

# Updating A Time-Series of Survey Questions: The Case of Abortion Attitudes in the General Social Survey

Sarah K. Cowan\*, Michael Hout and Stuart Perrett

July 2021

Forthcoming at *Sociological Methods & Research*

\*Corresponding Author

sarahkcowan@nyu.edu  
Department of Sociology  
New York University  
295 Lafayette Street  
New York  
NY 10012

GSS Methodological Report No. 133  
NORC at the University of Chicago  
Chicago, IL

# Updating A Time-Series of Survey Questions: The Case of Abortion Attitudes in the General Social Survey

## **Abstract**

Long-running surveys need a systematic way to reflect social change and to keep items relevant to respondents, especially when they ask about controversial subjects, or they threaten the items' validity. We propose a protocol for updating measures that preserves content and construct validity. First, substantive experts articulate the current and anticipated future terms of debate. Then survey experts use this substantive input and their knowledge of existing measures to develop and pilot a large battery of new items. Third, researchers analyze the pilot data to select items for the survey of record. Finally, the items appear on the survey-of-record, available to the whole user community. Surveys-of-record have procedures for changing content that determine if the new items appear just once or become part of the core. We provide the example of developing new abortion attitude measures in the General Social Survey. Current questions ask whether abortion should be legal under varying circumstances. The new abortion items ask about morality, access, state policy, and interpersonal dynamics. They improve content and construct validity and add new insights into Americans' abortion attitudes.

“The way to measure change is not to change the measure” (Smith, 2005).<sup>1</sup> This first law of studying societal change undergirds much of the content in repeated cross-sectional surveys; researchers ask the same questions the same way over time to quantify change in public opinion and behavior. Given enough time, though, and people might come to understand the same words differently. A key word or phrase in a question might go out of everyday use, or the terms of a question might be too narrow to encompass the issues that come to respondents’ minds when they think about the topic a question raises. The most commonly used surveys — the General Social Survey (GSS) and the American National Election Studies (ANES) — have such long histories that some items are likely afflicted by one or both of these problems. Both studies have changed measures from time to time throughout their histories, but not systematically using a set procedure. A standard protocol for updating questions and scales might prove useful.

Tom W. Smith, the long-serving Principal Investigator of the GSS articulated a second law: “When constant measures produce nonconstant measurement, change the measure to measure change” (Smith, 2005). As survey experts advise, scholars should constantly be monitoring measures to ensure their continued fitness (Turner and Martin, 1985). Indeed, the GSS has periodically changed its core (Marsden et al., 2020), and the ANES has a user-sourced process of introducing new measures.<sup>2</sup> The survey methodology literature, exemplified by Smith (2005), recognizes many reasons for changing measures. Smith (2005) presented several compelling examples ranging from the simple task of adjusting income measures for inflation to the complex task of measuring racial attitudes among the many groups contributing to US diversity. At times the change is to re-introduce old measures that are newly relevant, such as in 2008 items asking whether a respondent would vote for a female or black Presidential candidate reappeared in the GSS. At times the changes might be to add new items to a familiar scale, as the GSS is doing with its vocabulary quiz (“wordsum”). These changes are motivated by changing language or changing demographics.

We examine a social change of a different sort: public opinion on abortion in the United States. The current questions ask if abortion should be legal for pregnant women who face a variety of circumstances. They scale well together and were ideal when the terms of debate concerned birth defects and poor women (Hout 1999). But discourse and policy moved on. Morality was always part of the conversation but not part of the scale; that is still true. Abortion access, gestational age, and restrictions on the use of insurance are new terrains, written into bills debated in state legislatures. Exploratory work reported here also revealed a social dimension of support for keeping abortion legal — the willingness to help a close friend or relative who has chosen to abort. On the basis of this experience, we suggest a third law of studying social change and a propose a general protocol for updating measures: “When the terms of debate change, change the questions to include new terms.” Failing to do so could threaten the content validity of existing measures. If measures get too far out of sync with the public discussion, then surveys may miss or miscast changes in opinion. By way of example, we report our original work on abortion opinions, but we feel confident that the context for other measures in the GSS and ANES has changed, too. Indeed, as time goes by the need for updates will increase, and a protocol for systematic change will serve users better than ad hoc modifications.

We propose such a protocol based on our experience developing a module on abortion opinion for the 2018 GSS. Abortion opinion is a model case for three key reasons: (1) people care about the issue, (2) their opinions on this issue correlate with important attitudes and behaviors, and (3) six of the seven items in current use were written in 1965.<sup>3</sup> As we outline below, the social, legal, and political context have all changed a great deal since. To assess whether these changes were affecting measurement, we canvassed stakeholders, asking what they wished they knew more about, turned their replies into survey questions, and fielded them as part of the 2018 GSS. In this paper we compare GSS respondents’ answers to the current questions with their answers to the new ones. Although this article is mostly about methods, our analysis has substantive implications. The new items, by themselves and in

combination with some of the current items, give new insight into abortion public opinion in the United States. With the new measures, we unearth greater variation in public opinion on both the pro-choice and pro-life side of the opinion spectrum. In doing so, we find little evidence in support of abortion polarization, despite the long-held finding that unlike other moral issues, abortion remains polarized (DiMaggio, 1997; Baldassarri and Park, 2020).

Below we detail the process by which we developed new questions; we then assess the new questions. Having proven their fitness, we answer substantive questions about abortion opinion in the United States including dimensionality and extent of polarization. It is easy to argue for more questions but doing so always incurs costs – if not financial, then in terms of respondent fatigue and the quality of survey responses. We conclude by establishing criteria for adding *and* subtracting questions to a time-series in light of these constraints. We then make specific suggestions regarding the abortion time series in the General Social Survey.

## **Abortion Public Opinion and the General Social Survey**

The first GSS in 1972 included measures of abortion opinion, and every GSS since has too, with the exception of 1986.<sup>4</sup> These original six abortion questions share a common stem: “Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if ...”. Each item then stipulates the condition of conception: if pregnancy endangers the woman’s health, if the woman has become pregnant as a result of rape, if there is a strong chance of a serious defect in the baby, if she does not want any more children, if her family is very low income and cannot afford more children, or if she is unmarried and does not want to marry the man. Written by the late Alice S. Rossi, responses to these questions have frequently been summed as the “Rossi scale” since answers to these questions can be arranged to form a hierarchical order (Clogg and Sawyer, 1981;

Hout, 1999).

The content of the six hypothetical conditions reflect the abortion debate of the time. Rossi made birth defects the first condition because Thalidomide, a tranquilizer then being prescribed to counter morning sickness, was linked to ill-formed limbs and infant deaths (Finkbine, 1967). Rossi added the woman's health and rape as additional conditions because the American Legal Institute was advocating for those exceptions to laws against abortion at the time (Rossi, 1967). She added three other "social" conditions: being poor, single, or wanting no more children. The original items behaved well as a scale (Clogg and Sawyer, 1981) and, fifty years later, proved to be quite reliable (Hout and Hastings, 2016).

Later, the board that controls GSS content added "for any reason" as a seventh condition. As posed, this question has been difficult to interpret since 20 percent of people who disagreed with one or more of the Rossi items *agreed* with "any reason." In addition, the seven-item scale proved to be less reliable than the original six-item scale (Hout and Hastings, 2016).

Opinions on abortion have sorted over time in ways that increased the correlation between abortion attitudes and partisanship. Until the mid-1980s, opinion on abortion was weakly related to partisanship, with Democrats being somewhat less supportive of abortion rights than Republicans (Layman and Carsey, 2002). Since then some individuals changed their abortion view to conform to that of their party while others changed from one party to another over the abortion issue (Fiorina, 2017). The Rossi scale's power to predict political matters has increased as a consequence of this sorting.

Despite the politicization of opinion on abortion, the overall distribution of Americans' responses to these questions have remained remarkably stable since the 1970s (Wilcox and Norrander, 2002; Fiorina et al., 2011). The long-term average of the six-point scale has been 3.9 with a high of 4.2 in 1974 (right after *Roe v. Wade*) and a low of 3.6 in 2004. Thus, it is fair to say that most Americans believe that abortion should be legal, but only under

certain circumstances.

## **Shifting terms of the abortion debate threaten old measures' content validity**

The Rossi items have provided over fifty years of comparable data on the attitudes of the general public. Like other long-running surveys, the GSS's Rossi scale focuses on abortion's legality (Appendix A lists these measures from other surveys.) Such items were developed when many states and the U.S. Congress were debating the legality of abortion. They successfully measured abortion opinion relevant to the years before and after the U.S. Supreme Court case *Roe v. Wade* legalized abortion nationwide.

With a few notable exceptions, the terms of the policy battles have shifted from whether abortion is legal to how easy or hard it is to get an abortion, an issue we and others refer to as "accessibility". These restrictions are smaller in scope than the more fundamental question of legality, but they often make abortions harder to get within their jurisdictions. The first such restriction was federal; the 1977 Hyde amendment forbade Medicaid from funding abortions specifically to limit abortion access. Representative Henry Hyde objected to abortion on religious grounds. He described his goal with the legislation in this way: "I certainly would like to prevent if I could legally, anybody from having an abortion, a rich woman, a middle-class woman, or a poor woman. Unfortunately, the only vehicle available is the HEW Medicaid bill." The tactic to create abortion barriers that some women could not overcome remains a primary approach among abortion opponents.

In addition to the terms moving from legality to access, the debate itself moved from the federal level to states. Many states restrict abortion access with some combination of waiting periods, gestational limits, counseling requirements, physician and hospital requirements, bans on insurance coverage, and others (Cohen and Joffe, 2020). Some states legislated

other restrictions only to be overturned by courts.

With questions about legality on a national level, the Rossi scale misses these matters of accessibility that dominate states' abortion legislative agenda and markedly affect the experience of obtaining an abortion (Cohen and Joffe, 2020). The Rossi scale does not capture any opinions regarding access. It does not reflect the fact that policy before and after *Roe v. Wade* was more frequently set at the state level than at the federal level. It also cannot distinguish between respondents' personal moral position and their view on the law. It does not relate to interpersonal behavior.

We contend that given these changes in the political and legal context, new items are needed that align with the current context to increase content validity. Improving content validity, should, in turn, improve construct validity. They may help us answer substantive questions about abortion which we turn to next.

## **Substantive questions of interest**

We hope that by updating abortion items we will be able to answer two important substantive questions. First, the Rossi scale captured the unidimensional reality of abortion attitudes circa 1980 (Clogg and Sawyer, 1981). As the terms of the debate have evolved from legality into other terrains, do abortion attitudes remain unidimensional? Or are matters such as abortion access, abortion morality and, especially, interpersonal aspects of abortion distinct domains of opinion?

Second, for over forty years, whether a researcher uses the six- or seven-item form of the Rossi scale, approximately 40 percent of Americans score at the maximum. Figure 1 illustrates this surprising lack of change in pro-choice abortion opinion. Though there have been meaningful fluctuations in abortion opinion, the smoothed trend shows just how large the proportions with the top scores are, as well as how small any changes over time have



been.

[Figure 1 about here]

With the Rossi scale we can say with some confidence that the rise of new issues has not lessened opposition to (or support for) abortion, but we cannot say whether the 40 percent who score at the maximum are as homogeneous as their identical scale scores make them look. That is, is the scale accurately capturing a mass of staunch supporters or is it a truncated scale that captures pro-life sentiment in far more detail than pro-choice sentiment? The answer to this question will have implications for the longstanding finding that abortion opinion is polarized (Abramowitz and Fiorina, 2013).

We answer these substantive questions in the latter part of our Results section, after describing the new questions and their development.

## **A Process to Update Existing Measures**

In order to measure elements of abortion opinion missed by the Rossi items, we engaged in a lengthy and informative process in which we enlisted the help of two groups of experts. The first group were abortion experts that provided insight onto the substance of abortion and we used as key informants on the current and future of the abortion debate. The second group were survey experts who turned the insights from the abortion experts into survey questions.

We first identified a list of over twenty experts on abortion. We then categorized them by occupation: academic, pollster, or activist, as well as their political leanings. We chose nine to capture variety in occupation and opinion, and also chose based on our assessment of whether they would understand and comply with our request. All of the nine we invited to participate accepted our invitation.

We articulated our goals and commissioned memos responding to this prompt: “What topics do you anticipate being important enough to measure now and for the next decade or so? To be clear, we do not ask you to write the questions but based on your expertise, to raise the issues.” We paid them \$1,000 each. Given the heated debate on abortion, we promised that the memos would only be shared with the group developing the new questions.

A priori we identified that we would consider the memo to be helpful if the experts provided information that would have helped and detracted from their political case. Each expert fulfilled this criterion. For instance, experts on both sides of the abortion debate mentioned insurance and taxpayer funding of abortion and both mentioned parental and spousal notification laws.

One principal investigator and a graduate student research assistant then approached the memos as if they were qualitative data. They read the memos, developed codes for common themes and excerpted the memos accordingly. The research assistant then scoured existing surveys for questions that aligned with the themes.

We recruited five survey experts to help us design new questions. We chose people who had expertise in survey question design particularly questions dealing with health, women’s issues or complex public issues such as environmental risks. We paid them \$1,250. We sent each survey expert a document with the abortion experts’ themes and memo excerpts, the survey questions from the literature search (and their marginals) as well as the abortion experts’ full memos (de-identified). We asked them to write or recommend 15 questions based on the information they received. We then circulated the union of all questions from survey experts among them and asked them to write us a memo commenting on the questions. We then held a conference call in order to pick 20 questions to field on a pilot survey. We circulated the pilot questions to the abortion experts and paid an additional \$250 for memos commenting on these sample questions. These final memos helped to make small changes to the response options.

We then pilot-tested our 20 new items (along with the Rossi scale and respondent demographics) on AmeriSpeak, NORC's nationally representative online panel of survey respondents. Our goal was to identify the questions that would cover new topics within the domain of abortion attitudes and differentiate among respondents who agreed with all six Rossi items. Having chosen eleven questions, we conducted some cognitive interviews on those on the basis of which we dropped two questions. We submitted nine questions to the GSS Board of Overseers, who approved them for pre-testing. They were then fielded in the 2018 GSS along with the Rossi scale which is a part of the GSS core.

## **Reflections on the process**

As mentioned briefly above, we took account of the fact that abortion is more politically charged than most opinions in surveys. Were we updating the questions for a less-fraught topic, we would have shared the memos among all substantive experts and invited them to discuss the merits of different proposals. We also would have invited some of the substantive experts to join the conversation with the survey experts.

We had money to spend consulting substantive and survey experts, but there are ways to reconsider questions for less money. We offered payment to our experts in our initial invitation but several people mentioned they would have done the work for free. At first, one survey expert declined to be paid (we insisted for fairness reasons). While we paid a surplus for a nationally representative pilot, a diverse convenience sample might suffice.

Lastly, we were under time constraints to meet the deadlines set for the GSS. With more time we would have done more cognitive interviews both before and after the AmeriSpeak survey. This would have also saved money as there were questions that we eventually dropped for being too confusing that would have been discarded before the AmeriSpeak survey.

## The New Questions

The nine new<sup>5</sup> questions are:

1. Leaving aside what you think of abortion for yourself (Female)/for those close to you (Male), do you think a woman should continue to be able to have an abortion legally or not, or would you say it depends?
2. Leaving aside whether you think abortion should be legal, are you morally opposed to abortion or not, or would you say it depends?
3. Here in [current state], how easy or hard do you think it is for a woman to get an abortion? Response options: Very easy, easy, neither easy nor hard, hard, very hard.
4. Here in [current state], do you think that laws should be changed to make it easier for a woman to get an abortion, be changed to make it harder for a woman to get an abortion, or should the laws stay as they are now?
5. If a close family member or friend decided to have an abortion, which of the following kinds of help, if any, would you give if you were able...
  - Help with arrangements, like a ride or childcare?
  - Help paying for the abortion?
  - Help paying for costs other than the abortion, like for a ride or hotel if she needs to stay overnight?
  - Help by providing emotional support?
6. People use their health insurance to help cover the cost of receiving health care. Do you think people should be able to use their health insurance to help cover the cost of receiving an abortion? Response options: People should be able, people should not be able. <sup>6</sup>

NORC translated the questions into Spanish following established protocols. The Spanish versions are on the GSS website.

Most novel are the four items about helping a friend or family member who is seeking an abortion and the one about insurance coverage. Other new questions vary themes familiar in traditional questions. For instance, both the Pew and Gallup legality questions begin “do you think abortion should be legal . . .”. While there are certainly merits to this approach, the new legality question has three strengths. First, it specifically puts aside the question of what respondents or someone close to them would do with regard to abortion. Second, it establishes that abortion is legal. Third, it places the woman as a subject within the question. Abortion experts suggested the first two; survey experts turned their suggestions into questions.

Most recent changes in abortion access and law were at the state level. We therefore included two state-specific questions. We first asked people to characterize abortion access in their state as easy or hard. We did so in order to stimulate them to think about their local context and to capture their perceptions of it. We caution against interpreting these questions as if they are accurate reflections of the experiences of people seeking an abortion. While some respondents certainly considered the logistics of accessing abortion in their state or a recent law that was passed there, we suspect that the majority likely gave a general impression based on their local political environment. Within states, we see considerable disagreement regarding whether it is easy or hard to access an abortion which may reflect disagreement in a subjective assessment, within-state differences in abortion access, or the differences between people who know more or less about local abortion services. Next we asked if laws about abortion access in their state should change, again prompting them to consider their state context.

In short, the new questions improved the content validity by asking about abortion in the current terms of the debate. We now assess: (1) how the measures perform; (2) whether

that improved content validity resulted in improved construct validity and (3) whether the new items help answer the substantive questions about dimensions of abortion opinion and how polarized it is in the United States.

## Data & Methods

### Data

The GSS is designed to be representative of English or Spanish-speaking adults living in households. The GSS interviewed 2,348 people in 2018, 60 percent of those sampled.<sup>7</sup> In 2018, 92 percent of the interviews were in person; 8 percent were done by telephone. Complexities of the sampling design, particularly the strategy of following only half of the initial non-respondents, make it necessary to weight cases differently when reporting descriptive statistics (Smith et al., 2019, App A). The GSS has three ballots (A, B, and C) that differ in content. Respondents are randomly assigned a ballot. The new measures, demographic variables, and the political variables we use were on all three ballots of the 2018 GSS; The Rossi items were only on A and C ( $N = 1,395$ ).

Our analyses include the new measures as well as the Rossi scale, political variables and respondent demographic variables. The Rossi scale measures support for legal abortion on a 6-point scale, where 0 denotes opposing legal abortion under all six hypothetical circumstances asked, and 6 denotes favoring legal abortion under all six hypothetical circumstances. Individuals who responded “don’t know” on any item were dropped. Many researchers add to the Rossi scale a seventh item Any Reason. We replicate all our analyses involving the Rossi scale with the seven-item scale. The results were nearly identical so we include the six-item analysis here (results available on request).

We created a Help index by summing positive responses to each of the four helping

items; as such, individuals who responded “don’t know” or did not answer one or more were dropped (8 percent of the sample).

To assess the construct validity of our new questions, we compared how well they predict political identities with how well the Rossi scale predicts those outcomes. Political party identification is coded as a seven-point scale from strong Republican to strong Democrat (dropping those who identify with other parties). We treated it as a continuous scale but reversed the usual direction because Democrats tend score high on the Rossi scale, and we find it easier to write about positive correlations than negative ones.<sup>8</sup>

We conducted parallel analyses of political ideology and vote choice. Given that party identification is more reliable than ideology (Kinder and Kalmoe, 2017) and the results from the analyses are substantively similar, we present party identification in the main text and results from models predicting ideology and 2016 vote choice in Appendix C.

In our multivariate analyses, we include a standard set of demographic and political controls to shield against the crudest forms of excluded variable bias. They are gender, age, race, marital status, labor force status, religion, religious service attendance, education, income, and U.S. residency at age 16.

## Methods

### **Does exposure to the existing items change responses to the new ones?**

We begin our analysis of the relationship between the Rossi items and our new items by exploring whether exposure to the Rossi items influenced responses to the new ones. GSS core content comes first on all ballots, so when the Rossi items were asked (ballots A and C), they were asked approximately 15 minutes before the new ones. So if the Rossi items altered responses to the new items, respondents who got ballot B will have different distributions

of answers to the new questions than respondents who got ballots A or C. Simple chi-square tests failed to reject the null hypothesis that responses on ballot B did not differ to responses on ballots A and C.<sup>9</sup>

We also assessed each subsequent analysis where exposure to the Rossi items could change our conclusions. We ran analyses separately for these two sets of respondents. We found no meaningful variation across the two groups when evaluating correlations between the new items, the reliability of these items in a scale (using Cronbach's  $\alpha$ ), the distributions of responses, the dimensionality of the items, or predicting Democratic party identification.

### **Assessing the performance of the new measures**

We assess the performance of the new items before evaluating the two substantive questions that motivated our endeavor. First, we describe the distributions of the new items and ask how our new abortion items relate to the Rossi scale and the “any reason” question. We assess the correlations between the new abortion items and the Rossi scale.

Second, we use the Rossi scale and the new abortion items to predict political party identification. We view this as an exercise in testing the construct validity of our new items, not a causal analysis. We recognize that any correlation between party and abortion opinions stems from mutual influence and co-determination. Specifically, we ask if the new items add to what the Rossi scale can already tell us about the correspondence between abortion opinions and political party identification. If the new items associate with party after controlling for the Rossi scale, then they would have improved the construct validity of the familiar scale. If the Rossi scale does not associate with party identification after controlling for the new items, then the case for new measures is even stronger.

Third, we assess the reliability of our new abortion items and compare this to the reliability of the Rossi items, measuring the internal consistency of the two sets with Cronbach's



alpha. We then add each new item to the Rossi scale and evaluate the internal consistency of each of these iterations.

Once satisfied with the validity and reliability of our new items, we will turn to our substantive analyses which we describe in more detail below in the section “Analysis plan for the substantive analyses”.

## **Results: Assessing the New Measures**

### **Distributions of the new items**

Figure 2 presents people’s responses to each new abortion item, the Rossi items, and the Any Reason question. The rising bars for the items of the Rossi scale and the Help Index hint each may form a Guttman scale. In 2018, 84 percent of people answering all the Rossi questions followed the strict Guttman pattern and another 11 percent either reversed Rape and Defect or No more and Poor. For the helping items, 95 percent followed the strict Guttman pattern. In line with the lack of change in the core measures, responses to the Rossi items and Any Reason resemble the distributions of answers in other recent years. In line with the implicit assumption of Rossi’s scale, more people answered “It depends” to the Legality and Moral Opposition questions than took an unequivocal stand. Noteworthy, as well, between 10 and 20 percent of people answered “Don’t know” to the Access and Access Reform questions. Americans split evenly on whether insurance should be used to cover the costs of abortion.

[Figure 2 about here.]

## Correlation with existing items

In Table 1 we see the pairwise correlations for the Rossi scale and each of our new items. The correlations range from around 0.8 for items that measure similar concepts (such as legality with the Rossi scale or the Help index with its components) to near zero for Access and most other items. In broad terms, this is the result we hoped for, a mix of validation and new insight.

The Legality item is the one we expected would overlap most with the Rossi scale, and indeed it does. The overlap is enough that the new item performs much like part of the original scale. The correlation of 0.78 rivals the correlations between Rossi scores measured in successive waves of the GSS panels (Hout and Hastings, 2016). Nearly all people who answered all the Rossi items the same way gave an unequivocal answer to the new legality question: 82 percent of those who *approved* of all six Rossi items also said abortion should continue to be legal while 88 percent of respondents who *disapproved* of all six Rossi items also said that abortion should not be legal anymore. People who gave differing answers to the Rossi items tended to say “it depends” to the new item; the proportion saying “it depends” was highest (79 percent) for those who approved of half the Rossi items and disapproved of half of them.

The law can reflect morality when the populace agrees on moral questions. But for contentious issues like abortion, morality implies less for the law. In 2018, 82 percent of the opposition to legal abortion came from people who objected to it on moral grounds, but a significant number of people with moral objections supported legal abortion (11 percent). And, as we saw in Figure 2 more people said the morality of abortion “depends” than stated an unequivocal opinion. Two-thirds of those for whom morality “depends” also said legality “depends.”

[Table 1 about here.]

The most distinct item in our new battery is the question about how easy it is to get an abortion in the respondent's state – the item we refer to as Access. As we mentioned above, we caution against interpreting this as an expert assessment of the ease of getting an abortion in one's state. We observed relatively limited state to state variation in Access scores in restricted data that revealed respondents' states of residence, but within-state variance was substantial.<sup>10</sup> Further, respondents had mixed accuracy with regard to across-state perceptions.

People's answers to the Access question were not biased by their support or opposition to abortion, as indicated by the very low correlations between access and both the Rossi scale and the new Legality measure in Table 1. Nor did their moral stance inflect their answer. The only marked correlation between the Access answer and another variable is for Access Reform. People who thought abortion was hard to get in their state favored making it easier; people who thought abortion was easy to get favored making it harder.

Access Reform overlaps with the more common Legality measure ( $r = 0.54$  with the Rossi scale and  $r = 0.51$  with the Legality measure), but much of its signal is distinct. The new question about insurance coverage brings to the fore one practical element of accessibility. It has elements of personal freedom that the other items miss, and its correlations are for the most part indicative of this important but modest distinction. In many ways, it parallels access reform, but its correlation with access reform is of the same magnitude as its correlation with other items.

The Help index invokes the social aspect of abortion accessibility. The people who strongly support abortion are more likely to offer multiple forms of help, while abortion opponents mainly expect to offer emotional support or none at all. The correlation is strong but far from perfect. That is because 66 percent of abortion opponents (those who score a 0 on the Rossi scale) would offer emotional support, and 30 percent would offer logistical help, to a friend or relative seeking an abortion. Of course, these abortion opponents are

less likely than supporters to ever hear of an abortion (Cowan, 2014).

## **Construct validity: Predicting Party Identification**

While we are interested in the measures to illuminate public opinion on abortion, we are also interested in their capacity to help us understand other political outcomes. We assess whether these new abortion items improve our ability to predict political outcomes, over and above the old ones. This is also a test of whether we have improved the construct validity of the abortion items. We find first, that the new items render the Rossi scale uninformative in predicting party identification and second that abortion opinion is highly politicized.

We employ Ordinary Least Squares (OLS) models to predict party identification with each of the new abortion items and the Rossi scale. We compare the coefficients of the new abortion items in models with and without the Rossi scale as a control, to assess their association with party identification beyond what is captured by the Rossi scale.<sup>11</sup>

To compare coefficients for different abortion measures we needed a common metric; we followed the usual practice of standardizing them so each has a mean of zero and a standard deviation of one. We did not standardize political party identification, though we did reverse it (strong Republican = 1 to strong Democrat = 7).<sup>12</sup> The coefficients we present can be read as the expected shift in the direction of identifying as a Democrat that corresponds to a one standard deviation increase in support for abortion legality, morality, access, or help.

We use three kinds of prediction models, with coefficients plotted in Figure 3 below. First, we use the new scale (a combination of the new items which we discuss below), the Rossi scale, and each abortion measure, on its own, to predict party identification. This coefficient is labelled "Singular" in the figure. Then we pair the Rossi scale with each of the new items and predict party identification. Finally, we assess the net contribution of the Rossi scale in a model that includes all of the new items (Legality, Moral Opposition, Access,

Access Reform, Insurance, and the Help index). Coefficients from this model are labelled "Full". All models include additive controls for gender, age, race, education, marital status, labor force status, income, U.S. residency at age 16, religion, and attendance at religious services.

[Figure 3 about here]

To ease reading Figure 3, we provide some examples. The .58 marginal effect for the Rossi scale in the singular model, represented by the hollow grey circle at the top of the plot, has party identification as the dependent variable. This model, like all others, controls for the social and demographic characteristics listed above. A one standard deviation increase in the Rossi scale results in a .58 point increase toward Democratic identification, controlling for demographic characteristics. A change of that magnitude would move a politically independent respondent three-fifths of the way toward leaning to the Democrats; an additional standard deviation increase on the Rossi scale would move that individual twice as far. When including all of the new measures and the Rossi index along with controls for respondent characteristics, as we do in the "Full" model, the regression coefficient on the Rossi scale fell to near zero (no longer statistically significant) as we see in the black diamond results of the first row. Once we know where a person stands on the new abortion items, knowing where they stand on the Rossi items adds nothing to our ability to predict their political party. We expected to learn something of predictive value from the new items, but we did not expect them to render Rossi redundant.

Now we consider the new items. A one standard deviation increase in respondents' response to Legality (row two) is associated with a .56 increase toward Democratic when controlling for personal characteristics, as seen in the singular model. When we control for the Rossi scale, the marginal effect of Legality drops to 0.33 but remains statistically significant – this is represented by the hollow, light grey square in the figure. Further controlling for all the other new items reduces the coefficient for legality to 0.20 (the black diamond in

the figure) which is still significant, though barely one-third its original magnitude.

As expected, each new item adds something to our predictions in models that include the Rossi scale as a control (the results with the square). Each estimate in between the full and singular coefficients is less than the corresponding singular estimate but still statistically significant. Access Reform is the single best-performing new item on this task of predicting party identification. All else being equal, people who thought abortion should be easier to obtain in their state were 0.57 points more Democrat, on average, than were people who thought abortion should be harder to get in their state. Most other estimates were closer to 0.30.

The strong correlations among new items (recall Table 1) make separating their effects difficult in the full model which includes the Rossi scale and all of the new items. Each of the new items has a coefficient close to 0.2 (the black diamonds in the figure on rows 2-6), although Moral Opposition is no longer a significant predictor once the other items (principally Legality and Insurance) are in the equation. That the coefficients are so similar in the Full model motivate our combining the new items and the Rossi scale into a new additive scale which we discuss below.<sup>13</sup> This new scale is strongly predictive of party identification, as evidenced by its coefficient of 0.75 in the final row of the figure.

In sum, abortion opinion, as measured by the Rossi items, the new items, and different combinations of these items, is strongly predictive of partisan identification. Whether abortion opinion concerns the circumstances of conception, state policies, morality, or interpersonal help, abortion opinion is neatly sorted by partisanship. These results illustrate how completely the issue of abortion has been politicized and provide substantial evidence of the construct validity of the new measures. Notably, our results cannot speak to causal questions; the strong relationships here may reflect the impact of abortion attitudes on partisanship or vice versa.

## Reliability

We do not have repeated survey waves fielding our new abortion measures and therefore cannot estimate their test-retest reliability. However, we can use inter-item consistency to say something about the reliability of the new items.

Cronbach's alpha ( $\alpha$ ) is a workhorse in scale construction. We use it here in two ways. First we compare reliability for the new items with the reliability of the Rossi scale. Then we use an item analysis to assess how adding new items to the Rossi scale can improve reliability. The results are in Table 2.

[Table 2 about here]

The new items were as reliable as the items of the Rossi scale by the  $\alpha$  measure; the  $\alpha$  for the new items was 0.86 and the  $\alpha$  for the Rossi scale was 0.88.<sup>14</sup> Combined in a single scale, reliability rose to 0.92, an impressive outcome for survey items administered to a nationally representative sample.<sup>15</sup> As rows 2 through 15 of the table show, deleting an item either leaves  $\alpha$  unchanged or reduces it. That means that while no item reduces reliability, a scale with fewer items might be almost as reliable. The two items that add the least, though are the two with highest agreement – the Rossi item about the woman's health and the new item about helping emotionally.

In sum, we have established the new items' fitness as survey items. By developing new items that address contemporary issues such as abortion access, we have improved content validity. Doing so, we improved construct validity as shown in the analyses predicting political party identification. The new items are as reliable as the Rossi scale. These tests establish that the new items are adequate survey items. They also demonstrate that every aspect of abortion public opinion is politicized. We turn now to our substantive questions.

# Results: Abortion opinion is unidimensional and moderate

## Analysis plan for the substantive analyses

Our first substantive question is whether our abortion items measure opinion on one or more latent dimension. We use exploratory Principal Components Analysis to assess whether our items capture opinions on orthogonal aspects of abortion. Is it the case, for instance, that Americans are more deeply divided on one dimension than on another? Are all dimensions of abortion opinion equally sorted by partisanship? We also test dimensionality using structural equation models.

Next, we assess how individual items relate to these underlying abortion positions. Specifically, item response theory (IRT) can help us see more about how the Rossi scale and the new items measure different positions along these latent dimensions. We use the one-parameter logistic model introduced by Rasch to generate item characteristic curves (Cressie and Holland, 1983).

Then we re-evaluate claims of abortion polarization, utilizing our new abortion items. Given our dimensionality findings (one dimension), we combine all items in a new additive scale. This additive scale is also motivated by the results of the models predicting party identification, as well as models predicting each of the abortion items as a function of the demographic and political covariates (results available on request). The shape of this distribution, specifically the extent to which it is bimodal, is an important indication of how polarized opinion on abortion is.

To further our understanding of polarization, we examine the distribution of the new questions' scale by respondents' scores on the Rossi scale. This permits us to consider whether the new questions are revealing variation obscured by the Rossi items and if so, for



whom is it revealing variation. Finally, we consider the relationship between respondents' subjective understanding of abortion access in their state to their wishes for changes in state law.

## Dimensionality

The new questions traverse terrain untapped by the Rossi scale and so we anticipated that we could discover new, orthogonal dimensions of abortion public opinion. We held no strong theoretical beliefs about the exact number of latent dimensions that might organize responses to these items. Given this, combined with the relatively small number of observations and variables, we assessed our expectation of multidimensionality using exploratory principal components analysis.

This analysis encompasses all the measures. Much to our surprise, a principal components analysis on the Rossi scale, Any Reason, and the new abortion items reveals only one underlying abortion opinion dimension. It yielded a single eigenvalue greater than one, and this main factor accounted for 87 percent of the common variance among the items. This result confounded our expectations since the new items were explicitly developed by experts to measure qualitatively different aspects of abortion opinion. We further tested this finding by fitting structural equation models with two or three latent dimensions. In all specifications these dimensions were highly correlated with one another ( $r > 0.7$ ). That responses can be organized by a single dimension shows that for most Americans their opinion on any particular aspect of abortion is simply informed by a general, overall sentiment.

IRT models allow us to assess how the Rossi scale and the new items measure different positions along this single latent dimension of abortion opinion. Since only dichotomous items can be used with one-parameter models, we included the new helping items along with the Rossi scale items in our model illustrated in Figure 4. The item characteristic curves array items in order of increasing difficulty – in our case, increasing support for

abortion – from left to right along the  $x$ -axis.

[Figure 4 about here]

Looking across the figure we see that the new helping items fill gaps in difficulty measured by the items of the Rossi scale. Most importantly, helping pay for the abortion was substantially more difficult to agree with than any of the Rossi items, thereby distinguishing between individuals with strong pro-choice sentiments. At the other end of the scale, offering emotional support was almost as easy as assenting to abortion when the woman's health was threatened. These new items cover more evenly positions along the latent continuum than the Rossi items have done.

## **Polarization**

We assess whether abortion is polarized through numerous analyses – all of which point to abortion opinion moderation.

We examine the overall distribution of abortion opinion by examining weighted responses to the Rossi scale, Legality, Moral Opposition, Access Reform, Insurance, and the Help Index. This scale is produced by summing responses to these items, a simple procedure justified by the consistent results of our models testing their construct validity (shown above in Figure 3). Since each item has a coefficient close to .2 in the models predicting party identification, the units of the overall scale represent fairly even differences in abortion opinion. The histogram for the new scale is shown in Figure 5.

[Figure 5 about here]

Combining responses to the new items and the Rossi scale produces a much flatter distribution of abortion opinion than the Rossi scale alone (see the second panel of Figure

2). The greater dispersion could have come merely from unearthing variation among pro-choice Americans who were top-coded by the Rossi scale (6's). In order to consider this, we examined the distribution of the new-item portion of the new scale within each of the Rossi scores (0-6), as shown in Figure 6.<sup>16</sup> Here we see that the new items uncover variation across the whole spectrum of abortion opinion.

At the pro-life pole, those respondents who score a 0 on the Rossi scale believe abortion should be illegal even when the mother's health is endangered or in cases of rape. Yet, many demonstrate some support of abortion in their answers to the new questions such as extending resources to a woman seeking an abortion or stating that the legality of abortion "depends". At the pro-choice pole, many of those who show the strongest support for abortion rights on the Rossi scale (Rossi 6's), demonstrate some reservations in their answers to the new questions. For instance, only about half of those who score 6 on the Rossi scale are willing to help pay for the abortion, though nearly all are willing to help pay for the ancillary costs. Significant minorities of Rossi 6's say that legality depends, morality depends, or that insurance should not pay for abortion.

For those in between these poles, we also see the new questions reveal variation. If the new questions did not, then these distributions would have small standard deviations; instead they range from 1.5 to 2.2. Of course it is also clear from the histograms that people who score highly on the Rossi scale, score highly on the new-item subscale, too. Indeed, the linear correlation between the two subscales is 0.80, and the between panel variance in Figure 6 is 64% of the total variance of the new-item subscale. Comparing these two statistics implies that the relationship between the two subscales is quite linear, that is  $0.80^2$  equals 0.64, meaning that any nonlinear association between the two parts of the new scale is less than 0.01.

[Figure 6 about here]

Taken together, these results indicate: (1) The new questions reveal greater variation across the spectrum of support for abortion (2) the flatness suggests a lack of polarization since polarization would create a bimodal distribution of opinion, with high numbers of respondents at either end of our abortion scale.

In answer to the question posed earlier about a ceiling effect, we see that the Rossi items produce a truncated scale and that by unearthing variation among pro-choice leaners we also unearth variation among pro-life leaners. This provides additional evidence that abortion is not as polarized in terms of the extremity of positions taken on the issue.

Our finding of moderation across the new aspects of abortion measured is a surprise given the enormous political import of the issue and that scholarly work identifies it as uniquely polarized (DiMaggio et al., 1996). We explore this more thoroughly next. As our analyses above show, all aspects of abortion opinion are politicized. Whether the issue concerns the circumstances of conception, state policies, morality, or interpersonal help, abortion opinion is neatly sorted by partisanship. At the same time, and perhaps unexpectedly, most Americans – including partisans – do not have extreme views on abortion. The opinions of Republicans and Democrats on abortion differ systematically, but they do not differ by much.

“It depends” responses predominate when we ask about either legality or morality. In contrast to the claims of activists, and to divergent Democratic and Republican Party platforms, the ethical status of abortion is not a binary one for a great many Americans. This ambivalence on abortion may well be due to internal value conflict (Alvarez and Brehm, 2002). The response of “it depends” actually mirrors the way abortion is dealt with by the law: legal in some circumstances, illegal in others.

Many people who oppose abortion, nonetheless express willingness to help a woman they know if she chose to seek an abortion. Ultimately, only ten percent of Americans would refuse to provide any of the kinds of help we asked about — emotional, logistical, or

financial. Despite their willingness to help, Americans have some reservations. Regardless of their abortion attitude, they were reluctant to help pay for the abortion itself; they were more likely to help pay for ancillary costs.

While several states have recently passed laws making it harder for women to access abortion care, other states have passed laws to insure that abortion would be legal in that state even if the federal courts reverse or weaken *Roe v Wade*. Though the states' abortion laws are diverging, public opinion appears to favor moderation. Americans, by and large, are at ease with how difficult it is to acquire an abortion in their state (see Figure 7). Belying the amount of focus paid to the issue, over 40 percent want the laws in their state to stay the same. As we can see in Figure 7, Americans who believe abortion is difficult to access in their state want to make it easier and those who believe it is easy to access want to make it harder. This finding is in line with a thermostatic model of public opinion in which preferences are negatively related to public policy changes (Wlezien, 1995).

[Figure 7 about here]

Americans are seeking to moderate what they see to be current policy; if Americans held absolute views on abortion, we would not expect to see this systematic variation in opinion on Access Reform depending on perceptions of state context. In an America populated by abortion absolutists, Americans' opinion on changing accessibility would not depend on their perceptions of how difficult or easy it is for a woman to obtain an abortion (see Appendix D for results predicting Access Reform as a function of perceived differences in accessibility at the state level, with fixed effects for states).

When thinking about the morality or legality of a woman's decision to have an abortion, Americans take into account the differing circumstances in which she does so, and position themselves in the middle ground. Similarly, when thinking about changes to abortion access, Americans reject extremely permissive or restrictive changes to state law. As such, the wider

American public does not fall into a simple dichotomy between those who believe either there should be no constraints on abortion access and those who believe abortion should be illegal. Instead, these results from the new abortion items support previous findings that most Americans eschew extreme opinions on abortion (Cook et al., 1992; Fiorina et al., 2011).

## Discussion

### Updating A Long-Running Time Series

We posed the question of when to change long-time measures of divisive opinion. Smith (2005) provided some answers to this question. We added another: when the terms of the debate have shifted enough to undermine the content validity of the existing measure(s).

We have proposed some very general methods and practices. We began with input from stakeholders to insure that the issues addressed in questions were salient today. The stakeholders' input was discursive, seldom in the form of a question for a general-population survey. For that, we turned to a panel of question-writing experts. By design, they wrote more questions than we could ultimately include. We reduced the number of items empirically, by a field trial. An online panel answered all the questions, and we picked the nine that performed best for the “main event” – a module on the GSS. The new items proved to be reliable and valid based on our analyses of responses to the 2018 GSS, presented herein.

Our large investment in measurement was motivated by the particulars of our case: Americans' abortion attitudes. Abortion is both central to American politics and academic debates on polarization (Baldassarri and Park, 2019). Thus, it is important that we measure abortion attitudes accurately; there is a lot at stake. Yet, the abortion measures in common use are dated, as are the assessments of their reliability (Clogg and Sawyer, 1981) and

validity (Hout, 1999). Six of the seven abortion items in the GSS core were written in 1965; the seventh first appeared in 1977. Empirically the data resembled entrenchment. Support for abortion cycled over the long term in small changes that were dwarfed by the variance in the six- or seven-item scales in popular use. The distributions were distinctively bimodal when few others were (DiMaggio et al., 1996). This shape of the distribution, especially the concentration of between 35 and 40 percent of cases at the maximum, has fueled the debate about whether Americans are polarized. Much rests on whether the observed maximum is a true maximum or hides substantial variation of pro-choice opinion that varies from moderate to strong.

Our new measures improved content validity by asking questions about the morality of abortion, state-level policies, abortion access, and interpersonal support for abortion. They proved their worth in terms of construct validity, in fact they outperformed the Rossi scale in predicting political party identification in 2018. Importantly, the new items arrayed along the same latent dimension of opposition to or support for abortion. Scores on the longer scale are not bimodal; they are nearly uniform. Both sides of the polarization debate will need to consider this new evidence. Many respondents who were, by the Rossi scale, maximally pro-choice were nevertheless reluctant to help a close friend or family member pay for an abortion; while on the other side of the abortion debate, many people who want to restrict abortion were surprisingly likely to offer logistical support. Advancing into the new terrain of the interpersonal revealed hitherto obscured variation. Given that the helping items scaled on the same latent dimension as the original Rossi items, these more unexpected findings support our conclusion of opinion moderation.

On the one hand, our process had a textbook character. The broad outline of concept to operationalization, refinement, field trial, and public data collection is classic. That said, it is noteworthy that few items in the GSS or ANES came to those surveys via such a systematic development and vetting. We are in an era of scientific transparency. We hope

we have contributed to that by proposing some standards that are transparent precisely because they are familiar.

Generally, a time series will prove that it has run its course when it can no longer predict other outcomes that have historically been correlated such as party-identification or vote choice. The Rossi scale continues to perform as well as it does on those tasks because it is uni-dimensional. Nevertheless, even a scale that has proven itself valuable for so long can still be improved.

We turn now to the difficult task of suggesting criteria for identifying which questions to include or exclude when the number of questions is fixed or resources are constrained. That every survey question adds financial cost and contributes to respondent fatigue necessitates parsimony (Herzog and Bachman, 1981; Yammarino et al., 1991).

These are suggestions based on our experience with the GSS abortion questions and are based both on statistical tests as well as professional judgment. Certainly, the final decision for any given survey would be the discretion of the investigators.

We suggest beginning by identifying existing questions that could easily be cut from a time series. The priority is to cut questions that are redundant in that they do not reveal new variation in opinion or reveal some new aspect of a latent attitude. One approach to doing so is to generate item characteristic curves. Curves which overlap or are close indicate redundancy. To our knowledge, no statistical test can determine which duplicative item to drop; that choice must be made on the basis of substantive expertise and professional judgment.

We have, however, developed some general principles on which to guide that judgment. One principle is to value the time series. In action, this means replacing a question that has a long history only when it's quite necessary to do so. It also means prioritizing questions that can be useful for a long time. Questions that capture a general sentiment or behavior



tend to last longer than ones tied to a specific policy. Questions that are predicated on specific social trends (e.g. a predominance of births within marriage, having had no black President) will typically have less lasting-power than questions not predicated on specific social conditions. A general rule of thumb is that if respondents would not have understood the question twenty years ago, respondents twenty years from now may not either.

In short, the goals are simply to maintain a time series when possible and when necessary (and there is a high bar for necessary), to maximize revealing variation while simultaneously maximizing the likelihood of introducing a survey item that can also pass the test of time.

## **Measuring Abortion Opinion in the GSS**

To this point, our recommendations have been about standards and process. Do we have recommendations about specific items? The GSS cannot get any longer; to add a question, one must be cut. The ICCs in Figure 4 offer an empirical guide, at least for the new helping items. An item that has an ICC close to another item's ICC, covers the same part of the dimension; an item with an ICC that is not close to others covers an otherwise uncovered portion of the dimension. Two pairs of Rossi items were very close in Figure 4: the ICC for rape was close to the one for birth defect and the ICC for poor was close to the one for birth limitation. Cutting one from each pair would be less costly than any other changes to the Rossi scale. The easy and hard Rossi items leave a space about one standard deviation wide between them. Two helping items fill that space: helping with arrangements and helping pay logistics. The original problem was, of course, 40 percent of people clustered at the top of the scale. The new item about helping pay for the abortion remedies that by being substantially harder than any hard item (or Any Reason). Separately, the morality question adds variation while asking about a timeless, general part of the abortion debate and so may well be a good addition. We caution, however, that these are more recommendations

of how to use our results than specific recommendations about what to cut or add. The GSS Board controls content and has procedures for cutting and adding.

## **Conclusion**

Debates on the fundamental issues in American life — race and gender relations, abortion, or the role of government — evolve. As they do, opinions may change in ways that older items cannot discern. While most longstanding surveys have had a policy of reviewing and updating their content since the 1990s, changes have been made by independent groups with little coordination. Procedures and standards are opaque. We have described some here that yielded fruitful results on the measurement of abortion opinion and which we believe can be applied to a host of other measures. As our major data infrastructure ages, we must continue to assess its weaknesses even as we appreciate its strengths. Monitoring social change demands continuity and consistency in measurement as exemplified by the General Social Survey; we are not calling for a major overhaul. Rather, our call is for occasional, incremental revision to be undertaken systematically and with clear standards.

## Notes

1. This statement is usually attributed to Otis Dudley Duncan. James A. Davis said it often, as does Tom Smith.
2. <https://electionstudies.org/call-for-questionnaire-ideas-2019-pilot-survey/>
3. We do not know when the seventh was written but it first appeared in the GSS in 1977.
4. See Smith et al. (2019) for a description of how some GSS items were rotated across years prior to 1988 and across ballots since 1988. The abortion items have always been on Ballots A and C since the ballot scheme was introduced in 1988.
5. The questions are new to the GSS; some appeared in other surveys before and all were part of our AmeriSpeak field trial.
6. There is a survey experiment not analyzed here. It is an experiment regarding who should determine what information a woman should have when deciding about whether to have an abortion or regarding childbirth. The GSS variables are *abmedgov1* and *abmedgov2*.
7. This corresponds to the American Association for Public Opinion Research's RR5 definition.
8. This is frequent practice in the literature, e.g., Bafumi and Shapiro, 2009.
9. Combining ballots A and C gave our tests a little more power against the null.
10. The between-state variance in responses to Access measured on a 5-point scale was 0.05, and the within-state variance 1.08. See Appendix D for the mean and standard deviation of Access by state.
11. We checked for multicollinearity between the items measuring abortion opinion in all models with more than one abortion item as an independent variable. Variance inflation factors were below 3.

12. We reversed the scoring in the original data source, which gave Republicans the higher score, to reduce the cognitive load that comes with thinking of large negative effects as “big.”
13. It is also true that the new abortion items (included in the overall scale) are similarly predicted by the same set of demographic and political variables.
14. We limited the analysis to the 1,161 cases that had valid responses to all 14 items; 26 percent of respondents were missing on one or more of the 14 items.
15. Commonly used scales often have  $\alpha$  values closer to 0.7. The GSS vocabulary quiz has an  $\alpha$  of 0.59.
16. Note that the height of the bars in Figure 6 are proportional to frequencies.

## Tables and Figures

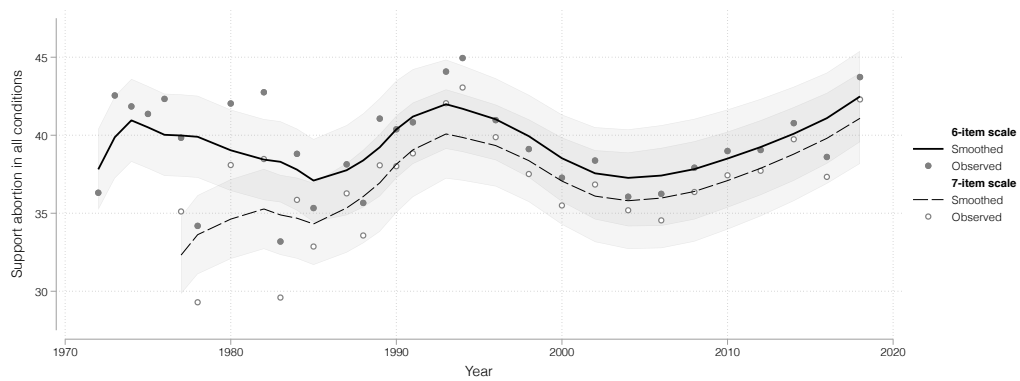


Figure 1: Observed and smoothed percentage approving of abortion in all circumstances by year and number of circumstances in the scale.

*Notes:* Data weighted to adjust for features of sampling design. Seven-item version of Rossi scale created with the addition of “any reason.” Trends lightly smoothed with bandwidth of 0.33. Shaded areas represent the 95% confidence intervals of the smoothed trends.

*Source:* General Social Surveys, 1972-2018.

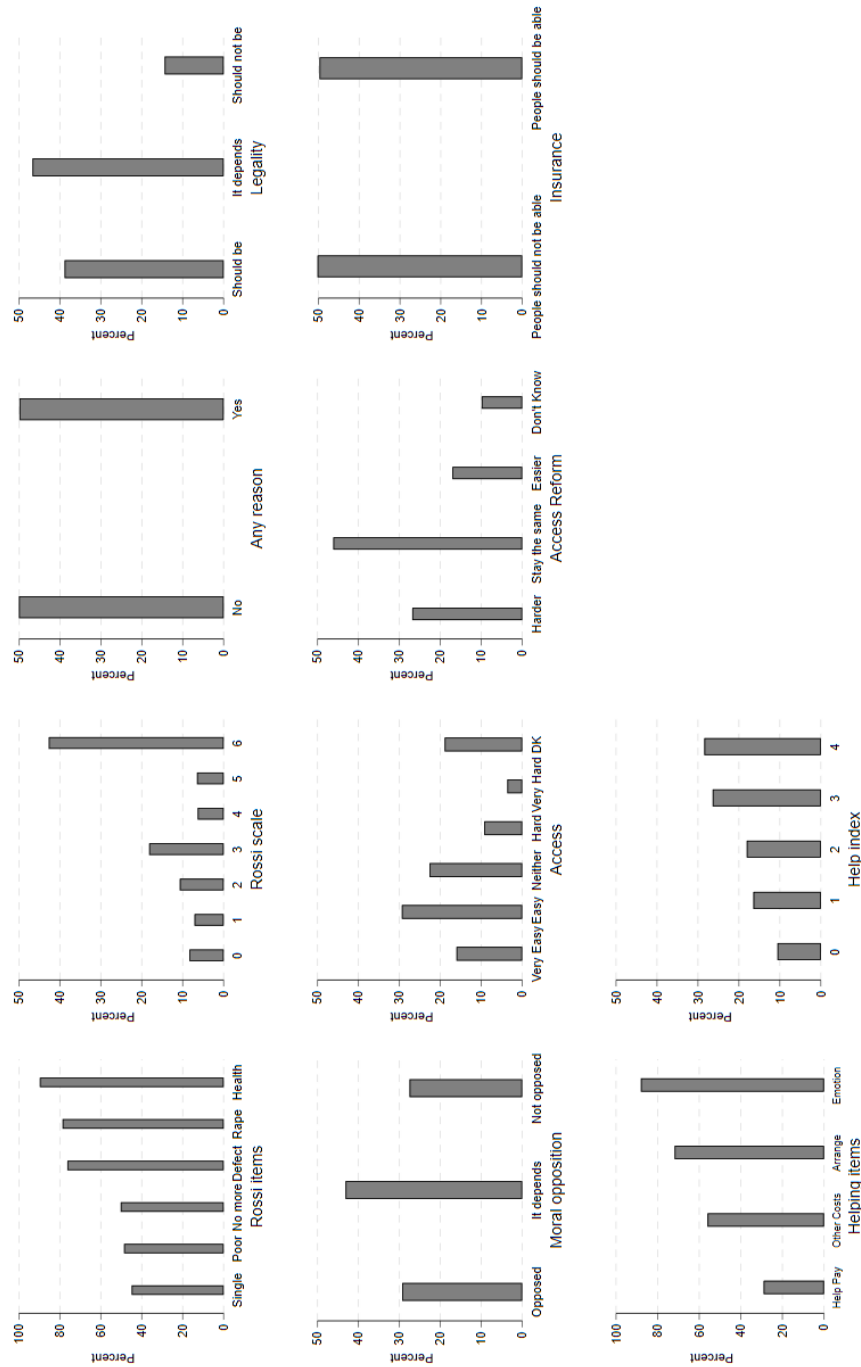


Figure 2: Histograms showing the distributions of each new abortion item, as well as “any reason,” the Rossi scale, and the Help index. Notes: Data weighted to adjust for features of sampling design. Rossi and helping items are shown individually (percentage of “yes” responses) and cumulatively as scale and index scores. The “Don’t know” replies are shown only for items where such responses accounted for over 5% of answers. Source: General Social Survey, 2018.

Table 1: Correlations among new abortion items, as well as “any reason”, the Rossi scale, and the Help index.

<i>Scale or item</i>	<i>Scale or item</i>											
	Rossi scale	Any reason	Legality	Morality	Access	Access reform	Insurance	Pay for abortion	Other costs	Arrange	Help index	
Rossi scale	1											
Any reason	0.79	1										
Legality	0.78	0.67	1									
Morality	0.58	0.52	0.56	1								
Access	0.05	0.04	0.05	0.04	1							
Access reform	0.54	0.48	0.51	0.42	0.38	1						
Insurance	0.56	0.50	0.51	0.44	0.09	0.43	1					
Help												
Pay for abortion	0.44	0.40	0.42	0.41	0.02	0.33	0.44	1				
Other costs	0.49	0.42	0.44	0.41	0.07	0.35	0.44	0.55	1			
Arrange	0.55	0.48	0.51	0.42	0.01	0.37	0.46	0.40	0.61	1		
Emotion	0.39	0.28	0.37	0.23	0.04	0.24	0.32	0.24	0.38	0.47	1	
Help index	0.61	0.53	0.57	0.49	0.05	0.43	0.55	0.75	0.86	0.81	0.62	1

*Notes:*  $N = 1,236$ . Data weighted to adjust for features of sampling design. Volunteered responses of “Don’t Know” have been included as response categories for both Access and Access Reform items, since these answers accounted for above 5% of those given. Correlations for these items without ‘Don’t Know’ as a response category are substantively similar.

*Source:* General Social Survey, 2018.

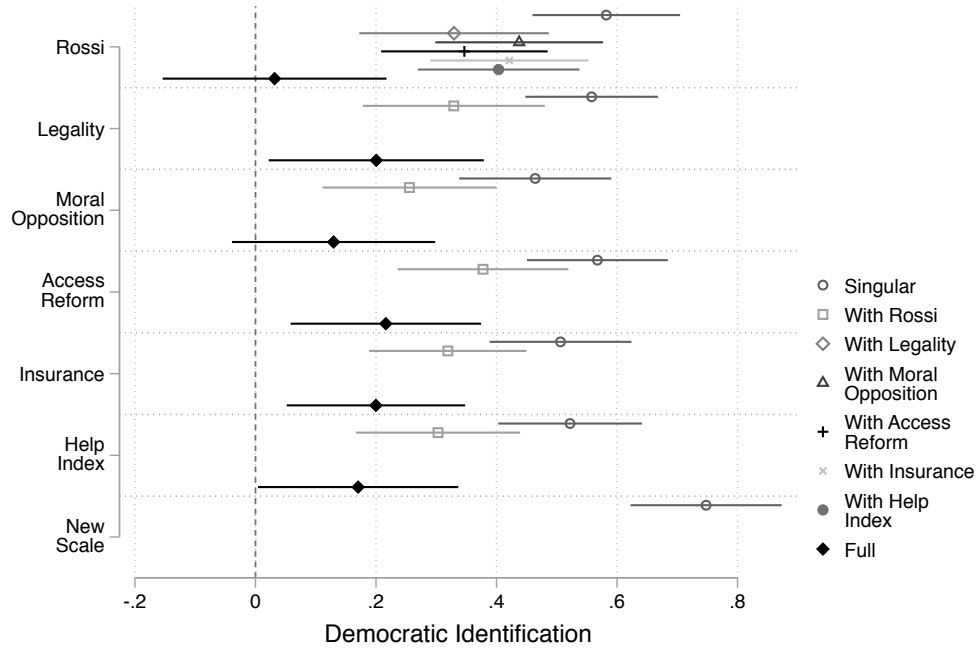


Figure 3: Modelling party identification as a function of each abortion item or scale singly; each new item or scale plus the Rossi scale; and a full model with all the new items and the Rossi scale.

(1) OLS models predicting party identification (1 = Strong Republican, 7 = Strong Democrat) with controls for gender, age, race, marital status, labor force status, religion, religious attendance, education, income, and U.S. residency at 16.

(2) Legality, Moral opposition, and Access reform are treated continuously with 3 levels, “It depends” or “Stay the same” in the middle. “Don’t know” respondents were dropped. All abortion measures were standardized.

(3) Horizontal lines show 95% confidence intervals.

Source: General Social Survey, 2018.



Table 2: Cronbach's  $\alpha$  values for items of the Rossi scale and new items for full scales and if each item is deleted from the scale

Item	Rossi scale	New item scale	All item scale
$\alpha$ for full scale	0.88	0.86	0.92
<u>Drop Rossi scale items:</u>			
Woman's Health	0.86		0.92
Rape	0.82		0.91
Birth Defect	0.86		0.91
No More	0.84		0.91
Poor	0.84		0.91
Single	0.84		0.91
<u>Drop new scale items:</u>			
Legality		0.83	0.91
Moral Opposition		0.84	0.91
Access Reform		0.83	0.91
Insurance		0.84	0.91
Arrange		0.84	0.91
Pay		0.85	0.92
Other Costs		0.84	0.91
Emotion		0.86	0.92

*Note:* The analysis is limited to the 1,161 cases that had valid responses to all 14 items.

*Source:* General Social Survey, 2018.

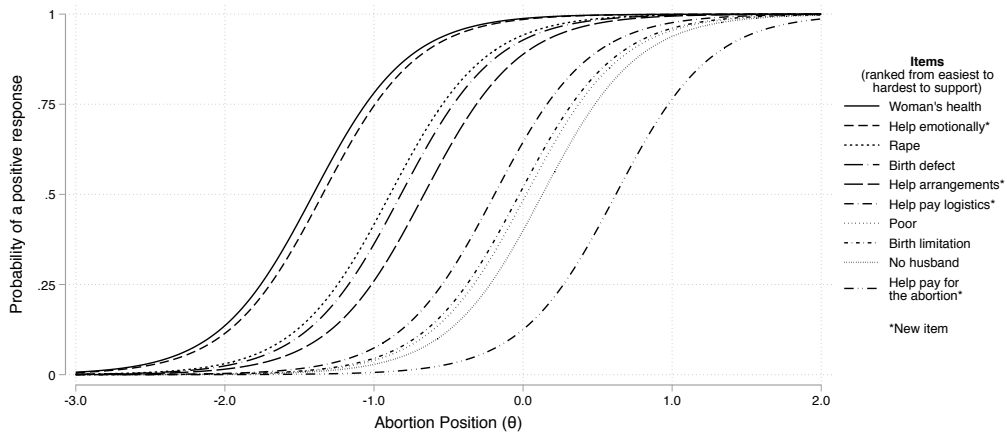


Figure 4: Item characteristic curves for the items of the combination of Rossi and helping items.

*Note:* Curves show predicted probability of a positive response according to a one-parameter logistic IRT model.

*Source:* General Social Survey, 2018.

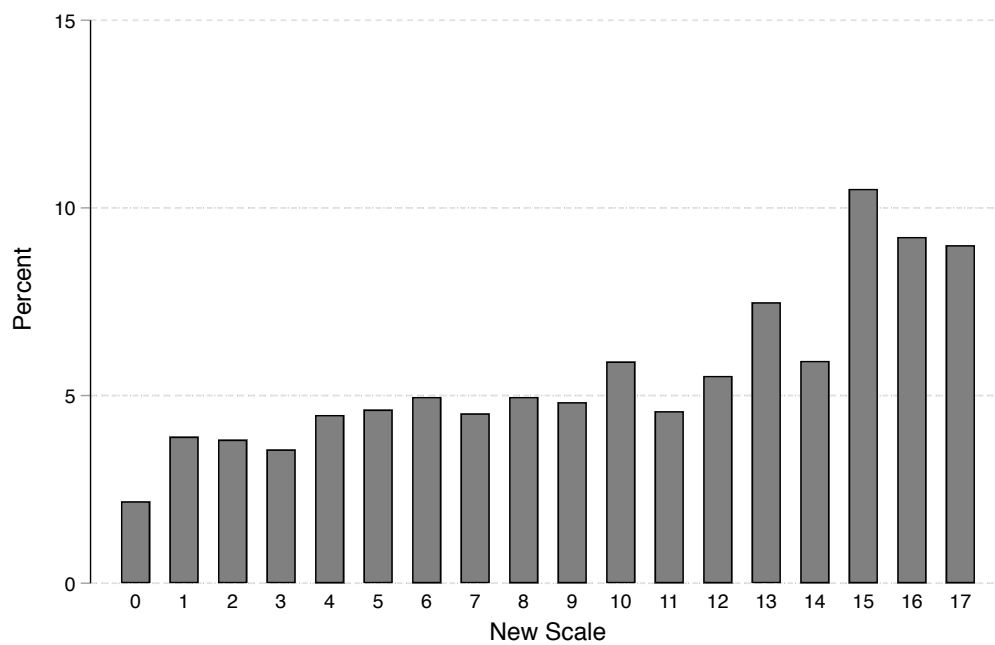


Figure 5: Weighted responses to the new abortion opinion scale, generated by summing responses to the Rossi scale, Legality, Moral Opposition, Access Reform, Insurance, and the Help Index.

*Source:* General Social Survey, 2018.

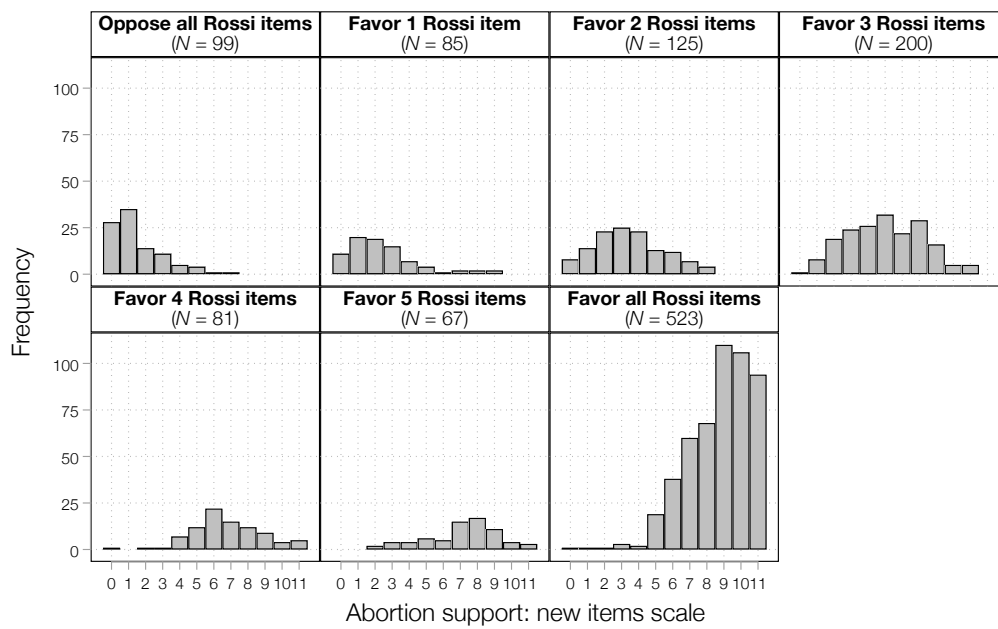


Figure 6: Histograms showing the distributions of the new-items scale by scores on the Rossi scale.

Source: General Social Survey, 2018.

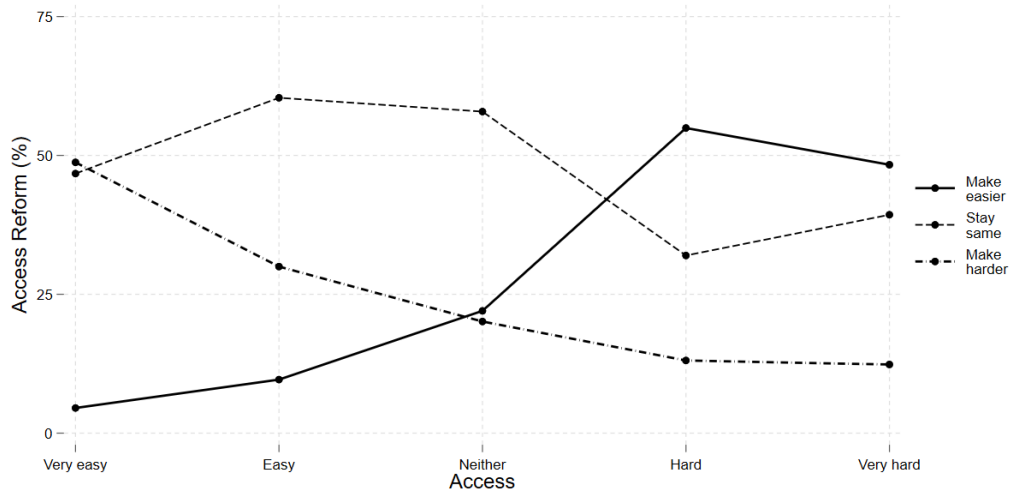


Figure 7: Preference to make abortion access easier or harder or stay the same by perceived difficulty of obtaining an abortion in the respondent's state.  
*Note:* Data weighted to adjust for features of sampling design.  
*Source:* General Social Survey, 2018.

## References

- Abramowitz, Alan I. and Morris P. Fiorina. 2013. "Polarized or Sorted? Just What's Wrong with our Politics, Anyway." *The American Interest* .
- Alvarez, Michael R. and John Brehm. 2002. *Hard choices, easy answers: Values, information, and American public opinion*. Princeton University Press.
- Bafumi, Joseph and Robert Y. Shapiro. 2009. "A New Partisan Voter." *The Journal of Politics* 71:1–24.
- Baldassarri, Delia and Barum Park. 2019. "Was There a Culture War? Polarization and Secular Trends in US Public Opinion." *Journal of Politics* Forthcoming.
- Baldassarri, Delia and Barum Park. 2020. "Was There a Culture War? Partisan Polarization and Secular Trends in U.S. Public Opinion." *Journal of Politics* 83:809–827.
- Clogg, Clifford C. and Darwin O. Sawyer. 1981. "A Comparison of Alternative Models for Analyzing the Scalability of Response Patterns." *Sociological Methodology* 12:240–280.
- Cohen, David S and Carole Joffe. 2020. *Obstacle course: The everyday struggle to get an abortion in America*. University of California Press.
- Cook, Elizabeth A., Ted G. Jelen, and Clyde Wilcox. 1992. *Between two absolutes: Public opinion and the politics of abortion*. Westview Press.
- Cowan, Sarah K. 2014. "Secrets and Misperceptions: The Creation of Self-Fulfilling Illusions." *Sociological Science* 1:466–492.
- Cressie, Noel and Paul W. Holland. 1983. "Characterizing the Manifest Probabilities or Latent Trait Models." *Psychometrika* 48:129–141.
- DiMaggio, Paul. 1997. "Culture and cognition." *Annual review of sociology* 23:263–287.
- DiMaggio, Paul, John Evans, and Bethany Bryson. 1996. "Have American's social attitudes become more polarized?" *American journal of Sociology* 102:690–755.
- Finkbine, Sherri. 1967. "The Lesser of Two Evils." In *The Case for Legalized Abortion Now*, edited by Alan F Guttmacher, pp. 15–25. Diablo Press.
- Fiorina, Morris P. 2017. *Unstable Majorities: Polarization, Party Sorting, and Political Stalemate*. Hoover Press.
- Fiorina, Morris P, Samuel J Abrams, and Jeremy C Pope. 2011. *Culture war? (3rd Ed.)*. Pearson.
- Herzog, A Regula and Jerald G Bachman. 1981. "Effects of questionnaire length on response quality." *Public Opinion Quarterly* 45:549–559.
- Hout, Michael. 1999. "Abortion politics in the United States, 1972–1994: From single issue to ideology." *Gender Issues* 17:3–34.
- Hout, Michael and Orestes P. Hastings. 2016. "Reliability of the core items in the General Social Survey." *Sociological Science* 3:971–1002.
- Kinder, Donald R. and Nathan P. Kalmoe. 2017. *Neither Liberal nor Conservative: Ideological Innocence in the American Public*. University of Chicago Press.
- Layman, Geoffrey C. and Thomas M. Carsey. 2002. "Party Polarization and "Conflict Extension" in the American Electorate." *American Journal of Political Science* 46:786.

- Marsden, Peter V., Tom W. Smith, and Michael Hout. 2020. "Tracking U.S. Social Change over a Half Century: The GSS at Fifty." *Annual Review of Sociology* 46:109–134.
- Rossi, Alice S. 1967. "Public Views on Abortion." In *The Case for Legalized Abortion Now*, edited by Alan F Guttmacher, pp. 26–53. Diablo Press.
- Smith, Tom W. 2005. "The Laws of Studying Societal Change." *Survey Research* 36:1–5.
- Smith, Tom W., Michael Davern, Jeremy Freese, and Stephen L. Morgan. 2019. *General Social Surveys, 1972–2018 [machine-readable data file]*. NORC.
- Turner, Charles and Elizabeth Martin. 1985. *Surveying Subjective Phenomena*, volume 1. Russell Sage Foundation.
- Wilcox, Clyde and Barbara Norrander. 2002. "Of moods and morals: The dynamics of opinion on abortion and gay rights." *Understanding Public Opinion* 2:121–48.
- Wlezien, Christopher. 1995. "The Public as Thermostat: Dynamics of Preferences for Spending." *American Journal of Political Science* 39:981.
- Yammarino, Francis J, Steven J Skinner, and Terry L Childers. 1991. "Understanding mail survey response behavior: A meta-analysis." *Public Opinion Quarterly* 55:613–639.

## Appendix

### A. Abortion Questions From Major U.S. Cross-Sectional Surveys

#### General Social Survey (Rossi items)

1. Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if...
  - (a) If there is a strong chance of serious defect in the baby?
  - (b) If she is married and does not want any more children?
  - (c) If the woman's own health is seriously endangered by the pregnancy?
  - (d) If the family has a very low income and cannot afford any more children?
  - (e) If she became pregnant as a result of rape?
  - (f) If she is not married and does not want to marry the man?
  - (g) The woman wants it for any reason?

#### Gallup

1. Do you think abortions should be legal under any circumstances, legal only under certain circumstances or illegal in all circumstances?
2. With respect to the abortion issue, would you consider yourself to be pro-choice or pro-life?
3. For each one, please say whether you are – very satisfied, somewhat satisfied, somewhat dissatisfied or very dissatisfied. If you don't have enough information about a particular subject to rate it, just say so. How about – the nation's policies regarding the abortion issue?
  - (a) (Asked of those dissatisfied with nations's policies on abortion) Would you like to see abortion laws in this country made more strict, less strict or remain as they are?
4. Would you like to see the Supreme Court overturn its 1973 Roe versus Wade decision concerning abortion, or not?
5. Regardless of whether or not you think it should be legal, for each one, please tell me whether you personally believe that in general it is morally acceptable or morally wrong. How about abortion?

#### Pew

1. Do you think abortion should be legal in all cases, legal in most cases, illegal in most cases? [Read in order to random half of sample, in reverse order to other half of sample]
2. Do you personally believe that each of the following is morally acceptable, morally wrong, or is it not a moral issue?...Having an abortion

#### American National Election Studies

1. Which one of the opinions on this page (1972: card) best agrees with your view? (1972-1980)
  - (a) Abortion should never be permitted.
  - (b) Abortion should be permitted only if the life and health of the woman is in danger.
  - (c) Abortion should be permitted if, due to personal reasons, the woman would have difficulty in caring for the child.
  - (d) Abortion should never be forbidden, since one should not require a woman to have a child she doesn't want.
2. Which one of the opinions on this page best agrees with your view? (1980-2016)
  - (a) By law, abortion should never be permitted



- (b) The law should permit abortion only in case of rape, incest, or when the woman's life is in danger.
- (c) The law should permit abortion for reasons other than rape, incest, or danger to the woman's life, but only after the need for the abortion has been clearly established.
- (d) By law, a woman should always be able to obtain an abortion as a matter of personal choice.

## B. Sensitivity Analyses

### Rossi as Treatment

Table B1: Sensitivity analysis of Rossi items as a potential treatment. Models predicting each of the new abortion items and the Help index, with full controls (results not shown) and a binary variable indicating whether a respondent was asked the set of Rossi items.

Asked Rossi	Legality	Moral Opposition	Access	Access Reform	Insurance	Pay for Abortion	Other Costs	Arrange	Emotion	Help Index
$\beta$	0.01	-0.05	-0.03	-0.02	-0.18	-0.08	-0.21	-0.05	-0.26	-0.10
t-test	(0.34)	(-1.50)	(-0.79)	(-0.54)	(-1.42)	(-0.59)	(-1.69)	(-0.37)	(-1.35)	(-1.60)
Model	OLS	OLS	OLS	OLS	Logit	Logit	Logit	Logit	Logit	OLS
$N$	1,939	1,942	1,608	1,774	1,835	1,906	1,919	1,904	1,931	1,858
$R^2$	0.29	0.25	0.06	0.26						0.22
pseudo $R^2$					0.16	0.14	0.13	0.15	0.12	

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

### Helping Question Order

Table B2: Sensitivity analysis of order of the helping items. Logit models predicting each of the help items with a categorical variable indicating the order in which that item was asked.

*Note:* The order of the four help questions was not randomized in the 2018 GSS, but was randomized in the AmeriSpeak survey conducted before the items were administered in the GSS. Data come from an AmeriSpeak survey, administered by NORC in the summer of 2017, on a nationally representative sample of American adults. This pilot was conducted to assess a variety of new items proposed for the 2018 GSS. The four help items from the AmeriSpeak survey are identical to those which were subsequently included in the 2018 GSS, except for the fact their order was randomized.

Question Order	Arrange		Pay for Abortion		Other Costs		Emotion	
	$\beta$	s.e.	$\beta$	s.e.	$\beta$	s.e.	$\beta$	s.e.
Order (vs. Asked First)								
Second	-0.04	(0.34)	-0.29	(0.27)	0.23	(0.28)	0.29	(0.66)
Third	0.45	(0.36)	-0.06	(0.28)	-0.19	(0.27)	-0.12	(0.62)
Fourth	0.11	(0.34)	-0.08	(0.28)	-0.13	(0.28)	-0.39	(0.65)
$N$	748		748		748		748	

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

## C. Additional Construct Validity Tests: Predicting Ideology, Whether Voted, and Vote Choice

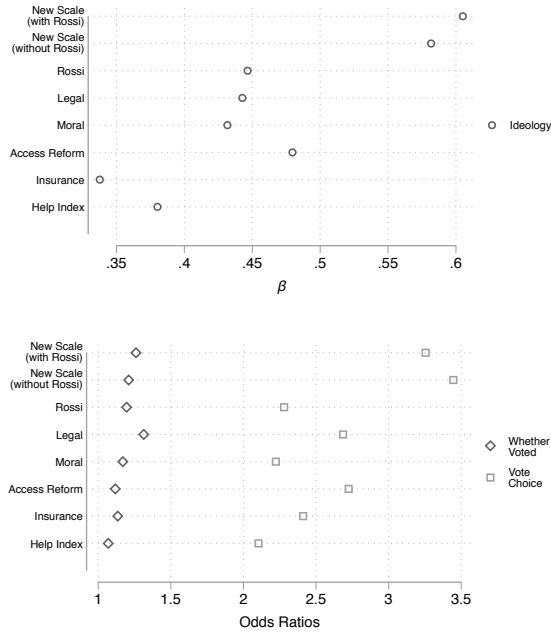


Figure C1: Results for models predicting political opinion and behavior.

(1) The models predict whether a respondent's political views (1 = Extremely Conservative, 7 = Extremely Liberal), whether they voted in the 2016 Presidential Election (1 = Voted), and whether they voted for Clinton or Trump (1 = Clinton).

(2) All models have one independent variable of interest (shown on the  $y$ -axis) and the same set of controls: gender, age, race, marital status, labor force status, religion, religious attendance, education, income, and U.S. residency at 16.

(3) Legality, Moral opposition, and Access reform are treated continuously with 3 levels, 'It depends' or 'Stay the same' in the middle. 'Don't know' respondents were dropped. All abortion measures were standardized.

(4) The first new abortion opinion scale was generated by summing responses to the Rossi scale, Legality, Moral Opposition, Access Reform, Insurance, and the Help Index. The second new scale was generated by summing responses to Legality, Moral Opposition, Access Reform, Insurance, and the Help Index.

Source: General Social Survey, 2018.

## D. Geographic Analyses

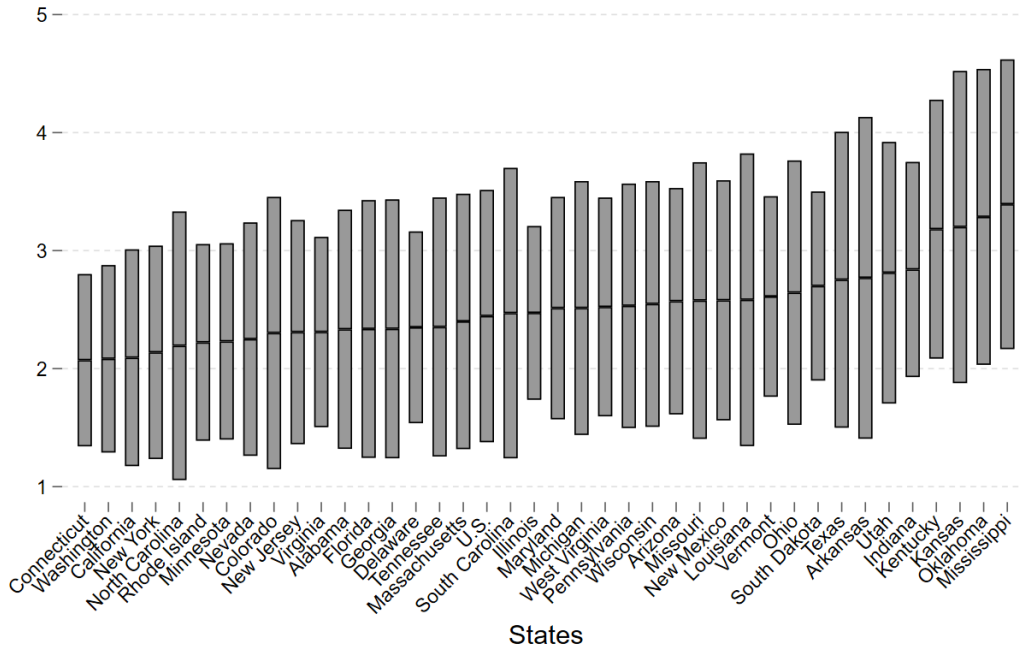


Figure D1: Mean and standard deviation of responses to respondent's perceptions of how easy or hard it is to get an abortion in their home state by U.S. state, and for the U.S. as a whole (5 = Very hard, 1 = Very easy).

*Notes:* (1) The GSS did not interview respondents from Alaska, Hawaii, Idaho, Iowa, Maine, Montana, New Hampshire, North Dakota, Oregon, or Wyoming.

(2) Statistics from Nebraska and the District of Columbia are not included since the number of observations were less than 10.

*Source:* Geocoded General Social Survey, 2018.

Table D1: Linear regression predicting Access Reform (-1 = Harder, 0 = Stay same, 1 = Easier). These analyses include fixed effects for the 41 U.S. States in which the GSS was fielded (results not presented).

	OLS Estimates	
	$\beta$	s.e.
<i>Access (vs. Very easy &amp; Easy)</i>		
Neither	0.24***	(0.04)
Very hard & Hard	0.60***	(0.05)
Don't Know	0.24***	(0.05)
Female	-0.05	(0.03)
Age	0.00	(0.00)
<i>Race (vs. White)</i>		
Black	0.06	(0.05)
Hispanic or Latino/a	-0.14**	(0.06)
All other	0.12	(0.11)
<i>Marital Status (vs. Married)</i>		
Widowed	0.03	(0.07)
Divorced	0.12*	(0.05)
Separated	0.02	(0.09)
Never married	0.09*	(0.05)
<i>Employment Status (vs. Full-time)</i>		
Part-time	0.00	(0.05)
Temporarily Unemployed	0.03	(0.11)
Unemployed	-0.08	(0.09)
Retired	-0.03	(0.05)
School	-0.09	(0.11)
Keeping House	-0.08	(0.05)
Other	-0.07	(0.14)
<i>Religion (vs. Conservative Protestant)</i>		
Mainline Protestant	0.03	(0.04)
Catholic	0.04	(0.05)
Jewish	0.21	(0.17)
Other Religion	0.19	(0.10)
No religion	0.15**	(0.05)
Weekly Attendance	-0.19***	(0.04)
<i>Education (vs. No HS Diploma)</i>		
HS Graduate	0.13*	(0.06)
Some college	0.12*	(0.06)
BA or above	0.17**	(0.06)
<i>Party Identification (vs. Independent)</i>		
Democrat	0.09*	(0.04)
Republican	-0.11*	(0.05)
<i>Ideology (vs. Moderate)</i>		
Liberal	0.30***	(0.05)
Slightly Liberal	0.13**	(0.05)
Slightly Conservative	-0.11*	(0.05)
Conservative	-0.20***	(0.05)
Income (\$90,000 and above)	0.04	(0.04)
U.S. Residence at 16	0.04	(0.07)
Constant	-0.70***	(0.15)
<i>N</i>	1,764	
<i>R</i> <sup>2</sup>	0.36	

Notes: (1) For reporting parsimony, age is treated as a continuous variable. Models with age by decade do not differ.

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

Source: Geocoded General Social Survey, 2018.