party. Similarly, Warshaw and Tausonovitch (2018) show that voters are largely unable to distinguish how members of Congress from the same party differ ideologically.

Consistent with these findings, other scholars have found evidence that voters increasingly evaluate candidates based on their party affiliations, rather than their individual records (Rahn 1993; Lau and Redlawsk 2001; Popkin 1994; Snyder and Ting 2003; Bonica and Cox 2017; Kim and LeVeck 2013). Additionally, voters may actually know a smaller number of facts about their individual representatives than they did in the past. One example of this phenomenon is that voters are now less likely to know the names of their Congressional representatives (Jacobson 2015). This second finding is particularly interesting, given the

fact that the rise of and popularity of the internet has made it much easier for individuals to find this type of factual information – and at a fairly low cost compared to past decades. Similarly, it should be fairly easy to learn about how an individual’s own representatives voted on a specific bill, but the vast majority of voters do not seem to possess this type of direct knowledge about their incumbents, even for salient bills (Dancey and Sheagley 2013; Ansolabehere and Jones 2010). By increasingly evaluating candidates based on their party instead of their individual records, this may signal a shift in the level of accountability in American politics from individual representatives to parties as a whole.

However, it is still an open question as to why this shift to “party-centric” voting has occurred. One prominent explanation is that the difference in what individuals know and how they vote is actually being driven by changes in how the parties behave in Congress. Votes made by legislators in Congress have become increasingly polarized (Poole, Rosenthal, and Koford 1991; McCarty, Poole, and Rosenthal 2006; Poole and Rosenthal 1991; Poole and Rosenthal 1997; Poole and Rosenthal 2011). In addition, legislators have also increasingly taken more extreme party positions (Bafumi and Herron 2010; Fiorina, Abrams, and Pope 2006; Clinton 2006). Over time, as the parties in Congress have polarized, meaning that they have become more ideologically distinct, the party label itself has become a more informative cue to voters (Dancey and Sheagley 2013; Kim and LeVeck 2013; Grynaviski 2006; Smidt 2017). With a polarized Congress, there are now greater differences between how the members of each party vote, in addition to greater homogeneity within each party. This makes it such that if you know how a candidate’s party voted on an issue, you probably also have a very good idea about how that individual candidate voted as well. Therefore, because the party label of a candidate is relatively cheap to acquire and has become increasingly informative, voters may increasingly focus on a candidate’s party label instead of other information.

This explanation is appealing, both because it is intuitive, and because increases in

polarization strongly correlate with increases in what voters know about the parties (Smidt 2017). However, we cannot say for certain that this correlation means that the information contained in increasingly polarized party records is actually *causing* party-centric knowledge and voting among voters<sup>3</sup>. Also, even if the relationship is causal, we still lack evidence to show how strong the causal relationship is.

Evidence in support of this claim that more informative party labels have caused more party-centric voting has historically been difficult to obtain, mainly because we lack a control group containing voters who lack access to the increasingly informative party brands at multiple points in U.S. history. Without this comparison, and especially with observational data, it is impossible to determine causality, leaving us with only the ability to determine the possible correlation between party-centric voting and the increasing informativeness of party records.

Here, I address this difficulty by using a measure that quantifies the amount of information that is contained in party records about how individual candidates vote. This measure is abstract and can be applied to many environments, including decision-theoretic experiments, where the amount of information available to subjects equals the amount of information that is conveyed by party brands at different points in history. Using an experiment in which I vary the amount of information conveyed by the party label to match the information environment at various points in time, I formally test the following hypothesis:

**Hypothesis:** As party labels become more informative about how individual candidates will vote, citizens will be less willing to invest costly effort in learning about how the candidate actually voted.

---

<sup>3</sup>For example, one potential confounder is illustrated by McCarty, Poole, and Rosenthal (2006), who argued that polarization in Congress has been driven by fundamental demographic changes in the electorate, such as economic inequality and increased levels of immigration. These types of demographic changes could certainly also affect changes in voters' knowledge and behavior in addition to Congressional polarization.

### 3 Measurement

To test this hypothesis, we first need a measure that quantifies how much information is conveyed by the party label while simultaneously describing how distinct the party brands are from each other at a given point in time. Information is defined here as a reduction in uncertainty. This means that the more information that is available, the more uncertainty is reduced. Here, we are interested in reducing uncertainty about how incumbents vote in Congress. Therefore, the more information that is contained in the party label, the more uncertainty about how incumbents will vote is reduced.

This concept can be applied to how people vote ideologically by using the Jensen-Shannon Divergence (Lin 1991). The Jensen-Shannon Divergence (JSD) uses entropy ( $H$ ), which is a measure of uncertainty, to characterize how much uncertainty is reduced if we know which distribution is generating a given set of data<sup>4</sup>. In this case, the distributions are the parties in Congress (Republican, Democrat) and we can quantify the amount of uncertainty about ideological voting records that is reduced or eliminated when we go from not knowing an incumbent's party label to knowing their party label.

To do this, we first measure the uncertainty that we would have about whether an incumbent cast a conservative vote on any given bill if we did not know the incumbent's party. Then, we measure how much uncertainty we have about whether the incumbent cast a conservative bill if we did know what party they belong to. The difference between these two measurements represents how much uncertainty would be reduced on average by knowing the party that the incumbent belongs to. This difference is the JSD, and is illustrated in Equation 1 below.

---

<sup>4</sup>See appendix for a more detailed characterization of the Jensen-Shannon Divergence (JSD), as well as its mathematical details.



$$\begin{aligned}
& \text{Uncertainty given that a vote is cast by a member of either party} & \text{Average uncertainty given that a vote is cast by a member of a specific party} \\
& \overbrace{H\left(\frac{1}{2}Dem + \frac{1}{2}Rep\right)} & \overbrace{\left(\frac{1}{2}H(Dem) + \frac{1}{2}H(Rep)\right)} \\
JSD_{\frac{1}{2}}(Dem, Rep) = & & \\
\end{aligned} \tag{1}$$

To briefly illustrate how the JSD might be used in the case of party records and how the information in party records has varied over time, I use roll call votes from the 45-113th Congresses from 1878-2014 to calculate the JSD for each year<sup>5</sup>. 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<sup>6</sup>. Given these assumptions, we could use Equation 1 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 partisan lawmaking, Equation 1 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). 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. The party JSD is calculated over the parties' entire legislative record because it is a measure of how much information a party's records convey about individual candidates from that party. It is not a measure of how much information is consumed by any particular voter.

---

<sup>5</sup>All data is obtained from <http://voteview.com>. Following Poole and Rosenthal (2007), all consensus votes were removed.

<sup>6</sup>The JSD does not require that we restrict ourselves to a single dimension. This is just to simplify the example.

For a given year, the party JSD is calculated as follows. For each bill, I code whether a yea vote is conservative or liberal using the following procedure: First, I take the median first dimension DW-NOMINATE score of the legislators who voted yea. Then, I 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” vote on the bill is classified as a conservative vote (1). Otherwise, it is coded liberal (0). For each party, I then calculate the proportion of conservative votes cast in a given year,  $pr(con)$ , and use this as the estimate of the probability that a candidate from the party takes a conservative vote on any particular bill. I use  $1 - pr(con)$  or  $pr(lib)$  to estimate the probability that party members take the liberal side of a vote. Using these estimated probability distributions, party JSD is calculated according to Equation 1 above. The JSD for each year from 1878-2014 as seen in Figure 1 below exhibits considerable variation over time.

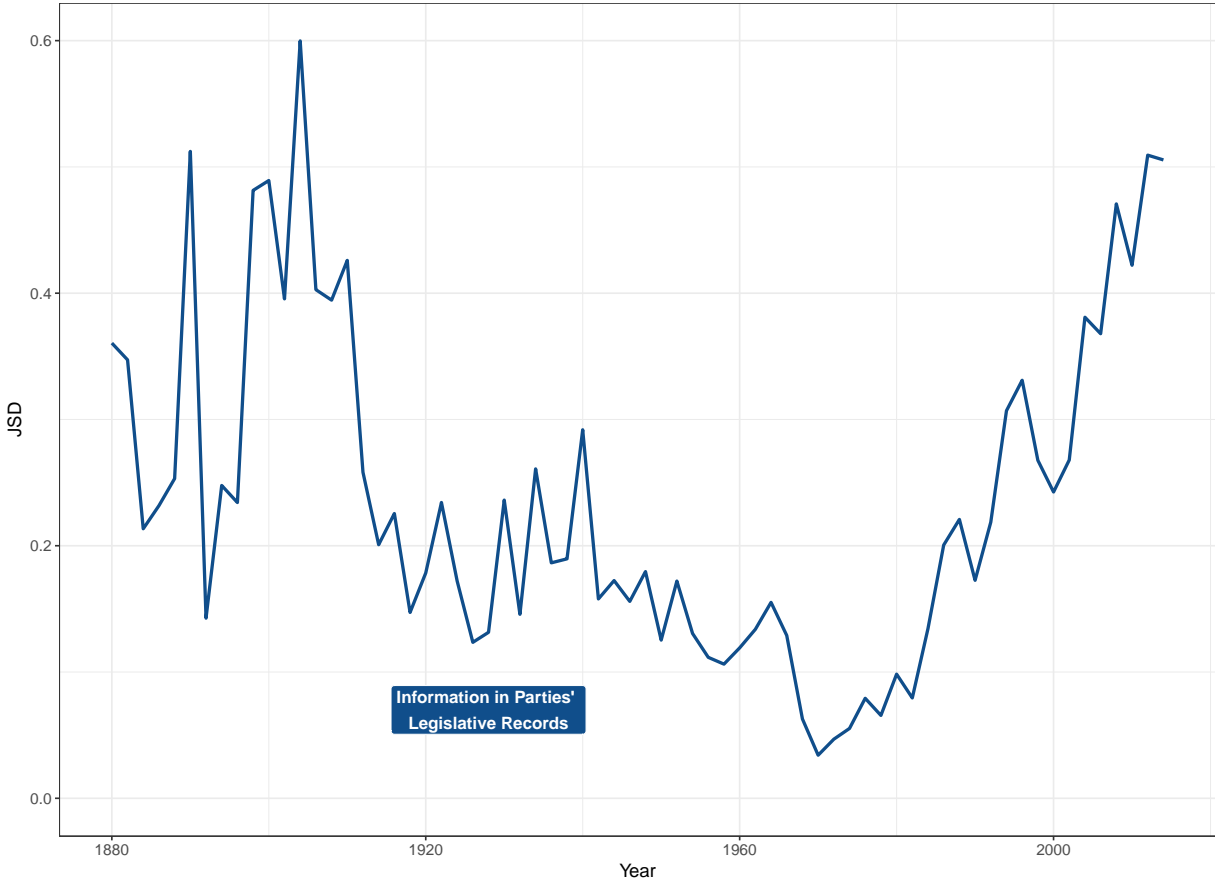
Theoretically, our measure of information, the JSD, can vary between 0 and 1. A JSD of 1 means that the parties are perfectly distinguishable from one another (most informative). In contrast, a JSD of 0 means that in terms of ideological voting records, the parties are not distinguishable at all.

Figure 2 illustrates how the JSD can be applied to how people voted (yea / nay) ideologically for the 101st Congress in 1990. Figure 2 tells us that for all votes made by legislators in 1990, the probability that a given vote by a legislator was conservative was 0.41, while the probability of a liberal vote was 0.59. If we are trying to guess how a legislator voted (without knowing which party they are from), this distribution of votes leaves us highly uncertain about the correct answer<sup>7</sup>. If we break this mixture of all votes down by party, we see that for Republican legislators (red box), the probability of a given legislator making a

---

<sup>7</sup>The entropy (H) for this mixture of votes represented by the purple boxes is very high (0.98 out of 1.00) meaning that we are very uncertain.

Figure 1: **Party JSD**: Variation in Party JSD in the last 130 years



conservative vote was 0.65 and the probability of making a liberal vote was 0.35. Similarly, for Democratic legislators (blue box), the probability of making a conservative vote was 0.18, while the probability of making a liberal vote was 0.82. In addition, the uncertainty that we have regarding the voting behaviors of Democrats is much smaller ( $H = 0.67$ ) than the uncertainty that we have regarding Republican legislators ( $H = 0.93$ ). Overall, the JSD for 1990 was 0.18, meaning that at this time in history, the party label contained a relatively small amount of information in terms of ideological voting records.

The party label JSD varies over bills and years, and in recent years, has become increasingly more informative than in the past. As seen in Table 1 and Figure 1, there is considerable variation in the party label JSD since 1970 and more generally, over the last

Figure 2: Party JSD Example: 101st Congress (1990)

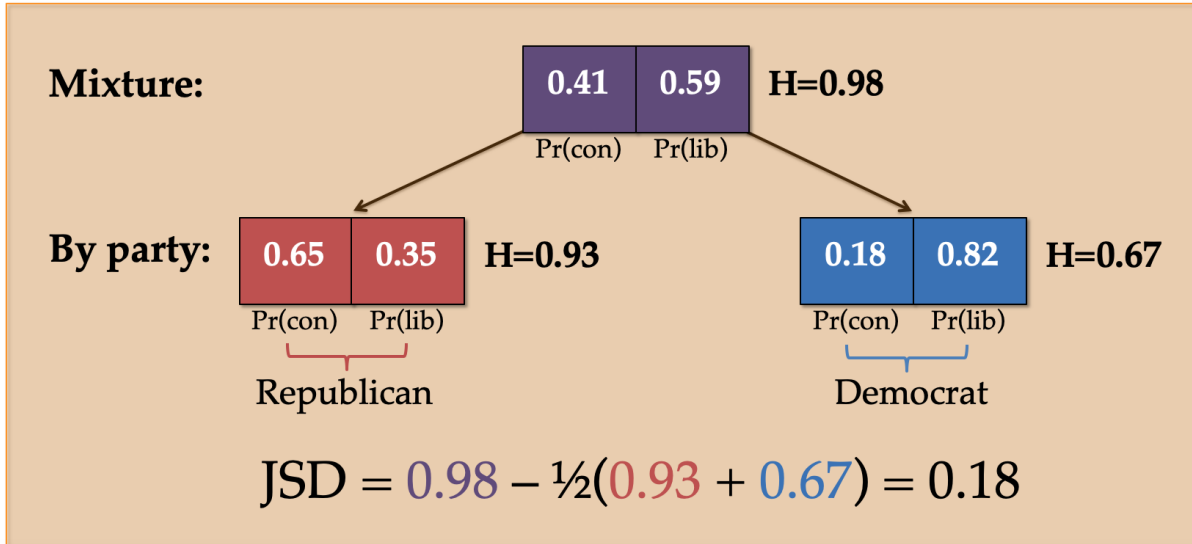


Table 1: Party Label JSD (1970-2010)

Year	Party Label JSD
1970	0.03
1980	0.10
1990	0.17
2000	0.24
2010	0.42

130 years. Because the party label JSD tells us exactly how much information the party label conveyed in a given year, we can use this measurement to recreate the information environment that an individual was exposed to during that time. To illustrate, recall that in the 101st Congress example, the mixed (purple) distribution of liberal / conservative votes included all votes taken during 1990. The JSD in 1990 thus gives us the amount of information that was conveyed by the party label for this year. To mimic the information environment of 1990, we can find a single bill where the distribution of yea / nay votes within each party matches the distribution of conservative / liberal votes in a given year.

For example, in 2004, the party label JSD was 0.43. We can mimic this level of information with a single bill that has the following characteristics:

**Overall, on this bill,**

213 legislators voted in favor of the bill and 214 legislators voted against the bill.

**In terms of the distribution of yea / nay votes by party,**

Democrats who voted in favor of the bill: 199/250

Republicans who voted in favor of the bill: 14/177

We can see here that knowing the yea / nay distribution on the bill overall is not helpful as it is 50 / 50. However, we know that the party label is highly informative (0.43) and this is demonstrated by the breakdown of yea / nay votes by party. When thinking about how a Democratic legislator voted on this bill, it is overwhelmingly likely that they voted in favor of the bill. Instead, if asked to guess about how a Republican legislator voted on this bill, it is highly likely that they voted against the bill. Here, knowing the party that a legislator is from greatly increases the probability that we would correctly guess how they voted on this bill.

To test the theory and the hypothesis that as party labels become more informative about how individual candidates will vote, citizens will be less willing to invest costly effort in learning about how the candidate actually voted, I will construct a set of bills where the distribution of yea / nay votes within each party matches with the distribution of liberal / conservative votes in a year. Here, each bill will match a different party label JSD for a given year, allowing for a comparison of information-seeking behavior over time.

The goal is to use this set of bills that represents the informativeness of the party label over time to see whether more informative yea / nay vote distributions (higher JSD and more informative party labels) cause individuals to be less willing to pay for information about whether a legislator voted yea or nay on the bill. This is analogous to how more informative distributions of conservative / liberal votes may have made people less willing to invest costly

effort into finding out how their individual representative actually voted in Congress. In the following section, I detail an experiment using a set of bills matching different party label JSDs to test this idea.

## 4 Decision-Theoretic Experiment

This paper seeks to examine the effect of the increasing information level of party labels on willingness to obtain additional information using actual information levels over the past five decades. To test the hypothesis, I employ an experimental research design. The design is based on the idea that there is a certain level of information that is contained in the party label and individual legislator voting records.

When the party label becomes more informative (JSD goes to 1), it is expected that individuals will place more weight on the party label and less weight on individual legislators' voting records. This is because a JSD of 1 indicates that knowing the party label allows voters to perfectly predict how an individual representative votes. Thus, as the party label increases in informativeness, respondents given the party label of a legislator should be less willing to pay for information on individual legislators, even when the goal is correctly identifying an individual legislator's vote. In this way, the party label acts as an information subsidy – even though it is cheaper to get candidate information nowadays, the high information level of the party label makes acquiring individual legislator information uneconomical.

To explain how individuals can gauge the informativeness of the party label using the bills chosen to mimic the information environment in a given year, imagine that there are 50 Democrats and 50 Republicans for a total of 100 members of Congress. For the first bill in Table 2, exactly half each party votes for the bill. This means that 25 of the Democratic members and 25 of the Republican members vote yea. Similarly, 25 Democratic members and 25 Republican members vote against the bill for a total of 50 yeas and 50 nays. Because

Table 2: Close Vote Example

Bill	Democrat Yea	Democrat Nay	Republican Yea	Republican Nay	Total Yea	Total Nay	JSD
1	25	25	25	25	50	50	0
2	50	0	0	50	50	50	1

this distribution gives us no information, the JSD is 0 – the parties are not distinguishable. In contrast, for the second bill, imagine that all 50 of the Democratic members vote yea and all 50 of the Republican members vote nay. Although the total is still 50 yeas and 50 nays, because the parties are completely distinguishable, Bill 2 has a JSD of 1.

In the no party label condition of the experiment, Bill 1 and Bill 2 have the same amount of information (50 yeas and 50 nays). If the task is to guess if a given member of Congress voted yea or nay, you have a 50/50 chance of being right for both Bill 1 and Bill 2, since you do not know the party label of the member. Therefore, individuals should pay the same amount for additional information about how the individual legislator in question voted in both of these scenarios (Bill 1 and Bill 2).

In contrast, in the party label condition, despite the fact that each bill has 50 yeas and 50 nays, the JSD of the two bills are very different. Since Bill 1 has a JSD of 0, individuals should pay for additional information on how the member in question voted. This is because even though they have the party label to use, it does not provide them with any additional information to use. However, on Bill 2, since the party label of the member is known and the JSD is 1, individuals should never pay for additional information because knowing the party label is all the information that is needed to make a correct prediction. Following this logic, as the JSD increases from 0 to 1, individuals should be less willing to pay for additional information in the party label condition as the party label increases in informativeness.

The basic experimental set-up is shown in Figure 3. Participants are told that they will receive a bonus (\$1) if they correctly guess how the legislator voted on the given bill. In the first part of the experiment, participants are presented with information on the distribution

of legislator votes for a specific bill that received a roll call vote (Stage 1). Bills used in the experiment were selected to match the overall JSD for the House of Representatives by year. The distribution of roll call votes on a bill within each party matched the distribution of conservative and liberal votes cast by each party in a given year. Therefore, the party label of a candidate conveys the same amount of information (about how the candidate voted on the bill) as the candidate’s party label would have conveyed about their propensity to vote conservatively in a given year. This means that the bill-level JSD for a given bill is identical to a yearly JSD between the years of 1970-2010 (Table 1). Furthermore, despite varying the underlying bill-level JSD, only close votes (50% yea, 50% nay) are used in this experiment. This allows each bill to have the same baseline level of information, absent a party label (Table 2).

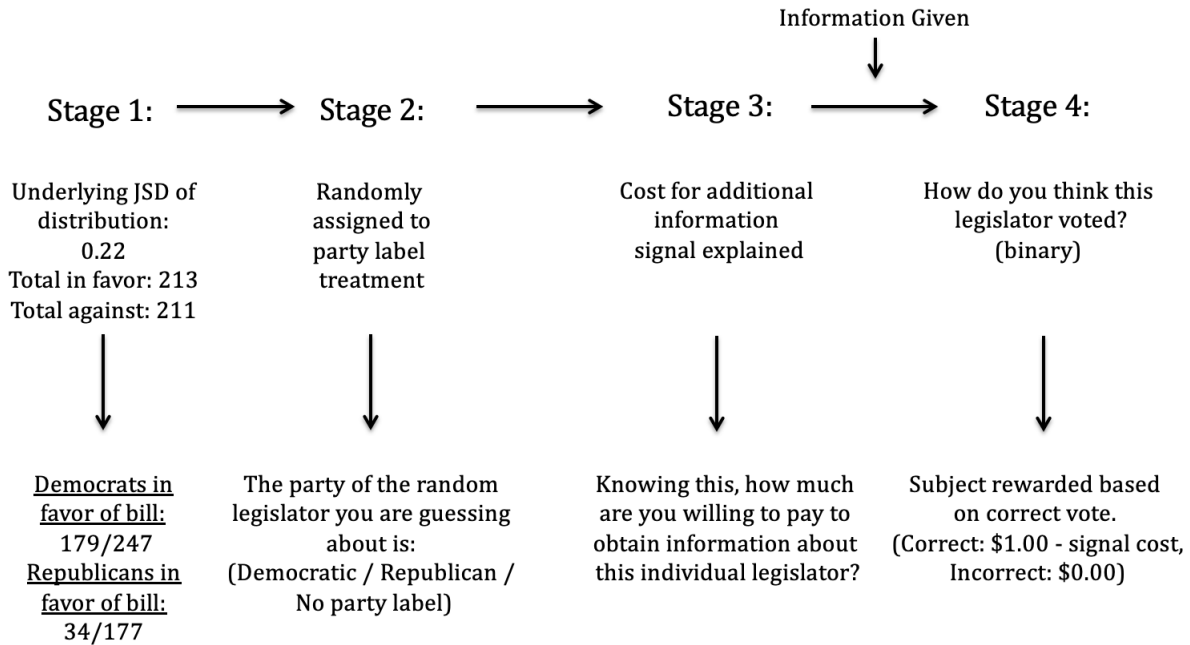
Participants are randomly assigned to receive one of two informational treatments: party label (Democrat or Republican) or no party label (Stage 2). Participants face a cost to acquire more information (Stage 3). If they choose to pay nothing, they will get a signal that is essentially a coin flip with a 50/50 chance of accuracy. On the other hand, if the participant chooses to pay \$0.02 or \$0.50, they will get a signal that is 52% or 100% accurate, respectively. Lastly, after receiving the signal according to their willingness to pay in Stage 3, the participant is asked to give their best guess of how the legislator in question voted on the bill (Stage 4). Participants are paid \$1.00 - signal cost for a correct guess and \$0 otherwise for one randomly selected trial.

This creates two distinct cases (Table 3). Participants complete a total of seven trials with varying bill-level JSDs under the same set of randomly assigned conditions, making this a between-subjects design.

Figure 4 is what participants in the experiment see. The top panel is a participant in the control condition which has “not available” as the party label. The bottom panel is a participant in the treatment condition who has “Republican” on the current trial, but



Figure 3: **Trial Set-up:** Example trial using 1980 JSD



randomly receives either “Democrat” or “Republican” for each trial. What is important to note here is that, for participants in the control condition (top), the only piece of information they can use is the 50/50 yea / nay distribution. They are not able to make use of the second piece of information – the distribution of yea / nay votes by party, because they do not have access to the party label of the legislator. We would therefore expect these participants<sup>8</sup> to pay some sum of money greater than zero for more information because they always have a 50/50 chance of answering correctly without more information. In contrast, for participants in the treatment condition (bottom), these individuals can use the distribution of yea / nay votes within each party. Therefore, we would expect treatment condition participants to be willing to pay less than individuals in the control condition because the distribution of votes within each party provides them with some information about how the candidate voted. In the particular case shown in Figure 4, subjects would know that the candidate (a Republican) was more likely than not to vote against the bill.

<sup>8</sup>Given that they have a normal range of risk aversion.

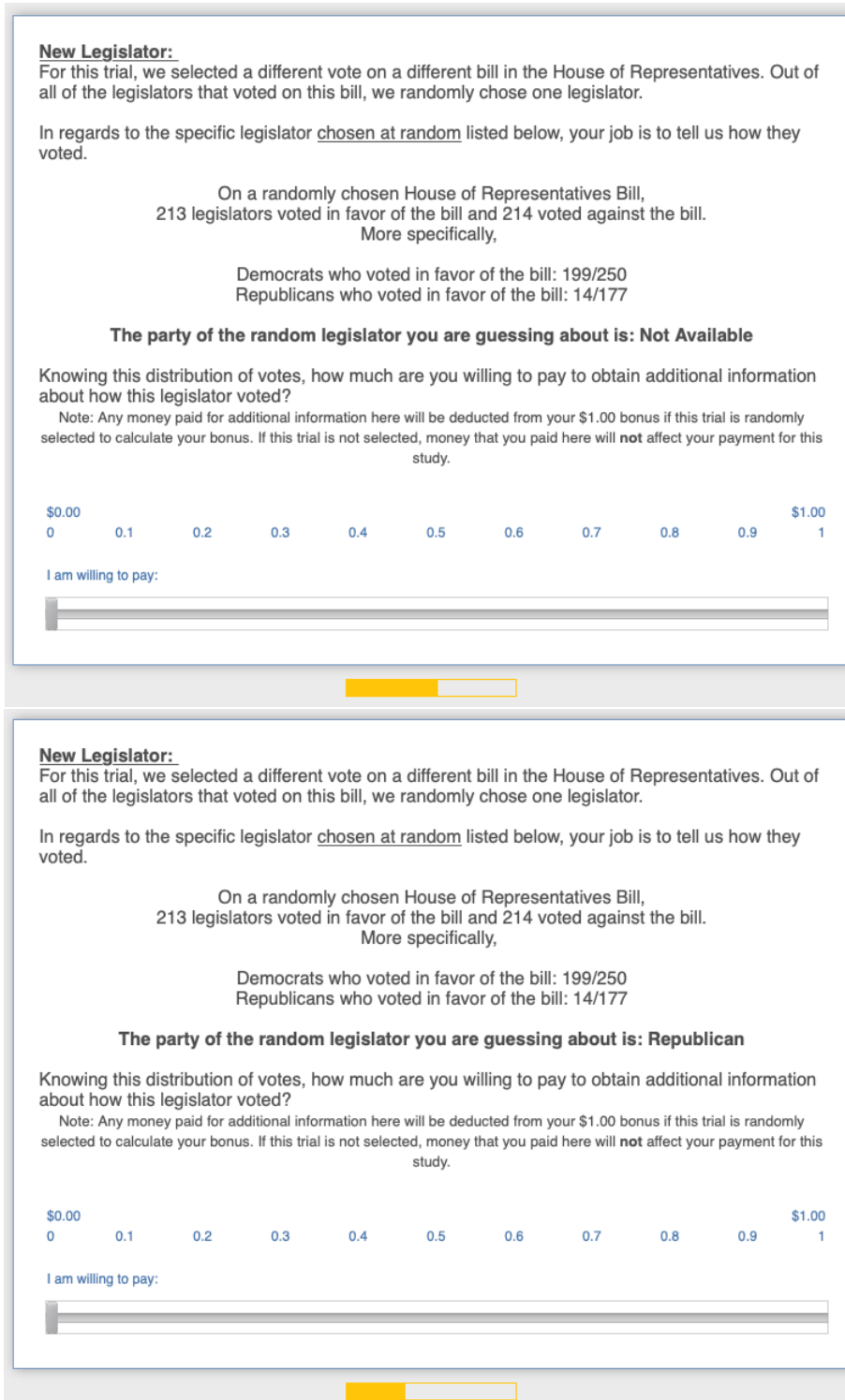


Figure 4: **Top:** A participant in the control condition. **Bottom:** A participant in the treatment condition.

Table 3: **Experimental Design:** informational level and conditions  
**Varying JSD by Year**

<b>Party Label</b>	$Bill_{1JSD}, Bill_{2JSD}, \dots, Bill_{7JSD}$
<b>No Party Label</b>	$Bill_{1JSD}, Bill_{2JSD}, \dots, Bill_{7JSD}$

Despite the final stage being the participant’s guess regarding the legislator’s vote, the dependent variable of interest in this experiment is willingness to pay for information. The hypothesis posits that as the party label becomes more informative, individuals will be less likely to invest costly effort into learning about an individual candidate’s voting record. Here, willingness to invest costly effort is modeled by monetary costs – to get more information about how the individual legislator that is being asked about voted, individuals must pay in cents. This is meant to capture the idea that in the real world, individuals must invest costly effort such as time to learn more about individual legislators’ voting records. It is acceptable to use willingness to pay for a hint in the experiment as a proxy for costly effort even though willingness to pay is not elicited in the real world. This is because, despite probably not having to pay an outright monetary cost to obtain more information in the real world, individuals incur other costs such as time, effort, and opportunity costs.

For each trial, participants are asked how much they would be willing to pay to get additional information to help them make their decision. This monetary value becomes the dependent variable. With multiple trials, the objective is to compare willingness to pay for an additional signal on bills with different party JSDs (over time).

In the no party label condition, I expect there to be a similar willingness to pay across JSDs because the vote distribution for close votes (50% yea / 50% nay) is not informative and the participant is unsure of the party label of the legislator in each trial. Therefore, they should be willing to pay a relatively similar amount for additional information across all JSD levels since they are not given any party labels to make use of.

In contrast, in the party label condition, I expect that as the informativeness of the

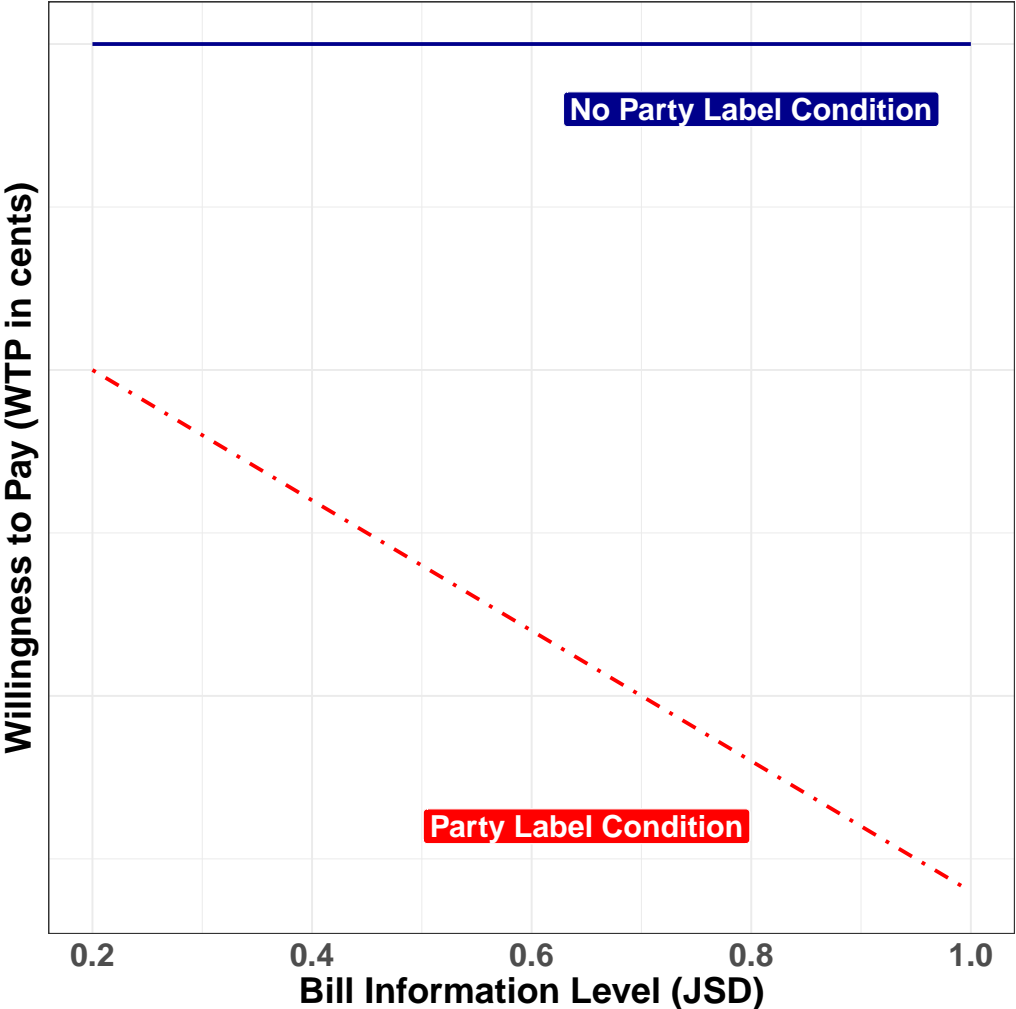
party label increases (JSD increases), individuals' willingness to pay for individual legislator information will decrease. As the distribution of votes between Democrats and Republicans becomes increasingly different, individuals who are given the party label will be able to make better predictions. Therefore, they will be less willing to pay for additional information in this case. As the distributions become more similar (JSD decreases), even individuals with the party label will have an increasingly hard time deriving information from it. Therefore, they will be more willing to pay for additional information.

With regard to willingness to pay for information, I compare the individuals who get the party label versus those that do not. Our theoretical expectations are presented graphically in Figure 5. Within each of the conditions, I vary the potential informativeness of the party label (JSD) while keeping the overall number of yeas and nays on each bill relatively equal. Individuals who receive the party label treatment, should be less willing to pay for additional information. The cost of obtaining additional information matters less for individuals with the party label given as we move from low informativeness of the party label to high informativeness (JSD). Therefore, the difference between willingness to pay in the party label condition (red) and willingness to pay in the no party label condition (blue) should become larger as the party JSD increases.

In the no party label condition, individuals' willingness to pay should be relatively the same across all of the different JSD levels. This is because they are given no party label and the yea / nay distribution of a close vote does not give them any additional information or the chance to increase their probability of making a correct guess without paying for additional information on the legislator's individual voting record.

While this experiment is quite abstract, this level of abstractness helps us to investigate the mechanism by which individuals have come to invest less in learning about individual candidates' voting records. By holding all else constant but the informativeness of the party label (JSD), this allows us to determine if increasingly informative party labels have

Figure 5: Pre-Registered Hypothesis Expected Results: Relationship between WTP and JSD



*caused* individuals to be less willing to invest effort into learning about individual candidate information.

## 5 Data and Methods

This experiment was pre-registered with the Open Science Foundation<sup>9</sup> and data was collected using Qualtrics and Amazon’s Mechanical Turk<sup>10</sup>. Participants were screened to be at least 18 years of age and reside in the United States. The 1199 participants were randomly assigned to one of two experimental conditions: party label or no party label. Each participant completed multiple trials under the same condition (i.e. party label or no party label).

Figure 6 shows the results from the experiment. Because the seven trials are equally spaced in terms of JSD jumps from one trial to another, there is no year associated with the last trial. For ease of discussion, since this trial’s JSD is closest to 2008, I will use 2008 as its year. The hypothesis is supported, as individuals in the party label condition are indeed less willing to pay for an additional piece of information, compared to individuals in the no party label condition, over time ( $p - value = 0.0004$ ). Additionally, as the party JSD increases, the difference in willingness to pay between the party label and the no party label conditions generally increases. From 1970 to 2008, willingness to pay for information has dropped by about 30%.

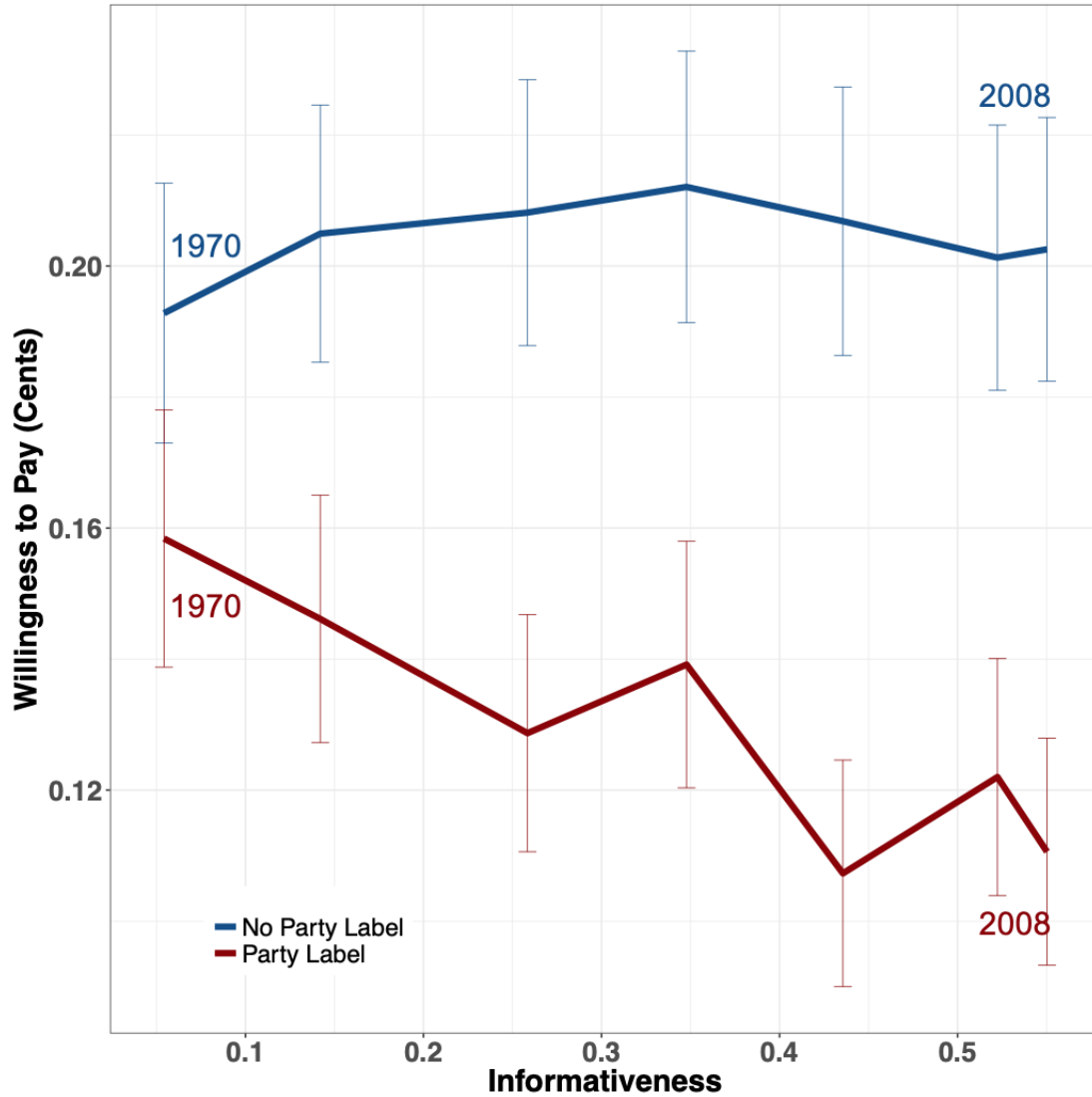
We would also expect that there is not a significant difference between the WTP of individuals in the no party label condition between 1970 and 2008. This is indeed the case, as the average willingness to pay for respondents is not significantly different between the two years ( $p - value = 0.4991$ ). Subsequently, we should expect a difference in WTP between 1970 and 2008 for participants in the party label condition. This is also the case. Participants in the party label condition are willing to pay significantly less for additional information in 2008 (when the JSD is over 4 times as high) than in 1970 ( $p - value = 0.0004$ ). Finally, there

---

<sup>9</sup><https://osf.io/dg43e/>

<sup>10</sup>This experiment was reviewed and approved by the University of California, Merced IRB (UCM2017-151).

Figure 6: **Results:** Relationship between WTP and JSD



is a significant difference between the amount that participants in the no party label versus participants in the party label condition are willing to pay in 2008 ( $p - value > 0.0000$ ).

## 6 Discussion and Conclusion

Although we have seen evidence that people have become more party-centric in their knowledge and evaluation of candidates, it has been an open question as to how much of

this shift was driven by the way that the parties act in Congress. We have seen here that a large part of the party-centric shift is driven by the information that is contained in party voting records and ultimately the party label.

Since the 1970s, both the informational level of the party label and the informational level of individual members' legislative records have increased. However, the informational level of the party label has increased at a higher rate. This may be why there has been an increase in party line voting because knowing the party label is easier and just as effective as knowing a legislator's entire voting record. This experiment investigates the effect of varying the level of information that is contained in the party label. More specifically, it tests whether increasing levels of information conveyed by the party label has caused voters to invest less in learning about their representatives' individual voting records. The results suggest that this may indeed be the case, as willingness to pay for information when the party label is known has decreased by about 30% over time.

This result is important because it might signal a change in the level of representativeness in the United States. There is evidence that people have become more party-centric regarding what they know about and how they evaluate candidates. We now know that a large part of this change in what voters are willing to learn about their representatives is driven by the increase of information contained in the party label. If parties continue to vote in blocks and in more ideologically homogeneous ways, using the party label is increasingly effective. However, this means that voters are holding representatives accountable on the basis of the party, not on their individual voting record in Congress. If voters are not aware of the voting behavior of their representatives in Congress and how it differs from their party's voting record, this weakens the incentive of members of Congress to vote according to their district's preferences. Finally, this type of behavior could decrease the incumbency advantage. One of the many benefits to being an incumbent is having an individual voting record to run on. If voters are not paying attention to individual records, however, this could



decrease this aspect of the incumbency advantage.

The results are also relevant to the debate on citizen competence and the use of heuristics. If individuals are competent in terms of seeking out information that will allow them to make a good decision, we would expect that they would use the party label more as it becomes more informative. On the other hand, incompetent individuals should be immune to the varying level of informativeness of the party label. By varying the informativeness of the party label, I find that individuals are more likely to use the party label when it is informative and more willing to pay for individual-level candidate voting information when the party label is less informative, suggesting that they exhibit competence.

While previous experiments have not controlled the benefit of voting correctly or the cost of acquiring more information, the design of the current experiment allows for control of both of these factors. In addition, this experiment varies the informational value of the party label to mimic real-world changes over the last five decades. This is done by first computing how much information is contained in the party label in a given year using Congressional roll-call votes. The resulting measure, the JSD, tells how distinct or distinguishable the votes of each party's members are along the left-right ideological spectrum (in terms of the propensity to make conservative or liberal votes). This measure is computed for every year and shows considerable variation throughout the years (Figure 1).

The JSD can be computed for years or bills, and in this case, each yearly JSD (or how informative the party label is) is matched to an individual bill-level JSD. This matching means that the party label of a candidate conveys the same amount of information (about how the candidate voted on the bill) as the candidate's party label would have conveyed about their propensity to vote conservatively in a given year. By matching a bill-level JSD with a yearly JSD, this experiment is able to vary the informational value of the party label over time, using actual Congressional bills for individual trials, without participants realizing what bill or year they are being asked about. This study compares participants'

willingness to pay for individual-level information over a variety of party label information levels (yearly JSDs) by using multiple trials (one year informational condition per trial). With this setup, we can use experimental manipulation to make meaningful inferences about changes in individuals' willingness to seek out individual candidate-level information when the party label was more or less informative at different points in history.

## 7 References

- Ansolabehere, Stephen, and Philip Edward Jones. (2010). "Constituents' Responses to Congressional Roll-Call Voting." *American Journal of Political Science* 54(3): 583-597.
- Bafumi, Joseph, and Michael C. Herron. (2010). "Leapfrog Representation and Extremism: A Study of American Voters and Their Members in Congress." *American Political Science Review* 104(3): 519-42.
- Bonica, Adam, and Gary W. Cox. (2018). "Ideological Extremists in the US Congress: Out of Step but Still in Office." *Quarterly Journal of Political Science* 13(2): 207-236.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. (1960). *The American Voter*. New York John Wiley and Sons.
- 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>>.
- Clinton, Joshua D. (2006). "Representation in Congress: Constituents and Roll Calls in the 106th House." *Journal of Politics* 68(2): 397-409.
- Converse, Philip E. (1964). "The Nature of Belief Systems in Mass Publics." *In Ideology and Discontent*, edited by David E. Apter, 206-61. New York: Free Press.
- Cox, Gary, and Mathew McCubbins. (2007). *Legislative Leviathan*. New York: Cambridge University Press.
- Cox, Gary, and Mathew McCubbins. (2005). *Setting the Agenda*. 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.
- Fiorina, Morris P, Samuel J. Abrams, and Jeremy C Pope. (2006). *Culture War? The Myth*

of a Polarized America. Vol. 2nd.

- 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." *The Journal of Politics* 77(3): 861-873.
- Jacobson, Gary C., and Jamie L. Carson. (2015). *The Politics of Congressional Elections*. Rowman & Littlefield.
- Kim, Henry A., and Brad L. LeVeck. (2013). "Money, Reputation, and Incumbency in US House Elections, or Why Marginals Have Become More Expensive." *American Political Science Review* 107(3): 492-504.
- Lau, Richard R., and David P. Redlawsk. (2001). "Advantages and disadvantages of cognitive heuristics in political decision making." *American Journal of Political Science* 45(4): 951-971.
- Levendusky, Matthew S. (2010). "Clearer Cues, More Consistent Voters: A Benefit of Elite Polarization." *Political Behavior* 32 (1): 111-31.
- Lin, Jianhua. (1991). "Divergence measures based on the Shannon entropy." *IEEE Transactions on Information Theory* 37(1): 145-151.
- McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge: MIT Press.
- Poole, Keith T., and Howard Rosenthal. (1991). "Patterns of Congressional Voting." *American Journal of Political Science* 35(1): 228-78.
- Poole, Keith T, and Howard Rosenthal. (1997). *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Poole, Keith T, and Howard Rosenthal. (2011). *Ideology and Congress*. Transaction Pub-

lishers.

- Poole, Keith T., Howard Rosenthal, and Kenneth J. Koford. (1991). "On Dimensionalizing Roll Call Votes in the U.S. Congress." *The American Political Science Review* 85(3): 955-76.
- Popkin, Samuel L. (1994). *The Reasoning Voter: Communication and Persuasion in Presidential Elections*. Chicago: University of Chicago: 49-71.
- Rahn, Wendy M. (1993). "The Role of Partisan Stereotypes in Information Processing about Political Candidates." *American Journal of Political Science* 37 (2): 472-96.
- Smidt, Corwin D. (2017). "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.
- Tausanovitch, Chris, and Christopher Warshaw. (2018). "Does the Ideological Proximity Between Candidates and Voters Affect Voting in US House Elections?" *Political Behavior* 40(1): 223-245.
- 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.

# Appendix

## Method of Measurement

Here, I 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:

$$H(X) = - \sum_{i=1}^n P(x_i) \log_2 P(x_i) \quad (2)$$

$H(X)$  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. This measure of uncertainty is maximized when there is an equal probability of each value  $x_i$ . 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

$$I(X, Y) = H(X) - H(X|Y) \tag{3}$$

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 about  $X$ ), 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.

## 7.1 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 the observer knows exactly which distribution is generating a given set of data. The JSD is the mutual information between the labels and the aggregate data (Lin 1991).

$$\begin{aligned}
 JSD_{\pi_1, \dots, \pi_n}(P_1, \dots, P_n) = & \underbrace{H\left(\sum_{i=1}^n \pi_i P_i\right)}_{\substack{\text{Uncertainty over a mixture} \\ \text{of } n \text{ unlabeled distributions}}} - \underbrace{\sum_{i=1}^n \pi_i H(P_i)}_{\substack{\text{Average uncertainty} \\ \text{of } n \text{ labeled distributions}}} \quad (4)
 \end{aligned}$$

In Equation 4 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 can 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 2 in the Appendix), the JSD is bounded between 0 and 1.