Identity as Dependent Variable: How Americans Shift Their Identities to Better Align With Their Politics

Patrick J. Egan*
Wilf Family Department of Politics
New York University

September 10, 2018

Political science generally treats demographic identities as “unmoved movers” in the chain of causality because these identities are conceptualized as being rooted in either ascriptive individual characteristics or hard-to-change aspects of individual experience. Here I hypothesize that the increasingly salient nature of partisanship and ideology as social identities leads liberal Democrats and conservative Republicans to shift their demographic identities to better align with the prototypes of their political groups, and thus the identity groups that make up the left and right coalitions in U.S. politics. I explore the hypothesis with a panel dataset that tracks Americans’ identities and political affiliations over four years. The data show that substantial numbers of Americans change how they identify over this span along the lines of national origin, sexual orientation, religion, and class. Furthermore, identity switching with regard to Latino origin, religion, class, and sexual orientation is significantly predicted by Americans’ partisanship and ideology in their pasts. All of these shifts are in directions that bring Americans’ identities into better alignment with their politics. Politics plays a particularly important role in identification with two identity groups—lesbians, gays and bisexuals, and those identifying as having no religious affiliation—in that the impact of politics on identity is large for these groups relative to their prevalence in the population. In showing how the process of identification can be imbued with politics, these findings both enrich and complicate our efforts to understand the relationship between identity and political behavior and indicate that caution must be taken in treating identities as firm, immovable political phenomena.

*patrick.egan@nyu.edu. I thank Karen Jusko, Katie McCabe, Katelyn Stauffer and participants at the 2018 APSA and MPSA meetings for helpful comments on this paper.
In recent years, identity has again emerged as a key explanatory variable in both academic and popular accounts of U.S. politics. The shift was reinforced by the 2016 presidential election, which revealed an American electorate deeply divided along the lines of identity dimensions like race, Latino origin, religion, and sexual orientation (Tyson and Maniam, 2016). Coupled with academic findings questioning the power of economic self-interest and policy preferences in shaping vote choice, some political scientists now point to group identity as a key independent variable predicting political behavior (e.g. Achen and Bartels 2016; Huddy 2018; Kinder and Kalmoe 2017), returning full circle to the focus placed on identity by some of the earliest academic accounts of voting and attitudes (Berelson, Lazarsfeld and McPhee 1954; Campbell et al 1960).

In political science research, it is often implicitly assumed that identities are stable and therefore can be confidently considered to be antecedent to political attitudes and behavior. From the perspective of drawing valid causal inferences, many identities have the appealing quality of being attached to ascriptive individual attributes that are fixed, or at least “sticky,” and thus thought to be unlikely to change in the short term. This in turn implies that the identities claimed by individuals at any given time are unlikely to be consequences of political behavior or attitudes. Thus when included as predictors in, say, a standard model of vote choice, it is often thought that we can have relative confidence in a lack of reciprocal causality between some political behavior (specified as the dependent variable) and any particular identity (the independent variable). Cross-sectional studies of political behavior by necessity take the assumption of fixed identities on faith. Many panel studies field items about ascriptive identities in only their first waves, making tests impossible with these datasets as well.

But for more and more Americans, politics has become key to the self-concept, leading “Democrat” and Republican” as well as “liberal” and “conservative” to become identities in themselves that have meanings far beyond shared policy preferences (Devine 2014; Huddy, Mason and Aarøe 2015; Iyengar, Sood and Lelkes 2012; Iyengar and Westwood
Social identity theory tells us that highly salient identities such as these can provide a definition of the self in terms of the defining characteristics of the identity group. Through a process called self-categorization, these characteristics are woven together into prototypes which become stylized representations of the kinds of persons who belong to the identity group (Turner 1985; Turner et al 1987). When an identity becomes salient, “self-perception and conduct become in-group stereotypical and normative,” leading identifiers’ beliefs and actions to converge toward those of prototypical group members (Hogg et al 1995, 260).

Here I extend self-categorization theory to hypothesize that the highly salient nature of political identities in contemporary U.S. politics can lead them to supersede other identities we typically think of as fixed, and thus counter-intuitively causing these identities to change to better align with partisan and ideological prototypes. I begin with the observation drawn from previous research that many identities can be labile and context-dependent, particularly among people who find themselves near the boundary demarcating one identity group from another. Some people who identify as liberal Democrats or conservative Republicans are therefore positioned to shift their identities to better conform with political prototypes—and thus the identity groups that make up the liberal Democratic and conservative Republican coalitions in American politics.

I explore the hypothesis using a nationally representative panel survey dataset in which questions about a range of identities were asked of empaneled respondents multiple times over four years. I find that during this span substantial numbers of Americans shift in and out of identities typically considered to be fixed, including identities associated with religion, sexual orientation, class, and national origin. Furthermore, I find that liberal Democrats and conservative Republicans are significantly likely to shift these identities in ways that conform with political group prototypes. Conservative Republicans are more likely than liberal Democrats to shift into identification as born-again Christian, Protestant, and national origins associated with being non-Hispanic white. Liberal Democrats are more likely than conservative Republicans to shift into identification as lesbian, gay or bisexual, having no religion, and being of Latino origin. Each of these shifts bring lib-
eral Democrats’ and conservative Republicans’ identities into better alignment with the identity groups that respectively make up the liberal and conservative coalitions in U.S. politics. This in turn suggests that the dynamics of identity maintenance outlined by social identity theory hold for partisanship and ideology in ways that can override identities that are usually assumed to be causally prior to political attitudes and behavior.

**Identities and Politics**

A straightforward definition of “identity” is a social category into which people are placed based upon one or more individual attributes. Attributes are mapped to identities according to membership rules that say which attributes are necessary for membership in the identity. Many important attributes are impossible to change (such as place of birth, ancestry of parents, and sexual attraction) or very hard to change (such as sex, skin tone, and other physical attributes), and for most purposes can be considered “fixed.” Another set of attributes can change but typically do so slowly, such as primary language, religion and socio-economic status; they are “sticky” (Chandra 2012). Fixed and sticky attributes are necessary for membership in many of the most highly salient identity categories in American politics, such as race, national origin, sexual orientation, religion, and class, which leads to the implicit assumptions that these identities are unchanging and that they are unmoved movers in models of political behavior.

But for two reasons, caution is called for in assuming identities like these do not change over time. The first concern arises from the distinction between objective group membership and subjective group identification (Huddy 2003). For group identifiers, group membership is incorporated into the conception of the self in ways that go beyond simply having the attributes necessary for membership in the group. In contrast to mere group members, group identifiers have a “subjective, or internalized sense of belonging to the group” (Huddy 2003: 513-514). While group membership is in many cases straightforward, group identification by contrast can be highly context dependent and up to a fair amount of individual discretion. Some identity categories, such as sexual orientation, religion, class, or national origin, are tied to attributes that can be acknowledged and em-
braced in some contexts but ignored or concealed in others, leading to variation in how some individuals identify over time. For example, people whose upbringings qualify them for objective membership in groups such as Irish Catholics, German Lutherans, or English Episcopalians can often exercise a fair amount of discretion in the extent to which they subjectively identify with these groups, essentially giving them “ethnic options” (Waters 1990). The prevalence of claimed identities can depend on many factors, including the extent to which the identity is stigmatized—or supported, as in the case of the dramatic growth in identification as American Indian in the U.S. in the mid-twentieth century (Nagel 1997). Survey responses about identity are governed by similar contingencies, as identity questions on surveys are typically self-assessments which are inevitably tied much more closely to one’s strength of group identification than one’s qualifications for group membership.

The second concern about treating identities as fixed is that many people have attributes that place them near the boundaries that demarcate one identity from another, and these boundaries tend to be fuzzy. Very few identities have clear-cut, knife-edge like membership rules that sharply separate identifiers from non-identifiers (Chandra 2006, 2012). Consider for example how in the U.S. the ability to claim or deny various racial and ethnic identities rests upon a bundle of attributes that can include fixed and sticky attributes like accent, skin tone, facial features, and body morphology, as well as changeable or concealable attributes like clothing, cultural practices, and the racial and ethnic identities of one’s parents, spouse or friends (Sen and Wasow 2016). Various mixes of these attributes can provide individuals with a lot, some, or very little discretion in identifying with a racial or ethnic group, which can in turn give rise to identity variation in different contexts, including surveys.

In models of politics, we should be particularly concerned if it is the case that political considerations are driving group identification, as that will lead standard cross-sectional studies that specify identity as predictors of political behaviors and attitudes to overestimate identity’s effect on these dependent variables. Recent research indicates this can be the case. Analyzing surveys of college students, Davenport (2016) finds that when Americans are of mixed-raced parentage, their own racial identification is strongly shaped by po-
Politically salient, causally prior characteristics like religion of upbringing, parents’ economic affluence and gender. Egan (2012) shows that the likelihood of coming out as lesbian, gay or bisexual (LGB) is significantly affected by characteristics of one’s upbringing that are causes of political attitudes, meaning that a substantial proportion of LGBs’ distinctively liberal political attitudes can be attributed to selection effects. The extent to which religion and religiosity are shaped by political attitudes has been the focus of a series of studies which include analyses of panel survey datasets of the kind employed in this paper (Putnam and Campbell 2010). Political orientations lead to shifts in religious orientations (Hout and Fischer 2014), especially when individuals experience a dissonance between their partisanship and their orientation toward secularism (Campbell et al 2018). Public opinion on “culture war” issues such as abortion and gay rights has changed religious beliefs more than vice-versa in recent years (Goren and Chapp 2017). When one is in the process of raising children, partisanship can affect parents’ religiosity (Margolis 2018). Political views can also affect the decision to leave specific religious denominations (Djupe, Neiheisel, and Sokhey 2018). The theory and analyses here present the first comprehensive explanation for why shifts across multiple identity categories can be predicted by political and ideological orientations.

Partisanship and Ideology as Social Identities

U.S. politics in our current era is characterized by historically high levels of partisan and ideological polarization among elites and masses (e.g. Abramowitz and Saunders 2008; Bafumi and Shapiro 2009; McCarty, Poole and Rosenthal 2016; Noel 2013; Pew Research Center 2017). Amid this wave of polarization, “Republican” and “Democrat,” as well as “liberal” and “conservative,” have become more than just bundles of policy preferences. They have also taken on the qualities of social identities, a hallmark characteristic of which is that in-group members make favorable comparisons between themselves and out-group members (Tajfel and Turner 1979). This is most markedly shown by the social distancing that liberals and conservatives exhibit toward one another. They rate one another negatively on survey questions; and they prefer to be friends with, date, marry, work and do
business with, and have next door neighbors from their own ideological group (Devine 2014; Huddy, Mason and Aarøe 2015; Iyengar, Sood and Lelkes 2012; Iyengar and Westwood 2015; Klofstad, McDermott and Hatemi 2012; Malka and Lelkes 2010; Mason 2018; McConnell et al 2018). Notably, in-group attachment and out-group antipathy with regard to partisan and ideological identities have been found to be stronger predictors of a variety of political behaviors and attitudes than actual policy preferences (Huddy, Mason and Aarøe 2015; Mason 2018).

One of the ways social identities become integrated into the self-concept is through self-categorization, a process by which group identifiers come to perceive themselves as similar to the prototypical identity group member and adopt beliefs and behaviors that conform to the prototype (Turner 1985; Turner et al 1987). Rich material is provided for the construction of partisan and ideological prototypes by the fact that the demographics of Democrats and Republicans, and liberals and conservatives, now differ substantially on many identity categories, a process Lilliana Mason calls “social sorting” (Mason 2016). As shown in Table 1a, data from the 2016 American National Election Studies (ANES) show that non-Hispanic whites and born-again Christians make up substantially greater shares of Republicans and conservatives than Democrats and liberals. By contrast, people of color (in particular, blacks and Latinos), LGBs, Jews, and those who claim no religious affiliation make up greater shares of Democrats and liberals than Republicans and conservatives. No substantial distinctions emerge between prototypes with regard to self-identified economic class.

The final columns of the table show that these differences generally become more acute when those who identify as both liberal and Democratic are compared to those who identify as both conservative and Republican. To quantitatively compare the degree to which the two political groups are different across identities, the final column of Table 1a reports each identity’s dissimilarity index score. A widely used measure of residential segregation, here the index measures the proportion of those claiming the identity who would have to switch political affiliations in order for equal numbers of the identity group’s members to call themselves liberal Democrats and conservative Republicans; the index is signed
Table 1: Sources of Partisan and Ideological Prototypes in U.S. Politics

a. Demographic characteristics of partisan and ideological groups, 2016 ANES

<table>
<thead>
<tr>
<th>Identity</th>
<th>Democrats</th>
<th>Republicans</th>
<th>liberals</th>
<th>conservatives</th>
<th>liberal Dems</th>
<th>conserv Reps</th>
<th>dissimilarity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>White, not Hispanic</td>
<td>57.0%</td>
<td>84.3%</td>
<td>66.2%</td>
<td>75.6%</td>
<td>65.3%</td>
<td>85.4%</td>
<td>0.28</td>
</tr>
<tr>
<td>Black, not Hispanic</td>
<td>19.0%</td>
<td>18.0%</td>
<td>12.3%</td>
<td>7.1%</td>
<td>13.4%</td>
<td>1.2%</td>
<td>-0.49</td>
</tr>
<tr>
<td>Asian/Pacific, not Hispanic</td>
<td>3.2%</td>
<td>2.6%</td>
<td>3.6%</td>
<td>2.8%</td>
<td>3.6%</td>
<td>2.5%</td>
<td>-0.10</td>
</tr>
<tr>
<td>Native, Not Hispanic</td>
<td>0.7%</td>
<td>0.5%</td>
<td>0.7%</td>
<td>0.4%</td>
<td>0.6%</td>
<td>0.6%</td>
<td>0.02</td>
</tr>
<tr>
<td>Other, not Hispanic</td>
<td>4.2%</td>
<td>3.4%</td>
<td>3.7%</td>
<td>3.8%</td>
<td>3.4%</td>
<td>3.6%</td>
<td>0.02</td>
</tr>
<tr>
<td>Hispanic</td>
<td>15.3%</td>
<td>6.7%</td>
<td>13.0%</td>
<td>9.7%</td>
<td>13.1%</td>
<td>6.3%</td>
<td>-0.20</td>
</tr>
<tr>
<td>Protestant</td>
<td>18.6%</td>
<td>37.3%</td>
<td>19.4%</td>
<td>34.4%</td>
<td>20.2%</td>
<td>41.7%</td>
<td>0.25</td>
</tr>
<tr>
<td>Catholic</td>
<td>22.2%</td>
<td>22.4%</td>
<td>19.4%</td>
<td>23.1%</td>
<td>20.5%</td>
<td>22.7%</td>
<td>0.03</td>
</tr>
<tr>
<td>Jewish</td>
<td>2.7%</td>
<td>1.2%</td>
<td>3.9%</td>
<td>0.8%</td>
<td>4.4%</td>
<td>0.9%</td>
<td>-0.36</td>
</tr>
<tr>
<td>Born-again Christian</td>
<td>27.7%</td>
<td>44.6%</td>
<td>19.9%</td>
<td>45.4%</td>
<td>18.3%</td>
<td>47.1%</td>
<td>0.32</td>
</tr>
<tr>
<td>Agnostic, atheist, none</td>
<td>28.3%</td>
<td>13.9%</td>
<td>31.9%</td>
<td>15.2%</td>
<td>33.0%</td>
<td>13.0%</td>
<td>-0.29</td>
</tr>
<tr>
<td>Lesbian, gay, bisexual</td>
<td>8.5%</td>
<td>2.1%</td>
<td>11.9%</td>
<td>2.3%</td>
<td>11.4%</td>
<td>1.4%</td>
<td>-0.45</td>
</tr>
<tr>
<td>Lower class</td>
<td>9.6%</td>
<td>5.4%</td>
<td>6.3%</td>
<td>7.8%</td>
<td>5.6%</td>
<td>4.5%</td>
<td>-0.06</td>
</tr>
<tr>
<td>Working class</td>
<td>38.7%</td>
<td>38.2%</td>
<td>34.8%</td>
<td>38.6%</td>
<td>31.2%</td>
<td>35.8%</td>
<td>0.05</td>
</tr>
<tr>
<td>Middle class</td>
<td>47.4%</td>
<td>52.4%</td>
<td>52.7%</td>
<td>50.4%</td>
<td>56.5%</td>
<td>55.4%</td>
<td>-0.01</td>
</tr>
<tr>
<td>Upper class</td>
<td>4.3%</td>
<td>3.9%</td>
<td>6.2%</td>
<td>3.2%</td>
<td>6.7%</td>
<td>4.3%</td>
<td>-0.12</td>
</tr>
</tbody>
</table>

b. Demographic characteristics of U.S. House of Representatives, 2015

<table>
<thead>
<tr>
<th>Identity</th>
<th>Democrats</th>
<th>Republicans</th>
<th>Progressive Caucus</th>
<th>Freedom Caucus</th>
</tr>
</thead>
<tbody>
<tr>
<td>White, not Hispanic</td>
<td>59.3%</td>
<td>94.4%</td>
<td>45.3%</td>
<td>97.2%</td>
</tr>
<tr>
<td>Black, not Hispanic</td>
<td>22.8%</td>
<td>0.8%</td>
<td>37.5%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Asian/Pacific, not Hispanic</td>
<td>5.3%</td>
<td>0.0%</td>
<td>6.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Native American, not Hispanic</td>
<td>0.0%</td>
<td>0.8%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>12.7%</td>
<td>4.0%</td>
<td>10.9%</td>
<td>2.8%</td>
</tr>
<tr>
<td>Protestant</td>
<td>45.0%</td>
<td>65.9%</td>
<td>46.9%</td>
<td>63.4%</td>
</tr>
<tr>
<td>Catholic</td>
<td>36.5%</td>
<td>29.4%</td>
<td>29.7%</td>
<td>25.0%</td>
</tr>
<tr>
<td>Jewish</td>
<td>9.5%</td>
<td>0.4%</td>
<td>12.5%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Religion: did not state</td>
<td>5.3%</td>
<td>0.0%</td>
<td>6.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Openly lesbian, gay or bisexual</td>
<td>3.7%</td>
<td>0.0%</td>
<td>6.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Ever had working-class job*</td>
<td>11.6%</td>
<td>4.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>From working-class background*</td>
<td>29.4%</td>
<td>13.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net worth &lt;$100,000*</td>
<td>16.1%</td>
<td>13.5%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Data from 2007.
negative for identities overrepresented among liberal Democrats and positive for those overrepresented among conservative Republicans.\(^1\) Index values regarding race, ethnicity, religion, and sexuality that are large in magnitude reflect the extent to which the nation’s political coalitions are distinctive with regard to these demographic categories.

Political elites further substantiate these prototypes. Table 1b shows that Democratic members of the House of Representatives are more heterogeneous than Republicans along the lines of race, religion, and sexual orientation in ways that parallel partisans in the general population. In addition, the backgrounds, occupations and net worth of House Democrats suggest their economic class status is somewhat lower than that of Republicans, although members of both parties are overwhelmingly from white-collar backgrounds (Carnes 2013). Demographic differences become even more pronounced upon examining the rosters of the House Freedom Caucus and House Progressive Caucus, composed of respectively the most conservative Republicans and most liberal Democrats in the House. Of the Freedom Caucus’s 36 members in 2015, 97% (all but one) were non-Hispanic white, 64% were white Protestants, and none were openly gay. By contrast, 45% of the Progressive Caucus’s members were non-Hispanic white, just 13% were white Protestants, and four of its members (6%) were among the few in the House to refuse to state a religious affiliation. Four of its members were openly gay.\(^2\)

Thus at both elite and mass levels, liberal Democrats and conservative Republicans present Americans with highly divergent prototypes along the lines of race, ethnicity, religion, and sexual orientation. Recent research indicates that Americans are not only aware of these demographic differences between the two political camps, but that they exaggerate them in their minds (Ahler and Sood 2018). Here I investigate the claim that as partisanship and ideology have become important social identities in U.S. politics, the process of self-categorization leads some people to adopt identities that conform with these prototypes and shed identities that do not.

\(\text{index} = \frac{1}{2} \left( \left| \frac{LD \cap I - LD \cap \sim I}{I - \sim I} \right| + \left| \frac{CR \cap I - CR \cap \sim I}{I - \sim I} \right| \right),\)

where \(I\) and \(\sim I\) are the shares of the population who do and do not identify as a \(j\), and \(LD\) and \(CR\) are the shares of the population identifying respectively as liberal Democrats and conservative Republicans.

\(^1\)For any identity group \(j\), the index is calculated as \(\frac{1}{2} \left( \left| \frac{LD \cap I - LD \cap \sim I}{I - \sim I} \right| + \left| \frac{CR \cap I - CR \cap \sim I}{I - \sim I} \right| \right),\)

\(^2\)Class background information was not available for members of Congress in 2015.
Data

Data come from the General Social Survey (GSS), the biennial study of Americans’ attitudes and behavior conducted by the National Opinion Research Center (Smith et al 2017). The GSS uses cluster-based sampling to obtain nationally representative samples of non-institutionalized adults in the contiguous United States. In the 2006, 2008 and 2010 GSS, respondents were empaneled to be reinterviewed two additional times, two and four years after their initial interviews. Most interviews were conducted face-to-face; successful recontact required that a greater proportion of interviews in later waves were done by phone.³

At the center of the analyses in this paper are data derived from questions asked in multiple waves about respondents’ identities with regard to race, ethnicity and national origin, religion, class, sexuality, and partisanship and ideology. Each identity category was scored dichotomously, taking on the value one if the identity is claimed by the GSS respondent and zero if not. While the GSS assessed some identities on a dichotomous basis with simple yes-or-no questions, for most identities respondents were given several responses from which to choose. These responses were recoded into dichotomous variables. For every identity, “don’t know” and “refuse to answer” responses were coded as zeroes rather than as missing data. Panelists who the GSS failed to contact for a reinterview in later waves of the panel were dropped from analyses; all analyses incorporate the panel non-response and post-stratification survey weights supplied by the GSS. Because question wording can have substantial impact on the measurement of identity, here I briefly discuss the survey items and recoding choices made for each identity category.⁴

Race. Respondents were asked “What is your race? Indicate one or more races that you consider yourself to be.” Respondents were then presented with a card featuring a list of choices. Respondents’ first reported race was coded into three dichotomous variables:

³As shown in Appendix Table 1, four-year recontact rates averaged roughly 60 percent over the two waves across the three panels. All analyses here apply the GSS’s post-stratification weights for panel attrition.
⁴Appendix Table 2 shows the GSS variables used to generate the dichotomous identity variables; Appendix Table 3 displays the number of respondents identifying with each identity in the analyses conducted in this paper.
“white,” “black,” or “Asian/Pacific.” This last category was created by collapsing several Asian and Pacific Islander identity categories from which respondents could choose, including “Asian Indian,” “Chinese,” “Japanese,” and “Native Hawaiian.”

**Hispanic/Latino origin.** A separate GSS question asked respondents, “Are you Spanish, Hispanic, or Latino [Latina if female]?” “Yes” and “no” responses were scored on a dichotomous basis. Note that therefore Latinos could be of any race according to their responses to the question described above.

**National origin.** All GSS respondents were asked the open-ended question “From what countries or part of the world did your ancestors come?” Multiple responses were permitted; here I analyze the country or place named first by respondents. The 11 most frequent responses to this item were used to create dichotomous variables, including two frequent responses—“American Indian” as well as “American only”—that do not refer to places outside the U.S.

**Sexuality.** Starting in 2008, the computer-administered self-interview (CASI) part of the GSS included the question “Which of the following best describes you?” followed by the choices (presented on the computer screen) of “gay, lesbian, or homosexual,” “bisexual,” and “heterosexual or straight.” The first two responses were collapsed to create the dichotomous variable “lesbian, gay or bisexual.”

**Religion.** GSS respondents were asked “What is your religious preference? Is it Protestant, Catholic, Jewish, some other religion, or no religion?” Responses of “Protestant,” “Catholic,” “Jewish,” and “no religion” were each coded as dichotomous variables. No other religions had substantial numbers of identifiers except for the many respondents (N = 124) who simply called themselves “Christian;’ this was also coded as a dichotomy.

---

5“American Indian or Alaska Native” was also offered as a response by the GSS, but too few people chose this category (N = 37 across valid cases in the first wave) for it to be included in analyses.

6In response to the race question, some participants volunteered that they were “Hispanic,” despite the fact that this category was not presented as a choice to respondents. (The GSS measured Hispanic/Latino identity separately as discussed below). Because volunteered responses are likely to be subject to a substantial degree of fluctuation from wave to wave of a panel survey, these responses were coded as zero for all racial categories.

7The GSS did not include a question about transgender identity.
A separate question later in the survey asked every respondent, “Would you say you have been ‘born again’ or have had a ‘born again’ experience—that is, a turning point in your life when you committed to Christ?” “Yes” and “no” responses to this question were scored as a separate dichotomous variable; born-again Christians could thus be of any religion.

**Class.** The GSS assesses self-described class by asking “If you were asked to use one of four names for your social class, which would you say you belong in: the lower class, the working class, the middle class, or the upper class?” Each of these responses was coded as a dichotomous variable.

**The Instability of Identities Over Time**

I first analyze the extent to which each identity is claimed by American adults and the extent to which identities are stable over time by focusing on identity claimed in Wave 2 of the three-wave panel. The multi-wave panel design permits expressing the total number of respondents claiming the identity in Wave 2 as those who (a) claimed the identity in both Waves 1 and 3; (b) claimed the identity in Wave 1 but not Wave 3; (c) claimed the identity in Wave 3 but not Wave 1; or (d) claimed the identity in neither Wave 1 nor Wave 3. Thus for each identity the quantity \( \frac{b+c+d}{a+b+c+d} \) is particularly revealing: ranging from zero to 100 percent, it is an estimate of the share of those claiming the identity at any given time who either did not hold the identity two years beforehand, will abandon the identity two years later, or both.

A graphical display of each identity’s switch rate is found in Figure 1a; the statistics plotted on this graph are shown in tabular form in Appendix Table 4. Identities associated with race and Latino origin exhibit the highest rates of stability, with the share of identifiers who are switchers falling below ten percent for each. By contrast, national origin identities exhibit higher rates of switching, which is likely in part due to the fact that the GSS question about national origin is open-ended. The range of switch rates among national origins is substantial. Mexican stands out as the most stable national origin, with more
Figure 1: Identity Switching Over Four Years in the Three-Wave GSS Panel Survey

a. Identity Switchers as Percent of Wave 2 Identifiers

b. Estimates of Identity Switching from Fixed-Effects Model
than 85% of Wave 2 identifiers consistently providing this response across all three waves. By contrast, two remarkably unstable national origin categories are American Indian and American Only; these data indicate that nearly all Americans claiming these identities at any given time are new to them, will decline to claim them later, or both.

Generally, religious identification is less stable than race and Latino origin, but more stable than most national origins. Those switching into and out of religious identities make up less than 20% each of Protestants, Catholics and Jews. A major exception to this pattern is found among the small share of the population who volunteer that they are simply “Christian,” as switchers make up 80 percent of these identifiers. More substantively important is the growing numbers of people who do not identify with a religious denomination: nearly four-in-ten of these identifiers have switched in, or will switch out of, this category over a four-year period. The identity of born-again Christian also exhibits a fair degree of instability; 29% of those saying at any given time they have had a born-again Christian experience either did not say so two years prior or did not say so two years later, or both.

Two identity categories exhibit much higher rates of instability: sexuality and economic class. Nearly half (47%) of those identifying as lesbian, gay or bisexual at any given time have recently switched into or will soon switch out of the identity, or both. As a society where class is not a particularly salient characteristic or organizing identity, it is not surprising that class identification in the United States exhibits relatively high rates of instability, particularly among the small share (a total of 10%) of Americans identifying as either lower class or upper class (the bottom and top of the four-class scale offered to GSS respondents, respectively). The data here indicate that more than two-thirds of those claiming these two class identities have recently switched in or will soon switch out of them. The other two class choices, working class and middle class, are claimed by much larger shares of the U.S. population; they too exhibit a fair amount of instability, with switchers making up roughly four-in-ten identifiers with each group.

\footnote{In analysis not shown here, those identifying as bisexual in Wave 2 were significantly ($p < .05$) more likely to switch in or out of LGB identity than those identifying as lesbian or gay in Wave 2.}
Measurement concerns

One concern that emerges in these analyses is the extent to which instability in identities is a function of response instability among those who select the identity rather than being attributable to the salience and centrality of the identity itself. That is, if the tendency to provide stable or unstable responses is correlated with the claiming of particular identities, a concern is that the analyses above may be confounding stability of the identity with the response stability of identifiers. To address this concern, I estimated a fixed-effects regression model in which each respondent essentially served as her own control in an analysis of the stability of identities among identifiers in Wave 2 of the panel surveys. This analysis generated estimates of switching rates that control for any time-invariant individual characteristics, including the extent to which one tends to provide inconsistent responses in surveys. Details of this estimation are found in the appendix; results are shown in Figure 1b. A comparison between these estimates and those in Figure 1a shows that the relative stabilities among identities estimated by these two approaches are broadly comparable, ruling out the concern that differences are driven by any correlation between identification and individual-level response instability.

A second concern to consider when analyzing any over-time change in panel surveys is the extent to which observed change is due to measurement error rather than change in true values over time. Unfortunately, the standard measurement model used to assess the reliability of measures in three-wave panel surveys (developed in Heise 1969 and Wiley and Wiley 1970) rests upon a crucial but untestable assumption about how true scores change that on its face is inappropriate for measurements of identity. This assumption is that the true scores change via a lag-1 (or Markovian) process, which is to say that after accounting for one lagged measure of the true score, no additional past values of the true score are meaningful predictors of the true score’s present value. This “memoryless” process is inappropriate for modeling change that unfolds over long periods of time, such as shifts in identity. In an extensive appendix to this paper, I show that when the lag-1 assumption is violated, estimates of reliability via the Heise/Wiley-Wiley model are biased,
and under conditions we would expect to be common in the measurement of identity this bias is in a negative direction. The appendix also provides evidence suggesting that these concerns apply to nearly all of the GSS measures of identity analyzed in this paper, and that therefore reliability coefficients for these measures calculated via the Heise/Wiley-Wiley model are likely too low. As shown in the appendix, this in turn leads the model to over-correct for unreliability and return what appear to be artificially inflated estimates of the stability of true scores, making it inappropriate for use in correcting for measurement error here. To help rule out the concern that measurement error explains these results, I conduct a series of placebo tests (described further below) with variables whose true values do not change over time.

In total, these data suggest that the identities survey respondents claim can be surprisingly fluid over time, even with regard to identities requiring fixed or sticky individual attributes for membership. The stability of identity categories in the contemporary U.S. ranges roughly from race and ethnicity as most stable, religion and LGB identity as less stable, and economic class as the least stable, while the stability of national origins is highly variable across different groups.

**How Americans Shift Their Identities to Align with Their Politics**

Having shown that identity switching is more commonplace than conventional wisdom suggests, I turn to an assessment of the social categorization hypothesis that liberal Democrats and conservative Republicans switch their identities to better conform with the partisan and ideological prototypes shown in Table 1. As is common in studies with panel data, the primary results in this paper are derived from models employing a lagged dependent variable specification: individual $i$’s identity at Wave 3 was modeled as a function of identity claimed four years earlier at Wave 1, partisanship and ideology at Wave 1, and controls
for each identity $j$:

$$
\text{logit } (\text{identify}_{ij,\text{wave}=3}) = \alpha + \beta_1 \text{identify}_{ij,\text{wave}=1} \\
+ \beta_2 \text{liberalism}_{i,\text{wave}=1} + \beta_3 \text{conservatism}_{i,\text{wave}=1} \\
+ \beta_4 \text{Democrat}_{i,\text{wave}=1} + \beta_5 \text{Republican}_{i,\text{wave}=1} + \text{controls} + \epsilon_i,
$$

(1)

where $\text{identify}_{ijt}$ takes on the value one if $i$ identifies as a $j$ at time $t$, and zero if not.

As shown in Equation 1, ideology and partisanship were entered into the model in a way that avoided constraining their impact on identity claiming to be monotonic. The GSS assesses ideology by asking respondents to place themselves on a seven-point scale anchored by “extremely liberal” on one side, “extremely conservative” on the other, and “moderate, middle of the road” at the center. I recoded this variable as the interval-level variable “liberalism,” scored 1 if the respondent identified as “extremely liberal,” .67 if “liberal,” .33 if “slightly liberal,” and zero if the respondent chose moderate, “don’t know,” or any of the conservative responses. “Conservatism” was analogously constructed from the conservative responses. Similarly, I recoded the GSS’s seven-point party identification variable into the interval-level variables “Democrat” and “Republican.”

Models also included controls for two variables correlated with political affiliations that could potentially confound the politics-identity relationship: age (which can be associated with shifts in identity over time due to life-cycle effects) and Wave 1 educational attainment (which is associated with response stability). Additional controls included respondents’ sex and an indicator variable for the GSS panel in which the respondent participated. The estimation incorporated survey weights for panel non-response supplied by the GSS; robust standard errors were clustered on the GSS’s primary sampling units.

Models were estimated for each of the identities shown in Figure 1, with one substantial adjustment that permitted the investigation of whether self-identified national origins shift in ways that align with the racial and ethnic prototypes of partisan and ideological identities. While most individual national origins are not strongly linked with ideol-
logical and partisan groups, these origins are associated with racial and ethnic identities that themselves distinguish ideological and partisan prototypes. Many national origins (like English, German, and Irish) are associated with non-Hispanic white identity; others (like Chinese, Mexican, and West Indian) are not. To test the hypothesis that individuals switch in and out of national origins in ways that align with partisan and ideological prototypes, I created a version of the national origin variable in which all origins associated with African, Asian, or Hispanic descent were scored one and the remainder scored zero.

Model parameter estimates of Equation 1 were used to calculate predictive margins for each identity at Wave 3 for Wave 1 conservative Republicans, Wave 1 liberal Democrats, and (as a baseline) all respondents in the GSS panels. For each identity, these predictive margins were calculated for each observation holding all other variables constant at their actual values and then averaged over the entire dataset using the GSS’s sampling weights. Figure 2 displays these estimates, with the baseline Wave 3 mean for each identity set to zero and the predictive margins for liberal Democrats and conservative Republicans plotted as departures from the baseline. These predictions can thus be interpreted as the net probability of each political group shifting into (if the prediction is positive) or out of (if negative) each identity over a four-year period compared to the general population. The left-hand side of the figure reports the differences between the predicted shifts of conservative Republicans and liberal Democrats for each identity; differences statistically significant at $p < .05$ are displayed in bold type.

The figure confirms that for many identities, the probability of claiming the identity in the present is endogenous to political affiliations in the past. Compared to conservative Republicans, liberal Democrats in Wave 1 were significantly more likely four years later in

---

9The nations and origins scored one (with values taken directly from the GSS’s nomenclature) were Africa, Arabic, China, India, Japan, Mexico, Philippines, Puerto Rico, the West Indies, and “other Asian” and “other Spanish.” The GSS’s “other Spanish” category does not include Spain, which was categorized separately and scored zero. This recoded variable includes all national origins, including those with too few identifiers to be included in Figure 1.

10For each identity $j$, these quantities are respectively $Pr(identify_{j, wave=3} = 1 | Republican_{wave=1} = .67, conservatism_{wave=1} = .67)$; $Pr(identify_{j, wave=3} = 1 | Democrat_{wave=1} = .67, liberalism_{wave=1} = .67)$; and $Pr(identify_{j, wave=3} = 1)$, holding all other individual Wave 1 characteristics constant (including whether one identified as a $j$ in Wave 1). Calculations were performed using the margins command in Stata.

11Regression output is reported in Appendix Table 7.
Wave 3 to switch into claiming identities as Latinos, lesbian, gay, or bisexual, nonreligious, lower class, and being of African, Asian or Hispanic national origin. In a similar fashion, after controlling for claimed identity in Wave 1, conservative Republicans in Wave 1 were significantly more likely than liberal Democrats four years later to identify as Protestant and as a born-again Christian. It should be noted that the sizes of these shifts are small, with the share of either political group’s members estimated to shift in or out of any identity in the low single digits. Nevertheless a substantial number of identity shifts are significantly predicted by partisanship and ideology in directions that reinforce existing political prototypes.

Source: predictive margins from estimated Equation 1. Differences between liberal Democrats and conservative Republicans that are statistically significant at \( p < .05 \) (two-tailed test) displayed in bold.
To address the concern that these results may be attributable to measurement error, I conducted placebo tests in which measures of variables whose true values do not change over time (including respondents’ zodiac sign, recalled region of residence at age 16, year of birth, and parents’ educational attainment) were substituted for the Wave 1 and Wave 3 identities in a series of regressions similar in every other respect to the estimated Equation 1. The same predictive margins were calculated for Wave 1 liberal Democrats and Wave 1 conservative Republicans, with the placebo test expectation that no significant differences should be found in the differences between the predicted shifts of the two political groups. As shown in Appendix Table 8, the share of the placebo tests (1 out of 27, or 3.7%) finding significant differences between liberal Democrats and conservative Republicans was less than would be expected by chance.

**Political Prototypes and Politicized Identity Change**

A comparison of the identity-switching patterns shown in Figure 2 with the dissimilarity index scores calculated for each identity in Table 1 provides strong support for the hypothesis that these identity shifts comport with political prototypes. Figure 3 displays this relationship, with the dissimilarity index scores again signed in the negative direction for identities that are over-represented among liberal Democrats and in the positive direction for those over-represented among conservative Republicans. What emerges is a remarkably strong relationship between prototypes and identity-switching: the more that an identity group’s members are concentrated in one of the two political groups, the greater the differences in rates at which political group members switch their identities align with political group prototypes. (The two measures are correlated at .71, $p < .01$.)

A more rigorous way to assess the relationship between prototype distinctiveness and identity switching is accomplished by pooling all observations across identities for each empaneled respondent, and then interacting the dissimilarity index score for each identity with the respondents’ partisanship and ideology variables as follows, where individuals
Figure 3. The Relationship between Political Prototypes and Politicized Identity Change

![Graph showing the relationship between political prototypes and politicized identity change.]

Sources: Dissimilarity index score: Table 1; effects on identity change: Figure 2.

are indexed \( i \) and identities are indexed \( j \):

\[
\text{identify}_{ij3} = \alpha + \sum_{j=1}^{J} \beta_j \text{identify}_{ij1} \\
+ \gamma_1 \text{liberalism}_{i1} + \gamma_2 \text{conservatism}_{i1} + \gamma_3 \text{Democrat}_{i1} + \gamma_4 \text{Republican}_{i1} \\
+ \gamma_1 \text{dissim}_j + \delta_2 (\text{dissim}_j \times \text{liberalism}_{i1}) + \delta_3 (\text{dissim}_j \times \text{conservatism}_{i1}) \\
+ \delta_4 (\text{dissim}_j \times \text{Democrat}_{i1}) + \delta_5 (\text{dissim}_j \times \text{Republican}_{i1}) + \text{controls} \\
+ \zeta_i + \xi_j + \epsilon_{ij}.
\]

In this model, the dependent variable is again individual \( i \)'s decision to identify as a \( j \) at Wave 3, controlling for identity claimed in Wave 1. Because the data are now pooled across
identities, random intercepts $\zeta_i$ are estimated for each individual and estimated standard errors are clustered at the individual level; the model now also includes fixed intercepts $\xi_j$ for each identity $j$. Here the key coefficients of interest are those on the terms interacting the identities’ dissimilarity index scores with individuals’ political variables (coefficients $\delta_2 - \delta_5$). Because the dissimilarity index scores are positively signed for identities over-represented among conservative Republicans, I expect $\delta_3$ and $\delta_5$ to be positive, reflecting conservatives’ and Republicans’ tendencies to switch identities to better comport with the conservative Republican prototype. For similar reasons, I expect $\delta_2$ and $\delta_4$ to be negatively signed, reflecting liberals’ and Democrats’ propensities to switch identities to align with the liberal Democratic prototype.

Table 2: Political Prototypes Predict Identity Shifts

<table>
<thead>
<tr>
<th></th>
<th>Wave 1 to Wave 3 identity shifts</th>
<th>$t - 1$ to $t$ identity shifts</th>
<th>model accounting for selection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>logit (1)</td>
<td>OLS (2)</td>
<td>logit (3)</td>
</tr>
<tr>
<td>identity$_j$ dissimilarity score...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ldots \times$ liberalism</td>
<td>-0.69*</td>
<td>-0.06*</td>
<td>-0.40*</td>
</tr>
<tr>
<td></td>
<td>[0.24]</td>
<td>[0.02]</td>
<td>[0.16]</td>
</tr>
<tr>
<td>$\ldots \times$ conservatism</td>
<td>0.76*</td>
<td>0.06*</td>
<td>0.54*</td>
</tr>
<tr>
<td></td>
<td>[0.28]</td>
<td>[0.02]</td>
<td>[0.20]</td>
</tr>
</tbody>
</table>

Dependent variable: claiming identity $j$ at wave $t$. Displayed are regression coefficients of interest from estimated Equation 2 and related models described in text.
Standard errors in brackets (clustered on respondent).
*Coefficients statistically different from zero at $p<.05$ (two-tailed test).

I test this hypothesis with several different specifications, as shown in Table 2, which reports the coefficients on the interaction terms of interest. Equation 2 is estimated via logit (column 1) and OLS (column 2). Two other estimation strategies incorporated Wave 2 measures of identity. The first approach (estimated with logit in column 3 and OLS in column 4) was a model similar to Equation 2 except identity at Wave $t$ was predicted by

\[12\] Regression output is shown in Appendix Table 9.
identity at Wave $t - 1$, political variables at $t - 1$, and controls, yielding two observations per identity for each GSS panelist with complete data across all three waves (that is, Wave 3 identity predicted by Wave 2 variables, and Wave 2 identity predicted by Wave 1 variables). This model also has the virtue of incorporating additional panelists who did not complete Wave 3 of the panel. The second approach (estimated with logit in column 5 and OLS in column 6) is a cross-sectional counterfactual model for three-wave panel data suggested by Morgan and Winship (2015, 385-388). It accounts for the possibility that past values of respondents’ identities may predict selection into the “treatment” of partisanship and ideology and control for any departure from the assumption that treatment and control groups exhibited parallel trends in identity claiming over time. This model includes indicator variables scored zero or one depending on whether respondents over the four years of the panel are ever in the treatment conditions of identifying as a liberal, conservative, Democrat, or Republican. These indicators are interacted with year of survey, creating separate intercepts and time-trends for each of the four treatments. The model therefore controls for the possibility that selection into any of the four treatments may be associated with either different levels of the dependent variable or with different trends in values of the dependent variable over time. All of these selection variables are interacted with the dissimilarity score; the coefficients of interest remain those on terms interacting the identities’ dissimilarity index scores with individuals’ political variables. Only respondents who completed all three waves of the survey are included in this analysis.13

The model is:

\[
\text{identify}_{ijt} = \alpha + \beta_1 \text{liberalism}_{it} + \beta_2 \text{conservatism}_{it} + \beta_3 \text{Democrat}_{it} + \beta_4 \text{Republican}_{it} \\
+ \gamma_1 \text{dissim}_j + \gamma_2 (\text{dissim}_j \times \text{liberalism}_{it}) + \gamma_3 (\text{dissim}_j \times \text{conservatism}_{it}) \\
+ \gamma_4 (\text{dissim}_j \times \text{Democrat}_{it}) + \gamma_5 (\text{dissim}_j \times \text{Republican}_{it}) \\
+ \delta_1 \text{year}_t + \delta_2 \text{everliberal}_i + \delta_3 \text{everconservative}_i + \delta_4 \text{everDemocrat}_i + \delta_5 \text{everRepublican}_i \\
+ \delta_6 (\text{year}_t \times \text{everliberal}_i) + \delta_7 (\text{year}_t \times \text{everconservative}_i) + \delta_8 (\text{year}_t \times \text{everDemocrat}_i) \\
+ \delta_9 (\text{year}_t \times \text{everRepublican}_i) + \delta_{10} (\text{dissim}_j \times \text{year}_t) + \delta_{11} (\text{dissim}_j \times \text{everliberal}_i) \\
+ \delta_{12} (\text{dissim}_j \times \text{everconservative}_i) + \delta_{13} (\text{dissim}_j \times \text{everDemocrat}_i) + \delta_{14} (\text{dissim}_j \times \text{everRepublican}_i) \\
+ \delta_{15} (\text{dissim}_j \times \text{year}_t \times \text{everliberal}_i) + \delta_{16} (\text{dissim}_j \times \text{year}_t \times \text{everconservative}_i) \\
+ \delta_{17} (\text{dissim}_j \times \text{year}_t \times \text{everDemocrat}_i) + \delta_{18} (\text{dissim}_j \times \text{year}_t \times \text{everRepublican}_i) \\
+ \text{controls} + \xi_i + \zeta_j + \epsilon_{ijt},
\]

where random intercepts $\xi_i$ are estimated for each individual $i$ and fixed intercepts $\zeta_j$ are estimated for each identity $j$, with standard errors clustered on $i$. The coefficients reported in Table 2 are $\gamma_2 \sim \gamma_5$. 

13The model is:
Across all of these specifications, all but one coefficient is signed in the theoretically expected direction. Negative coefficients on the interaction terms between the dissimilarity index score and liberalism and Democrat confirm that liberal Democrats shift their identities to align with dissimilarity index scores signed in a negative direction. Positive coefficients on the other two interaction terms demonstrate that the opposite is true for conservative Republicans. Most of the interaction term coefficients are statistically significant at the .05 level, but because the four ideology and partisanship variables are highly multicollinear an assessment of their joint statistical significance is the more appropriate test. As shown in the table, the coefficients are highly jointly significant (at $p < .001$) across all specifications. Together, these estimates provide rigorous evidence confirming the pattern shown in Figure 3: over time, a small but significant number of Americans shift their identities to align with partisan and ideological identity prototypes.

When Identities Are Infused with Politics

An additional insight that these findings offer about identity politics in the United States is that they point to identity groups whose memberships are particularly influenced by politics, in that partisanship and ideology predict shifts in identification with the group that are large relative to the size of the group in the population. As an example, consider lesbians, gays and bisexuals, a group that makes up about three percent of the U.S. adult population according to the GSS panel data. While Figure 2 shows that the association between political variables and identity shifts with regard to LGB identity appears to be small (the difference between liberal Democrats and conservative Republicans in net probability of change is two percentage points), this is actually quite a substantial change given the relatively small size of this group in the U.S. population. This is seen by considering another way to assess the magnitude of the shift, which is to transform it into the change in odds of claiming the identity at Wave 3 controlling for Wave 1 identity between the two political groups. Controlling for Wave 1 identity, the probability that conservative Republicans claim LGB identity at Wave 3 is .021 compared to .045 for liberal Democrats—a more than a doubling of the odds of claiming LGB identity. Identities for which this change in odds
is particularly large are arguably unusually shaped by politics, in that political affiliations and views can dramatically change the likelihood of identifying with the group.

Figure 4 displays this alternative measure of the influence of politics on identity group membership. On the x-axis, this graph displays the differences between conservative Republicans and liberal Democrats in the net probability of claiming identities in Wave 3, controlling for Wave 1 identity. (These are the same values displayed on the left-hand side of Figure 2.) Plotted on the y-axis are these same predictive margins now expressed
as the absolute value of the net difference in odds of claiming the identity between the two political groups. Here we see that of all the identities in this study, by this measure lesbian, gay and bisexual identity is by far the identity most endogenous to politics. In line with previous research (Hout and Fischer 2014), identification as a religious “none” is also heavily infused with politics. By contrast, the memberships of other identity groups whose membership is significantly predicted by political variables—including born-again Christian, Latino, and Protestant—are not substantially infused with politics, as political variables are associated with relatively small changes in the odds of identification.

Conclusion

These findings yield new insight on the nature of politically salient American identities and how they can be shaped by the liberal-conservative, Democrat-Republican divide. Inter-temporal stability varies highly among identities, running from relatively high (for race, Latino origin and most religions) to moderate (for party identification and some national origins) to low (for most national origins, sexual orientation, and class). Many of the identities commonly understood to be highly stable can in fact shift over time, and those who have switched in or will soon switch out of identities make up very large shares of those identifying as sexual minorities, religious “nones,” and any economic class.

These analyses permit us to see for the first time the extent to which over-time instability in identification is associated with politics, with liberalism and Democratic party identification predicting shifts toward identification as Latino, lesbian, gay, or bisexual, as nonreligious, lower class, and claiming national origin associated with being non-white; and conservatism and Republican party ID yielding movement toward identification as being a member of Protestant faith, and having had an experience as a born-again Christian. This is no small discovery: many of these identities are at the center of important American policy debates, and those who claim these identities are key blocs of voters, party activists and political donors. The data show us how in our era, which is so polarized that political affiliations become identities in themselves, politics can create and reinforce identities even thought to be as fixed as racial and ethnic categories. They thus reveal that “social sort-
ing,” while predominantly the result of individuals changing their politics to align with their identities, is also due in some part to people shifting their identities to better align with their politics.

Nearly sixty years ago, the “Michigan school” authors of The American Voter noted that the influence of group membership on political behavior might be overstated, as members of many identity groups often “come to identify with the group on the basis of pre-existing beliefs and sympathies.” (Campbell et al 1960, 323). The findings presented here join mounting evidence that this concern was well-placed, and that more rich discoveries await those who continue to make use of powerful tools and data to understand the origins of important identities in American politics.
References


APPENDIX
Appendix Table 1. General Social Survey Panels, 2006, 2008 and 2010

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FTF</td>
<td>Phone</td>
<td>Recontact rate of Wave 1 Rs</td>
</tr>
<tr>
<td>N, Wave 1</td>
<td>1,867</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>N, Wave 2 (two years after Wave 1)</td>
<td>1,251</td>
<td>285</td>
<td>76.8%</td>
</tr>
<tr>
<td>N, Wave 3 (four years after Wave 1)</td>
<td>974</td>
<td>301</td>
<td>63.8%</td>
</tr>
</tbody>
</table>
### Appendix Table 2. Identity categories, GSS variable names, and dichotomous variables generated for analysis

<table>
<thead>
<tr>
<th>Religion (RELIG)</th>
<th>Class (CLASS)</th>
<th>Born-Again Christian (REBORN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protestant</td>
<td>Lower</td>
<td>Born-again Christian</td>
</tr>
<tr>
<td>Catholic</td>
<td>Working</td>
<td></td>
</tr>
<tr>
<td>Jewish</td>
<td>Middle</td>
<td></td>
</tr>
<tr>
<td>Christian</td>
<td>Upper</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sexual Orientation (SEXORNT)</th>
<th>National origin (ETH1)</th>
<th>Non-white, non-Hispanic national origin (ETH1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesbian, gay or bisexual</td>
<td>African</td>
<td>African</td>
</tr>
<tr>
<td></td>
<td>American Indian</td>
<td>American Indian</td>
</tr>
<tr>
<td></td>
<td>American Only</td>
<td></td>
</tr>
<tr>
<td></td>
<td>English</td>
<td>Arabic</td>
</tr>
<tr>
<td></td>
<td>French</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>India</td>
</tr>
<tr>
<td></td>
<td>Irish</td>
<td>Japan</td>
</tr>
<tr>
<td></td>
<td>Italian</td>
<td>Mexico</td>
</tr>
<tr>
<td></td>
<td>Mexican</td>
<td>Philippines</td>
</tr>
<tr>
<td></td>
<td>Polish</td>
<td>Puerto Rico</td>
</tr>
<tr>
<td></td>
<td>Scottish</td>
<td>Other Asian</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other Spanish (does not include Spain)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hispanic/Latino Origin (HISPANIC)</th>
<th>Race (RACECEN)</th>
<th>Party Identification (PARTYID)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>White</td>
<td>Democrat (scored 0 to 1)</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>Republican (scored 0 to 1)</td>
</tr>
<tr>
<td></td>
<td>Asian/Pacific Islander (collapsed from following):</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian Indian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Filipino</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Japanese</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Korean</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vietnamese</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other Asian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Native Hawaiian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Guamanian or Chamorro</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Samoan</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other Pacific Islander</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ideology (POLVIEWS)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Liberalism (scored 0 to 1)</td>
<td></td>
</tr>
<tr>
<td>Conservatism (scored 0 to 1)</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix Table 3. Respondents Claiming Identities in Analyses

<table>
<thead>
<tr>
<th>Identity</th>
<th>$N$, Wave 1 switching analysis</th>
<th>$N$, Wave 1 party and ideology identity analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASIAN PACIFIC AMERICAN</td>
<td>114</td>
<td>113</td>
</tr>
<tr>
<td>BLACK</td>
<td>560</td>
<td>556</td>
</tr>
<tr>
<td>WHITE</td>
<td>2,975</td>
<td>2,962</td>
</tr>
<tr>
<td><strong>Hispanic/Latino origin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LATINO</td>
<td>381</td>
<td>378</td>
</tr>
<tr>
<td><strong>National origin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFRICAN</td>
<td>331</td>
<td>328</td>
</tr>
<tr>
<td>AMERICAN INDIAN</td>
<td>101</td>
<td>101</td>
</tr>
<tr>
<td>AMERICAN ONLY</td>
<td>102</td>
<td>101</td>
</tr>
<tr>
<td>ENGLISH</td>
<td>454</td>
<td>453</td>
</tr>
<tr>
<td>FRENCH</td>
<td>99</td>
<td>99</td>
</tr>
<tr>
<td>GERMAN</td>
<td>673</td>
<td>669</td>
</tr>
<tr>
<td>IRISH</td>
<td>425</td>
<td>424</td>
</tr>
<tr>
<td>ITALIAN</td>
<td>183</td>
<td>182</td>
</tr>
<tr>
<td>MEXICAN</td>
<td>201</td>
<td>200</td>
</tr>
<tr>
<td>POLISH</td>
<td>101</td>
<td>101</td>
</tr>
<tr>
<td>SCOTTISH</td>
<td>119</td>
<td>118</td>
</tr>
<tr>
<td><strong>Sexual orientation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGB</td>
<td>37</td>
<td>79</td>
</tr>
<tr>
<td><strong>Religion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTESTANT</td>
<td>2,001</td>
<td>1,994</td>
</tr>
<tr>
<td>CATHOLIC</td>
<td>880</td>
<td>872</td>
</tr>
<tr>
<td>JEWISH</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>NONE</td>
<td>647</td>
<td>645</td>
</tr>
<tr>
<td>CHRISTIAN</td>
<td>124</td>
<td>123</td>
</tr>
<tr>
<td>BORN-AGAIN CHRISTIAN</td>
<td>1,466</td>
<td>1,459</td>
</tr>
<tr>
<td><strong>Class</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOWER</td>
<td>281</td>
<td>278</td>
</tr>
<tr>
<td>WORKING</td>
<td>1,785</td>
<td>1,778</td>
</tr>
<tr>
<td>MIDDLE</td>
<td>1,673</td>
<td>1,666</td>
</tr>
<tr>
<td>UPPER</td>
<td>112</td>
<td>110</td>
</tr>
<tr>
<td><strong>Valid $N$ on all identities across all waves</strong></td>
<td>3,875</td>
<td>3,856</td>
</tr>
</tbody>
</table>

*Sexual orientation was not asked of all respondents, lowering the valid $N$ for analyses on change in LGB identity over time. Valid cases on sexual orientation: switching analysis: $N = 1,099$; party and ideology analysis: $N = 2,235$. 
## Appendix Table 4. Identity Claiming and Switching in the U.S. Adult Population

<table>
<thead>
<tr>
<th>Identity</th>
<th>% of population claiming identity in Wave 2</th>
<th>% of Wave 2 identity claimers switching identity (did not hold identity in Wave 1, abandoned identity in Wave 3, or both)</th>
<th>mean</th>
<th>se</th>
<th>mean</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASIAN/PACIFIC</td>
<td>3.6 (1.0)</td>
<td>6.3 (2.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLACK</td>
<td>13.6 (1.7)</td>
<td>7.3 (1.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHITE</td>
<td>76.6 (2.3)</td>
<td>5.6 (0.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LATINO*</td>
<td>12.2 (2.2)</td>
<td>6.6 (2.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National origin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFRICAN</td>
<td>7.7 (1.0)</td>
<td>46.0 (3.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMERICAN INDIAN</td>
<td>2.6 (0.4)</td>
<td>85.9 (4.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMERICAN ONLY</td>
<td>2.3 (0.4)</td>
<td>93.8 (2.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENGLISH</td>
<td>11.7 (0.9)</td>
<td>54.0 (2.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRENCH</td>
<td>2.4 (0.3)</td>
<td>69.2 (5.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GERMAN</td>
<td>16.8 (1.5)</td>
<td>39.5 (2.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRISH</td>
<td>10.3 (0.7)</td>
<td>50.1 (3.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITALIAN</td>
<td>4.8 (0.7)</td>
<td>34.8 (5.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEXICAN</td>
<td>6.6 (1.6)</td>
<td>13.4 (2.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POLISH</td>
<td>2.5 (0.3)</td>
<td>36.8 (5.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCOTTISH</td>
<td>3.1 (0.4)</td>
<td>73.5 (5.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual orientation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LESBIAN, GAY, BISEXUAL (LGB)</td>
<td>3.3 (0.5)</td>
<td>46.5 (9.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Religion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTESTANT</td>
<td>49.5 (2.6)</td>
<td>17.9 (1.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CATHOLIC</td>
<td>24.2 (1.9)</td>
<td>13.6 (1.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JEWISH</td>
<td>2.2 (0.6)</td>
<td>18.5 (5.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHRISTIAN</td>
<td>3.1 (0.0)</td>
<td>80.0 (5.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO RELIGION</td>
<td>17.2 (1.1)</td>
<td>39.1 (2.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BORN-AGAIN CHRISTIAN**</td>
<td>38.4 (2.3)</td>
<td>29.1 (1.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOWER</td>
<td>7.0 (0.5)</td>
<td>68.2 (3.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WORKING</td>
<td>46.7 (1.4)</td>
<td>44.0 (1.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIDDLE</td>
<td>43.2 (1.4)</td>
<td>38.1 (1.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPPER</td>
<td>2.8 (0.3)</td>
<td>69.0 (4.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Latinos can be of any race. **Born-again Christians can be of any religion.
Valid N = 3,875 for all identities except LGB, where N = 1,099.
Description of fixed-effects estimation of identity switching

This analysis gives rise to the estimates displayed in Figure 1b. It was conducted by restacking the dataset so that the unit of observation was at the individual \((i) \times \text{identity} (j)\) level, generating 26 observations—one for each of the \(J=26\) identities listed in Table 3—for each individual \(i\). Each \(ij\) observation included three indicator variables \((\text{identify}_{ij1}, \text{identify}_{ij2}, \text{identify}_{ij3})\) scored one if \(i\) identified as a \(j\) at wave \(t\) and zero if not. A final indicator variable, \(\text{switch}_{ij}\), was scored one if \(i\)’s responses regarding identity \(j\) switched at all during the panel; that is if \(\text{identify}_{ij2} \neq \text{identify}_{ij1}\), or \(\text{identify}_{ij2} \neq \text{identify}_{ij3}\). (A stylized representation of this restacked dataset is shown below.) I then estimated the model

\[
\text{switch}_{ij} = \alpha_j + \beta_j \text{identify}_{ij2} + \gamma_j \left( \text{identify}_{ij2} \times J \right) + \sum_{i=1}^{N} \nu_i + \epsilon_{ij},
\]

where \(J\) is an indicator variable assigned to each of the \(j = 1…26\) identity groups; and \(\nu_i\) is a fixed-effects estimate of each \(i\)’s tendency to render switch answers to survey questions about identity. The interaction \(\text{identify}_{ij2} \times J\) is included in the specification to allow for different rates of switching among those who identify as a \(j\) in wave 2 and those who do not.

Taken together, the parameter estimates \(\alpha_j, \hat{\beta}_j, \text{ and } \gamma_j\) were used to estimate

\[
\Pr\left(\text{switch}_{ij} \mid \text{identify}_{ij2} = 1\right),
\]

a measure of the average consistency with which people
identify with group $j$, controlling for any correlation in individual response stability with the tendency to identify with $j$. I estimated this equation as a linear probability model, using survey weights for panel non-response and clustering standard errors by individual $i$. I then generated the predicted probabilities of an individual’s switching in or out of an identity $j$ given that she identified as a $j$ in Wave 2, or

$$\Pr\left(\text{switch}_j \mid j \neq 1 \right).$$

These estimates are displayed in Figure 1b, a comparison of which with Figure 1a suggests that the relative stability of most identities is substantively similar across both measures.

Stylized Representation of the Dataset Analyzed in the Fixed-Effects Model

<table>
<thead>
<tr>
<th>id ($i$)</th>
<th>identity ($j$)</th>
<th>identify$_{ij1}$</th>
<th>identify$_{ij2}$</th>
<th>identify$_{ij3}$</th>
<th>switch$_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ASIAN PACIFIC AMERICAN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>BLACK</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>WHITE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>LATINO</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>AFRICAN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>AMERICAN INDIAN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>AMERICAN ONLY</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>ENGLISH</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Implications of a Violation of the Lag-1 Assumption in the Heise/Wiley-Wiley Model of Reliability in Three-Wave Panels

The Heise/Wiley-Wiley Model

Estimates of the reliability of survey items in three-wave panel designs are derived from a model introduced by Heise (1969) and further developed by Wiley and Wiley (1970). The Heise/Wiley-Wiley (HWW) model begins with a measurement model of the form

\[
Y_{i1} = T_{i1} + E_{i1} \\
Y_{i2} = T_{i2} + E_{i2} \\
Y_{i3} = T_{i3} + E_{i3},
\]

where for any individual \(i\), \(T_{it}\) is the (mean-deviated) unobserved true score at time \(t\), \(Y_{it}\) is the observed score at time \(t\), and \(E_{it}\) is random measurement error at time \(t\).\(^1\) Rather than being constant over time, the true scores are assumed to vary according to the structural model

\[
T_{it} = \beta_{t,t-1} T_{i,t-1} + Z_{it},
\]

where \(\beta_{t,t-1}\) is a parameter specifying how stable \(T\) is from one period to the next, and \(Z_{it}\) is a random disturbance associated with individual \(i\). Crucially, true scores are assumed to change via a lag-1, or Markovian, process in which the value of \(T\) at time \(t\) is dependent only on its lagged value—and not on the value of \(T\) at any previous time.

Applied to a three-wave panel, this model is written (dropping \(i\) for simplicity):

\[
T_1 = Z_1 \\
T_2 = \beta_{21} T_1 + Z_2 \\
T_3 = \beta_{32} T_2 + Z_3 \\
= \beta_{32}(\beta_{21} T_1 + Z_2) + Z_3,
\]

Plugging this into the measurement model gives

\[
Y_1 = Z_1 + E_1 \\
Y_2 = \beta_{21} Z_1 + Z_2 + E_2 \\
Y_3 = \beta_{32}(\beta_{21} Z_1 + Z_2) + Z_3 + E_3.
\]

All disturbances \(Z\) are assumed to have an expected value of zero, to be uncorrelated over time, and to be uncorrelated with the true scores and with measurement error. Measurement errors \(E\) are also assumed to have an expected value of zero, to be uncorrelated over time and to be uncorrelated with true scores. We can therefore simply write the variances of the observed scores as

\[
VAR(Y_1) = VAR(Z_1) + VAR(E_1) \\
VAR(Y_2) = \beta_{21}^2 VAR(Z_1) + VAR(Z_2) + VAR(E_2) \\
VAR(Y_3) = \beta_{32}^2[\beta_{21}^2 VAR(Z_1) + VAR(Z_2)] + VAR(Z_3) + VAR(E_3),
\]

\(^1\)Here I largely follow Alwin’s (2007, pp 104-107) notation of the model.
and the covariances of the observed scores as

\[
\text{COV}(Y_1, Y_2) = \beta_{21} \text{VAR}(Z_1) \\
\text{COV}(Y_1, Y_3) = \beta_{32}\beta_{21} \text{VAR}(Z_1) \\
\text{COV}(Y_2, Y_3) = \beta_{32}[\beta_{21} \text{VAR}(Z_1) + \text{VAR}(Z_2)].
\]

Reliability in the HWW Model

Reliability, denoted \( \rho^2 \), is defined as the ratio of the variance of true scores to that of observed scores. \( \rho^2 \) approaches one as observed scores approach perfect alignment with true scores and the variance of \( E \) approaches zero:

\[
\rho^2 = \frac{\text{VAR}(T)}{\text{VAR}(Y)} = \frac{\text{VAR}(Z)}{\text{VAR}(Y)} = \frac{\text{VAR}(Y) - \text{VAR}(E)}{\text{VAR}(Y)} = 1 - \frac{\text{VAR}(E)}{\text{VAR}(Y)}.
\]

Here Heise and Wiley and Wiley part ways: Heise assumes that reliability is constant across time periods; Wiley and Wiley do not. This distinction is not relevant here, as I focus on reliability at \( t = 2 \), which both approaches agree is

\[
\rho_{\text{HWW}}^2 = 1 - \frac{\text{VAR}(E_2)}{\text{VAR}(Y_2)} = 1 - \frac{\text{VAR}(Y_2) - \beta_{21}\text{VAR}(Z_1) - \text{VAR}(Z_2)}{\text{VAR}(Y_2)} = \frac{\beta_{21}\text{VAR}(Z_1) + \text{VAR}(Z_2)}{\text{VAR}(Y_2)} = \frac{\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)}{\beta_{21}\text{VAR}(Z_1) + \text{VAR}(Z_2) + \text{VAR}(E_2)}.
\]

Here, \( \rho_{\text{HWW}}^2 \) (which for simplicity I denote going forward as \( \rho^2 \)) is written as a function of primitive, unobservable elements of the model. Heise and Wiley and Wiley show that the same quantity can be calculated from the following ratio of the variances and covariances of observable scores \( Y \):

\[
\frac{\text{COV}(Y_1, Y_2)\text{COV}(Y_2, Y_3)}{\text{COV}(Y_1, Y_3)\text{VAR}(Y_2)} = \frac{\beta_{21}\text{VAR}(Z_1)\beta_{32}[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)]}{\beta_{32}\beta_{21}\text{VAR}(Z_1)[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2) + \text{VAR}(E_2)]} = \frac{\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)}{\beta_{21}\text{VAR}(Z_1) + \text{VAR}(Z_2) + \text{VAR}(E_2)} = \rho^2.
\]

When the Lag-1 Assumption Fails to Hold

As noted above, a key assumption of the HWW model is that true scores \( T \) change via a lag-1 autocorrelation process. It is a “short memory” model, in that after accounting for one lagged value of the true score \( T \), it is assumed that no previous information about \( T \) meaningfully predicts its current value. This is an untestable assumption that on its face would seem difficult to justify in analyses of long-term processes that unfold over many years, such as the formation of identity.
What happens to estimates of $\rho^2$ when the lag-1 assumption fails to hold? I investigate this question by developing a model that is similar the HWW model in every respect except that true scores follow a lag-2 process. In this revised model, all assumptions remain the same with one exception: the lag-2 nature of the process is reflected by adding the term $\beta_{31}T_1$ to the equation for $T_3$, where $\beta_{31}$ represents the lagged direct effect of $T_1$ on $T_3$, controlling for $T_2$:

\[
\begin{align*}
T_1 &= Z_1 \\
T_2 &= \beta_{21}T_1 + Z_2 \\
T_3 &= \beta_{32}T_2 + \beta_{31}T_1 + Z_3 \\
    &= \beta_{32}(\beta_{21}T_1 + Z_2) + \beta_{31}Z_1 + Z_3.
\end{align*}
\]

The observed scores are now

\[
\begin{align*}
Y_1 &= Z_1 + E_1 \\
Y_2 &= \beta_{21}Z_1 + Z_2 + E_2 \\
Y_3 &= \beta_{32}(\beta_{21}Z_1 + Z_2) + \beta_{31}Z_1 + Z_3 + E_3,
\end{align*}
\]

and the variances and covariances of the observed scores become

\[
\begin{align*}
\text{VAR}(Y_1) &= \text{VAR}(Z_1) + \text{VAR}(E_1) \\
\text{VAR}(Y_2) &= \beta_{31}^2\text{VAR}(Z_1) + \text{VAR}(Z_2) + \text{VAR}(E_2) \\
\text{VAR}(Y_3) &= \beta_{32}^2[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{31}^2\text{VAR}(Z_1) + \text{VAR}(Z_3) + \text{VAR}(E_3)
\end{align*}
\]

\[
\begin{align*}
\text{COV}(Y_1, Y_2) &= \beta_{21}\text{VAR}(Z_1) \\
\text{COV}(Y_1, Y_3) &= (\beta_{32}\beta_{21} + \beta_{31})\text{VAR}(Z_1) \\
\text{COV}(Y_2, Y_3) &= \beta_{32}[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{31}\beta_{21}\text{VAR}(Z_1).
\end{align*}
\]

Calculations of $\rho^2$ using the standard HWW formula for reliability now return a biased estimate of $\rho^2$, denoted here as $\rho_B^2$:

\[
\rho_B^2 = \frac{\text{COV}(Y_1, Y_2)\text{COV}(Y_2, Y_3)}{\text{COV}(Y_1, Y_3)\text{VAR}(Y_2)}
= \frac{\beta_{21}\text{VAR}(Z_1)[\beta_{32}[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{21}\beta_{31}\text{VAR}(Z_1)]}{(\beta_{32}\beta_{21} + \beta_{31})[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{21}\beta_{31}\text{VAR}(Z_1)}
= \frac{\beta_{21}[\beta_{32}[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{21}\beta_{31}\text{VAR}(Z_1)]}{(\beta_{32}\beta_{21} + \beta_{31})[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{21}\beta_{31}\text{VAR}(Z_1)}.
\]

Substituting the expression for $\rho^2$ from (1) above, write the bias of $\rho_B^2$ for $\rho^2$ as

\[
\text{BIAS}[\rho_B^2] = \rho_B^2 - \rho^2
= \frac{\beta_{21}[\beta_{32}[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{21}\beta_{31}\text{VAR}(Z_1)]}{(\beta_{32}\beta_{21} + \beta_{31})[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{21}\beta_{31}\text{VAR}(Z_1)} - \frac{\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)}{\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2) + \text{VAR}(E_2)}
= -\frac{\beta_{31}(\beta_{21}(\beta_{21} - 1)\text{VAR}(Z_1) + \text{VAR}(Z_2))}{(\beta_{21}\beta_{32} + \beta_{31})[\beta_{21}^2\text{VAR}(Z_1) + \text{VAR}(Z_2)] + \beta_{21}\beta_{31}\text{VAR}(Z_1)).}
\]
Four properties of $\rho_B^2$ provide cause for concern that the HWW formula produces biased estimates of reliability for the identity measures in this paper.

1. $\rho_B^2$ is biased for $\rho^2$ whenever the lag-1 assumption is violated (i.e., $\beta_{31}$ is nonzero).

By inspection, expression (2) indicates that $\rho_B^2$ simplifies to $\rho^2$ whenever $\beta_{31} = 0$, but $\beta_{31} \neq 0 \implies \rho_B^2 \neq \rho^2$ in all but trivial, knife-edge cases. This confirms that a violation of the assumption of a lag-1 model—which on its face would seem likely in a “long memory” process like identity change—leads the HWW formula to return a biased calculation of reliability.

Additional analysis indicates that the direction of this bias is likely negative—that is, that $\rho_B^2$ will underestimate the true $\rho^2$—for many of the identity measures in this paper. As auto-regressive parameters, $\beta_{21}, \beta_{31}$ and $\beta_{32}$ each fall in the range $(0, 1)$, making the denominator of the ratio in expression (2) always positive. Thus $\rho_B^2$ is negatively biased when:

$$- \beta_{31} (\beta_{21} (\beta_{21} - 1) \text{VAR}(Z_1) + \text{VAR}(Z_2)) < 0$$

and thus when

$$(\beta_{21} - \beta_{21}^2) \text{VAR}(Z_1) - \text{VAR}(Z_2) < 0.$$

This expression gives rise to two additional concerns about $\rho_B^2$ with relevance to the reliability of the measurement of identities in this paper:

2. When the lag-1 assumption is violated, reliability is more likely to be underestimated for measures of true scores, such as identities, that exhibit a high degree of stability between Wave 1 and Wave 2.

Inequality (3) is more likely to be satisfied as the quantity $\beta_{21} - \beta_{21}^2$ grows smaller. $\beta_{21}$ is the auto-regressive parameter representing the stability of identity between Wave 1 and Wave 2, and as shown below, $\beta_{21} \in (.5, 1)$ for nearly all of the identities in this paper. In this range, $\beta_{21} - \beta_{21}^2$ declines monotonically in $\beta_{21}$, approaching zero as $\beta_{21}$ approaches one. Inequality (3) therefore tells that when the lag-1 assumption is violated, the HWW formula is more likely to underestimate the reliability of measures of identities that are highly stable between Wave 1 and Wave 2.

3. When the lag-1 assumption is violated, reliability is more likely to be underestimated for measures of true scores with a low variance in the population, and thus measures of dichotomous identities with low prevalence in the population.

Inequality (3) shows that $\rho_B^2$ underestimates $\rho^2$ when $\text{VAR}(Z_1)$ is small relative to $\text{VAR}(Z_2)$. Recall that $\text{VAR}(Z_1) = \text{VAR}(T_1)$, the variance of the true scores in the population. By contrast, $\text{VAR}(Z_2)$ is the variance of individual random disturbances to true scores at Wave 2. Thus ceteris paribus, any measure of true scores that exhibits a low variance about its mean in the population is more likely to be inaccurately determined unreliable by the HWW method. An important implication of this result relevant to this paper is that when true scores are dichotomous—as are the identities measured in this paper, where $T_1 \in \{0, 1\}$
and thus $VAR(T_1) = \mu_T(1 - \mu_T)$, then the calculated reliabilities of measures of variables with a low prevalence in the population (and thus low variance) will be more likely to underestimate the measure’s true reliability when the lag-1 assumption is violated.

A final result identifies additional circumstances under which the underestimation of reliability by $\rho_B^2$ is aggravated:

4. As the lag-2 effect of $T_1$ on $T_3$, grows, $\rho_B^2$ becomes more negatively biased for $\rho^2$.

This is shown by taking the derivative of (2) with respect to $\beta_{31}$, which yields

$$\frac{\partial BIAS[\rho_B^2]}{\partial \beta_{31}} = -\frac{\beta_{21}\beta_{32}(VAR(Z_2) + (\beta_{21}^2 - \beta_{21})VAR(Z_1))}{(VAR(Z_2) + \beta_{21}^2VAR(Z_1) + VAR(E_2)) (\beta_{21}\beta_{32} + \beta_{31})^2},$$

the sign of which is the sign of

$$(\beta_{21} - \beta_{21}^2)VAR(Z_1) - VAR(Z_2),$$

which is negative whenever inequality (3) is satisfied. Thus if $\rho_B^2$ is negatively biased for $\rho^2$, it is more so to the extent that $\beta_{31}$ is large.

Evidence that the HWW Formula May Underestimate Reliability for Identities Measured in the GSS Panel Surveys

To what extent do these concerns apply to the measures of identities in the GSS? While three-wave panels do not provide enough data to assess the degree of bias directly, here I report estimates that strongly suggest that for the measures of identities in this paper, calculations of reliability using the HWW model are vulnerable to concerns about negative bias.

For each of the 27 identities in the paper, I produced estimates of $\beta_{21}$ and $VAR(Z_2)$ with a bivariate regression that generated the estimated model

$$y_{i2} = \hat{\alpha}_2 + \hat{\beta}_{21}y_{i1} + e_{i2}$$

which produces $\hat{\beta}_{21}$ as an estimate of $\beta_{21}$ and the variance of the residuals $VAR(e_{i2})$ as an estimate of $VAR(Z_2)$. I generated estimates of $\beta_{31}$ with a trivariate regression yielding the estimated model

$$y_{i3} = \hat{\alpha}_3 + \hat{\beta}_{32}y_{i2} + \hat{\beta}_{31}y_{i1} + e_{i3}.$$ 

Finally, as noted above, $VAR(Z_1)$ is simply the variance of the true scores in the population at Wave 1; I estimated this quantity using $VAR(Y_1)$.

Appendix Table 5 displays all of these quantities for each of the identities analyzed in this paper as well as the calculated values of the left-hand side of inequality (3). It shows that for every identity, $\beta_{31}$ is estimated to be significantly greater than zero, leading us to be concerned about biased estimates of reliability for these identities. Furthermore, for all but one identity—white racial identity—inequality (3) is satisfied in that $(\beta_{21} - \beta_{21}^2)VAR(Z_1) -$
VAR(Z₂) is estimated to be negative, which means that the HWW formula for reliability is expected to underestimate the reliability of the GSS measures of these identities. In addition, for several identities this negative bias is further aggravated because β₃₁ is estimated to be relatively high. This can be seen in Appendix Figure 1, which plots estimated values of β₃₁ against the left-hand side of inequality (3). We expect ρ² to be negatively biased for ρ² for measures of identities located to the left of zero on the x-axis; the negative bias is greater for identities located toward the top of the figure, which include Latino, lesbian, gay or bisexual, Mexican, and African, Asian or Hispanic national origin.

A final piece of suggestive evidence that the MWW model underestimates the reliability of the GSS measures is found when the MWW estimates of ρ² are used in errors-in-variables (EIV) regression models to estimate the lag-1 stability parameter β₂₁ for each of the identities in this paper. As shown in Appendix Table 6, the EIV models generate estimates of β₂₁ that are near—and in some cases, greater than—one, which unrealistically suggest nearly perfect continuity in most identities over the four-year period due to over-correction for measurement error. Tellingly, among the measures there is a strong correspondence between unrealistically high estimates of β₂₁ and the measures’ vulnerability to underestimated ρ² as indicated by their locations toward the top of the plot in Appendix Figure 1.

Conclusion

This analysis provides reason for general concern that the HWW formula for calculating the reliability of survey measures administered in a three-wave panel—the standard way reliability is calculated for purposes of correcting for measurement error—is vulnerable to concerns about bias when the HWW model’s lag-1 assumption fails to hold. Specifically, this analysis suggests that the HWW model underestimates the reliability of the GSS measures of identity employed in this paper, particularly for identities that are (1) relatively stable from one period to another (β₂₁ is high); (2) relatively rare in the population (VAR(Z₁) is low); and (3) are “long memory” in that current identity is explained by more than one lagged value of identity (β₃₁ is high). Nearly all—26 out of 27—of the identities analyzed in this paper appear to meet these criteria, making it difficult to have confidence in the MWW estimates of their measures’ reliability.

References


Table 5: Parameter estimates for evaluation of the HWW formula for reliability

<table>
<thead>
<tr>
<th>Identity</th>
<th>$\beta_{21}$</th>
<th>se</th>
<th>$VAR(Z_1)$</th>
<th>$VAR(Z_2)$</th>
<th>LHS of ineq (3)</th>
<th>$\beta_{31}$</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Race/ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASIAN/PACIFIC</td>
<td>0.89</td>
<td>0.04</td>
<td>0.04</td>
<td>0.00</td>
<td>-0.001</td>
<td>0.32</td>
<td>0.10</td>
</tr>
<tr>
<td>BLACK</td>
<td>0.95</td>
<td>0.01</td>
<td>0.14</td>
<td>0.01</td>
<td>-0.005</td>
<td>0.35</td>
<td>0.10</td>
</tr>
<tr>
<td>WHITE</td>
<td>0.83</td>
<td>0.03</td>
<td>0.76</td>
<td>0.05</td>
<td>0.058</td>
<td>0.32</td>
<td>0.04</td>
</tr>
<tr>
<td>LATINO</td>
<td>0.94</td>
<td>0.02</td>
<td>0.12</td>
<td>0.01</td>
<td>-0.005</td>
<td>0.46</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>National origin</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFRICAN</td>
<td>0.63</td>
<td>0.03</td>
<td>0.08</td>
<td>0.04</td>
<td>-0.025</td>
<td>0.39</td>
<td>0.05</td>
</tr>
<tr>
<td>AMERICAN INDIAN</td>
<td>0.28</td>
<td>0.05</td>
<td>0.02</td>
<td>0.03</td>
<td>-0.020</td>
<td>0.20</td>
<td>0.05</td>
</tr>
<tr>
<td>AMERICAN ONLY</td>
<td>0.12</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.022</td>
<td>0.16</td>
<td>0.04</td>
</tr>
<tr>
<td>ENGLISH</td>
<td>0.55</td>
<td>0.03</td>
<td>0.11</td>
<td>0.07</td>
<td>-0.047</td>
<td>0.35</td>
<td>0.03</td>
</tr>
<tr>
<td>FRENCH</td>
<td>0.57</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.012</td>
<td>0.32</td>
<td>0.07</td>
</tr>
<tr>
<td>GERMAN</td>
<td>0.65</td>
<td>0.02</td>
<td>0.17</td>
<td>0.08</td>
<td>-0.044</td>
<td>0.35</td>
<td>0.03</td>
</tr>
<tr>
<td>IRISH</td>
<td>0.57</td>
<td>0.03</td>
<td>0.11</td>
<td>0.06</td>
<td>-0.038</td>
<td>0.37</td>
<td>0.03</td>
</tr>
<tr>
<td>ITALIAN</td>
<td>0.77</td>
<td>0.04</td>
<td>0.05</td>
<td>0.02</td>
<td>-0.008</td>
<td>0.41</td>
<td>0.06</td>
</tr>
<tr>
<td>MEXICAN</td>
<td>0.93</td>
<td>0.02</td>
<td>0.06</td>
<td>0.01</td>
<td>-0.004</td>
<td>0.61</td>
<td>0.08</td>
</tr>
<tr>
<td>POLISH</td>
<td>0.67</td>
<td>0.06</td>
<td>0.03</td>
<td>0.01</td>
<td>-0.007</td>
<td>0.41</td>
<td>0.08</td>
</tr>
<tr>
<td>SCOTTISH</td>
<td>0.49</td>
<td>0.05</td>
<td>0.03</td>
<td>0.02</td>
<td>-0.018</td>
<td>0.29</td>
<td>0.05</td>
</tr>
<tr>
<td>AFRICAN, ASIAN OR HISPANIC ORIGIN</td>
<td>0.77</td>
<td>0.02</td>
<td>0.21</td>
<td>0.07</td>
<td>-0.028</td>
<td>0.44</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Sexuality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LESBIAN, GAY, BISEXUAL</td>
<td>0.73</td>
<td>0.07</td>
<td>0.03</td>
<td>0.02</td>
<td>-0.013</td>
<td>0.57</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>Religion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTESTANT</td>
<td>0.78</td>
<td>0.01</td>
<td>0.49</td>
<td>0.10</td>
<td>-0.016</td>
<td>0.38</td>
<td>0.03</td>
</tr>
<tr>
<td>CATHOLIC</td>
<td>0.86</td>
<td>0.01</td>
<td>0.25</td>
<td>0.04</td>
<td>-0.008</td>
<td>0.38</td>
<td>0.04</td>
</tr>
<tr>
<td>JEWISH</td>
<td>0.95</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.002</td>
<td>0.37</td>
<td>0.24</td>
</tr>
<tr>
<td>NO RELIGION</td>
<td>0.67</td>
<td>0.02</td>
<td>0.16</td>
<td>0.08</td>
<td>-0.046</td>
<td>0.35</td>
<td>0.03</td>
</tr>
<tr>
<td>CHRISTIAN</td>
<td>0.29</td>
<td>0.06</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.023</td>
<td>0.22</td>
<td>0.04</td>
</tr>
<tr>
<td>BORN-AGAIN CHRISTIAN</td>
<td>0.70</td>
<td>0.02</td>
<td>0.36</td>
<td>0.12</td>
<td>-0.047</td>
<td>0.36</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOWER CLASS</td>
<td>0.45</td>
<td>0.03</td>
<td>0.07</td>
<td>0.06</td>
<td>-0.040</td>
<td>0.22</td>
<td>0.03</td>
</tr>
<tr>
<td>WORKING CLASS</td>
<td>0.47</td>
<td>0.02</td>
<td>0.45</td>
<td>0.19</td>
<td>-0.075</td>
<td>0.30</td>
<td>0.02</td>
</tr>
<tr>
<td>MIDDLE CLASS</td>
<td>0.52</td>
<td>0.02</td>
<td>0.45</td>
<td>0.18</td>
<td>-0.068</td>
<td>0.31</td>
<td>0.02</td>
</tr>
<tr>
<td>UPPER CLASS</td>
<td>0.48</td>
<td>0.05</td>
<td>0.03</td>
<td>0.02</td>
<td>-0.016</td>
<td>0.20</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Table 6: HWW reliability and lag-1 stability estimates of GSS identity measures

<table>
<thead>
<tr>
<th>Identity</th>
<th>$\rho^2$</th>
<th>$\beta_{21}$</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Race/ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASIAN/PACIFIC</td>
<td>0.93</td>
<td>0.97</td>
<td>0.00</td>
</tr>
<tr>
<td>BLACK</td>
<td>0.96</td>
<td>0.99</td>
<td>0.00</td>
</tr>
<tr>
<td>WHITE</td>
<td>0.89</td>
<td>0.94</td>
<td>0.01</td>
</tr>
<tr>
<td>LATINO</td>
<td>0.92</td>
<td>1.01</td>
<td>0.00</td>
</tr>
<tr>
<td>AFRICAN</td>
<td>0.68</td>
<td>0.96</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>National origin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMERICAN INDIAN</td>
<td>0.29</td>
<td>0.94</td>
<td>0.04</td>
</tr>
<tr>
<td>AMERICAN ONLY</td>
<td>0.19</td>
<td>0.83</td>
<td>0.07</td>
</tr>
<tr>
<td>ENGLISH</td>
<td>0.54</td>
<td>1.04</td>
<td>0.02</td>
</tr>
<tr>
<td>FRENCH</td>
<td>0.60</td>
<td>0.92</td>
<td>0.02</td>
</tr>
<tr>
<td>GERMAN</td>
<td>0.67</td>
<td>0.93</td>
<td>0.01</td>
</tr>
<tr>
<td>IRISH</td>
<td>0.60</td>
<td>0.95</td>
<td>0.02</td>
</tr>
<tr>
<td>ITALIAN</td>
<td>0.80</td>
<td>0.96</td>
<td>0.01</td>
</tr>
<tr>
<td>MEXICAN</td>
<td>0.91</td>
<td>1.02</td>
<td>0.01</td>
</tr>
<tr>
<td>POLISH</td>
<td>0.73</td>
<td>0.93</td>
<td>0.01</td>
</tr>
<tr>
<td>SCOTTISH</td>
<td>0.46</td>
<td>1.04</td>
<td>0.02</td>
</tr>
<tr>
<td>AFRICAN, ASIAN OR HISPANIC ORIGIN</td>
<td>0.77</td>
<td>0.99</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Sexuality</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LESBIAN, GAY, BISEXUAL</td>
<td>0.59</td>
<td>1.12</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>Religion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROTESTANT</td>
<td>0.80</td>
<td>0.97</td>
<td>0.01</td>
</tr>
<tr>
<td>CATHOLIC</td>
<td>0.90</td>
<td>0.96</td>
<td>0.01</td>
</tr>
<tr>
<td>JEWISH</td>
<td>0.95</td>
<td>0.98</td>
<td>0.01</td>
</tr>
<tr>
<td>NO RELIGION</td>
<td>0.70</td>
<td>0.95</td>
<td>0.01</td>
</tr>
<tr>
<td>CHRISTIAN</td>
<td>0.39</td>
<td>0.70</td>
<td>0.03</td>
</tr>
<tr>
<td>BORN-AGAIN CHRISTIAN</td>
<td>0.73</td>
<td>0.96</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Class</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOWER CLASS</td>
<td>0.58</td>
<td>0.84</td>
<td>0.02</td>
</tr>
<tr>
<td>WORKING CLASS</td>
<td>0.54</td>
<td>0.91</td>
<td>0.02</td>
</tr>
<tr>
<td>MIDDLE CLASS</td>
<td>0.53</td>
<td>0.96</td>
<td>0.02</td>
</tr>
<tr>
<td>UPPER CLASS</td>
<td>0.53</td>
<td>0.88</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Figure 1: Parameter estimates for evaluation of the HWW formula for reliability
Appendix Table 7: Regression Estimates of Equation 1
(Output Giving Rise to Figure 2)

<table>
<thead>
<tr>
<th></th>
<th>ASIAN/PACIFIC</th>
<th>BLACK</th>
<th>WHITE</th>
<th>LATINO</th>
<th>AFRICAN, ASIAN, OR HISPANIC ORIGIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identity, Wave 1</td>
<td>9.974***</td>
<td>7.593***</td>
<td>5.240***</td>
<td>8.633***</td>
<td>4.423***</td>
</tr>
<tr>
<td></td>
<td>[0.851]</td>
<td>[0.310]</td>
<td>[0.276]</td>
<td>[0.552]</td>
<td>[0.214]</td>
</tr>
<tr>
<td>Democrat ID strength, Wave 1</td>
<td>0.304</td>
<td>0.966</td>
<td>-0.566*</td>
<td>1.013</td>
<td>0.753**</td>
</tr>
<tr>
<td></td>
<td>[1.167]</td>
<td>[0.522]</td>
<td>[0.237]</td>
<td>[0.867]</td>
<td>[0.244]</td>
</tr>
<tr>
<td>Republican ID strength, Wave 1</td>
<td>0.796</td>
<td>-1.429*</td>
<td>0.2</td>
<td>-0.936</td>
<td>-0.513</td>
</tr>
<tr>
<td></td>
<td>[0.966]</td>
<td>[0.546]</td>
<td>[0.385]</td>
<td>[0.683]</td>
<td>[0.304]</td>
</tr>
<tr>
<td>Liberalism, Wave 1</td>
<td>0.222</td>
<td>-0.198</td>
<td>0.379</td>
<td>-0.723</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>[0.926]</td>
<td>[0.638]</td>
<td>[0.350]</td>
<td>[0.739]</td>
<td>[0.338]</td>
</tr>
<tr>
<td>Conservatism, Wave 1</td>
<td>1.442</td>
<td>0.618</td>
<td>-0.279</td>
<td>-1.799**</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>[1.324]</td>
<td>[0.789]</td>
<td>[0.413]</td>
<td>[0.657]</td>
<td>[0.461]</td>
</tr>
<tr>
<td>Education, Wave 1</td>
<td>0.202**</td>
<td>-0.061</td>
<td>-0.033</td>
<td>-0.162***</td>
<td>-0.075**</td>
</tr>
<tr>
<td></td>
<td>[0.076]</td>
<td>[0.050]</td>
<td>[0.025]</td>
<td>[0.035]</td>
<td>[0.024]</td>
</tr>
<tr>
<td>Age, Wave 1</td>
<td>0.009</td>
<td>-0.017*</td>
<td>0.006</td>
<td>-0.016</td>
<td>-0.018**</td>
</tr>
<tr>
<td></td>
<td>[0.016]</td>
<td>[0.008]</td>
<td>[0.005]</td>
<td>[0.011]</td>
<td>[0.006]</td>
</tr>
<tr>
<td>Sex</td>
<td>0.677</td>
<td>0.19</td>
<td>0.041</td>
<td>0.356</td>
<td>-0.345**</td>
</tr>
<tr>
<td></td>
<td>[0.367]</td>
<td>[0.291]</td>
<td>[0.220]</td>
<td>[0.350]</td>
<td>[0.128]</td>
</tr>
<tr>
<td>Wave 1 conducted in 2008</td>
<td>1.738*</td>
<td>0.940*</td>
<td>-0.667*</td>
<td>-0.793</td>
<td>-0.062</td>
</tr>
<tr>
<td></td>
<td>[0.675]</td>
<td>[0.413]</td>
<td>[0.255]</td>
<td>[0.449]</td>
<td>[0.210]</td>
</tr>
<tr>
<td>Wave 1 conducted in 2010</td>
<td>1.608*</td>
<td>1.101*</td>
<td>-0.802**</td>
<td>-1.717**</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>[0.660]</td>
<td>[0.493]</td>
<td>[0.238]</td>
<td>[0.545]</td>
<td>[0.193]</td>
</tr>
<tr>
<td>Intercept</td>
<td>-13.124***</td>
<td>-4.598***</td>
<td>-0.885</td>
<td>-1.732</td>
<td>-0.903</td>
</tr>
<tr>
<td></td>
<td>[2.208]</td>
<td>[1.112]</td>
<td>[0.622]</td>
<td>[1.035]</td>
<td>[0.539]</td>
</tr>
</tbody>
</table>

N = 3,856. * p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests)
<table>
<thead>
<tr>
<th></th>
<th>LGB</th>
<th>PROTEST-ANT</th>
<th>CATHOLIC</th>
<th>JEWISH</th>
<th>NONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identity, Wave 1</td>
<td>6.222***</td>
<td>3.954***</td>
<td>5.715***</td>
<td>8.871***</td>
<td>3.333***</td>
</tr>
<tr>
<td></td>
<td>[0.480]</td>
<td>[0.120]</td>
<td>[0.193]</td>
<td>[0.873]</td>
<td>[0.129]</td>
</tr>
<tr>
<td>Democrat ID strength, Wave 1</td>
<td>-0.128</td>
<td>0.094</td>
<td>-0.094</td>
<td>1.427</td>
<td>-0.759***</td>
</tr>
<tr>
<td></td>
<td>[0.508]</td>
<td>[0.204]</td>
<td>[0.322]</td>
<td>[0.895]</td>
<td>[0.150]</td>
</tr>
<tr>
<td>Republican ID strength, Wave 1</td>
<td>-1.192</td>
<td>0.29</td>
<td>0.288</td>
<td>1.476</td>
<td>-1.337***</td>
</tr>
<tr>
<td></td>
<td>[0.746]</td>
<td>[0.214]</td>
<td>[0.359]</td>
<td>[0.827]</td>
<td>[0.280]</td>
</tr>
<tr>
<td>Liberalism, Wave 1</td>
<td>0.746</td>
<td>-0.426</td>
<td>-0.275</td>
<td>0.573</td>
<td>0.907***</td>
</tr>
<tr>
<td></td>
<td>[0.431]</td>
<td>[0.225]</td>
<td>[0.319]</td>
<td>[0.624]</td>
<td>[0.222]</td>
</tr>
<tr>
<td>Conservatism, Wave 1</td>
<td>-0.728</td>
<td>0.429</td>
<td>-0.207</td>
<td>-1.535</td>
<td>-0.239</td>
</tr>
<tr>
<td></td>
<td>[0.926]</td>
<td>[0.225]</td>
<td>[0.391]</td>
<td>[2.097]</td>
<td>[0.266]</td>
</tr>
<tr>
<td>Education, Wave 1</td>
<td>-0.175*</td>
<td>-0.037</td>
<td>-0.027</td>
<td>0.134**</td>
<td>0.077***</td>
</tr>
<tr>
<td></td>
<td>[0.082]</td>
<td>[0.020]</td>
<td>[0.036]</td>
<td>[0.045]</td>
<td>[0.021]</td>
</tr>
<tr>
<td>Age, Wave 1</td>
<td>-0.009</td>
<td>0.015***</td>
<td>0.01</td>
<td>-0.035*</td>
<td>-0.020***</td>
</tr>
<tr>
<td></td>
<td>[0.013]</td>
<td>[0.004]</td>
<td>[0.007]</td>
<td>[0.017]</td>
<td>[0.004]</td>
</tr>
<tr>
<td>Sex</td>
<td>0.201</td>
<td>0.174</td>
<td>0.037</td>
<td>0.752</td>
<td>-0.616***</td>
</tr>
<tr>
<td></td>
<td>[0.364]</td>
<td>[0.121]</td>
<td>[0.186]</td>
<td>[1.028]</td>
<td>[0.116]</td>
</tr>
<tr>
<td>Wave 1 conducted in 2008</td>
<td>0.145</td>
<td>0.037</td>
<td>1.053</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.138]</td>
<td>[0.224]</td>
<td>[0.645]</td>
<td>[0.184]</td>
<td></td>
</tr>
<tr>
<td>Wave 1 conducted in 2010</td>
<td>0.176</td>
<td>0.052</td>
<td>-0.248</td>
<td>0.426</td>
<td>0.448*</td>
</tr>
<tr>
<td></td>
<td>[0.384]</td>
<td>[0.126]</td>
<td>[0.186]</td>
<td>[0.518]</td>
<td>[0.176]</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.057</td>
<td>-2.720***</td>
<td>-3.878***</td>
<td>-9.325***</td>
<td>-1.634***</td>
</tr>
<tr>
<td></td>
<td>[1.311]</td>
<td>[0.429]</td>
<td>[0.676]</td>
<td>[1.935]</td>
<td>[0.397]</td>
</tr>
</tbody>
</table>

\( N = 3,856 \) except for LGB, where \( N = 2,235 \). * \( p<0.05 \), ** \( p<0.01 \), *** \( p<0.001 \) (two-tailed tests)
<table>
<thead>
<tr>
<th>Identity, Wave 1</th>
<th>CHRIS-TIAN</th>
<th>BORN- AGAIN</th>
<th>LOWER CLASS</th>
<th>WORK- ING CLASS</th>
<th>MIDDLE CLASS</th>
<th>UPPER CLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.760***</td>
<td>3.236***</td>
<td>2.528***</td>
<td>1.924***</td>
<td>1.975***</td>
<td>3.490***</td>
</tr>
<tr>
<td></td>
<td>[0.262]</td>
<td>[0.102]</td>
<td>[0.159]</td>
<td>[0.084]</td>
<td>[0.111]</td>
<td>[0.303]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Democrat ID strength, Wave 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Democrat ID strength, Wave 1</td>
<td>0.719**</td>
<td>0.185</td>
<td>-0.377</td>
<td>0.281*</td>
<td>-0.173</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>[0.259]</td>
<td>[0.163]</td>
<td>[0.221]</td>
<td>[0.132]</td>
<td>[0.142]</td>
<td>[0.413]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Republican ID strength, Wave 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Republican ID strength, Wave 1</td>
<td>0.464</td>
<td>-0.161</td>
<td>-0.829**</td>
<td>-0.138</td>
<td>0.336*</td>
<td>0.193</td>
</tr>
<tr>
<td></td>
<td>[0.388]</td>
<td>[0.174]</td>
<td>[0.311]</td>
<td>[0.153]</td>
<td>[0.162]</td>
<td>[0.476]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Liberalism, Wave 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Liberalism, Wave 1</td>
<td>-1.188**</td>
<td>-0.483*</td>
<td>0.201</td>
<td>-0.699***</td>
<td>0.667***</td>
<td>0.168</td>
</tr>
<tr>
<td></td>
<td>[0.432]</td>
<td>[0.205]</td>
<td>[0.255]</td>
<td>[0.187]</td>
<td>[0.178]</td>
<td>[0.587]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conservatism, Wave 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservatism, Wave 1</td>
<td>-0.68</td>
<td>0.818***</td>
<td>-0.032</td>
<td>-0.256</td>
<td>0.294</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>[0.439]</td>
<td>[0.179]</td>
<td>[0.346]</td>
<td>[0.198]</td>
<td>[0.178]</td>
<td>[0.579]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education, Wave 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Education, Wave 1</td>
<td>0</td>
<td>-0.105***</td>
<td>-0.161***</td>
<td>-0.132***</td>
<td>0.179***</td>
<td>0.146***</td>
</tr>
<tr>
<td></td>
<td>[0.032]</td>
<td>[0.021]</td>
<td>[0.027]</td>
<td>[0.018]</td>
<td>[0.021]</td>
<td>[0.041]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age, Wave 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, Wave 1</td>
<td>-0.035***</td>
<td>0.007*</td>
<td>-0.008*</td>
<td>-0.018***</td>
<td>0.018***</td>
<td>0.017*</td>
</tr>
<tr>
<td></td>
<td>[0.007]</td>
<td>[0.003]</td>
<td>[0.004]</td>
<td>[0.003]</td>
<td>[0.003]</td>
<td>[0.007]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>0.461*</td>
<td>0.387***</td>
<td>0.223</td>
<td>-0.047</td>
<td>-0.028</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>[0.200]</td>
<td>[0.101]</td>
<td>[0.188]</td>
<td>[0.104]</td>
<td>[0.093]</td>
<td>[0.268]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wave 1 conducted in 2008</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave 1 conducted in 2008</td>
<td>-0.24</td>
<td>0.174</td>
<td>0.24</td>
<td>-0.014</td>
<td>0.004</td>
<td>-0.413</td>
</tr>
<tr>
<td></td>
<td>[0.298]</td>
<td>[0.133]</td>
<td>[0.186]</td>
<td>[0.104]</td>
<td>[0.108]</td>
<td>[0.290]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wave 1 conducted in 2010</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave 1 conducted in 2010</td>
<td>0.13</td>
<td>-0.014</td>
<td>0.188</td>
<td>-0.073</td>
<td>0.044</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>[0.253]</td>
<td>[0.129]</td>
<td>[0.193]</td>
<td>[0.112]</td>
<td>[0.116]</td>
<td>[0.334]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intercept</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.690***</td>
<td>-1.451***</td>
<td>-0.705</td>
<td>1.647***</td>
<td>-4.560***</td>
<td>-7.175***</td>
</tr>
<tr>
<td></td>
<td>[0.604]</td>
<td>[0.418]</td>
<td>[0.533]</td>
<td>[0.306]</td>
<td>[0.368]</td>
<td>[0.860]</td>
</tr>
</tbody>
</table>

N = 3,856. * p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests).
Appendix Table 8: Placebo Tests

<table>
<thead>
<tr>
<th></th>
<th>predictive margins for ...Wave 1</th>
<th>predictive margins for ...Wave 1</th>
<th>difference</th>
<th>p-value of difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>liberal Democrats</td>
<td>conserv Republicans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Astrological sign</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aries</td>
<td>0.082</td>
<td>0.081</td>
<td>-0.001</td>
<td>0.781</td>
</tr>
<tr>
<td>taurus</td>
<td>0.080</td>
<td>0.080</td>
<td>0.000</td>
<td>0.958</td>
</tr>
<tr>
<td>gemini</td>
<td>0.087</td>
<td>0.085</td>
<td>-0.001</td>
<td>0.712</td>
</tr>
<tr>
<td>cancer</td>
<td>0.085</td>
<td>0.087</td>
<td>0.002</td>
<td>0.689</td>
</tr>
<tr>
<td>leo</td>
<td>0.077</td>
<td>0.083</td>
<td>0.006</td>
<td>0.224</td>
</tr>
<tr>
<td>virgo</td>
<td>0.097</td>
<td>0.094</td>
<td>-0.002</td>
<td>0.515</td>
</tr>
<tr>
<td>libra</td>
<td>0.085</td>
<td>0.090</td>
<td>0.005</td>
<td>0.088</td>
</tr>
<tr>
<td>scorpio</td>
<td>0.075</td>
<td>0.073</td>
<td>-0.002</td>
<td>0.533</td>
</tr>
<tr>
<td>sagittarius</td>
<td>0.077</td>
<td>0.073</td>
<td>-0.004</td>
<td>0.327</td>
</tr>
<tr>
<td>capricorn</td>
<td>0.082</td>
<td>0.079</td>
<td>-0.003</td>
<td>0.385</td>
</tr>
<tr>
<td>aquarius</td>
<td>0.088</td>
<td>0.088</td>
<td>-0.001</td>
<td>0.859</td>
</tr>
<tr>
<td>pisces</td>
<td>0.086</td>
<td>0.088</td>
<td>0.001</td>
<td>0.286</td>
</tr>
<tr>
<td>Region of residence at age 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>foreign</td>
<td>0.085</td>
<td>0.078</td>
<td>-0.006</td>
<td>0.275</td>
</tr>
<tr>
<td>NEW ENGLAND</td>
<td>0.040</td>
<td>0.039</td>
<td>0.000</td>
<td>0.842</td>
</tr>
<tr>
<td>MIDDLE ATLANTIC</td>
<td>0.158</td>
<td>0.148</td>
<td>-0.010</td>
<td>0.142</td>
</tr>
<tr>
<td>E. NOR. CENTRAL</td>
<td>0.187</td>
<td>0.189</td>
<td>0.002</td>
<td>0.658</td>
</tr>
<tr>
<td>W. NOR. CENTRAL</td>
<td>0.068</td>
<td>0.062</td>
<td>-0.005</td>
<td>0.067</td>
</tr>
<tr>
<td>SOUTH ATLANTIC</td>
<td>0.153</td>
<td>0.159</td>
<td>0.006</td>
<td>0.495</td>
</tr>
<tr>
<td>E. SOU. CENTRAL</td>
<td>0.052</td>
<td>0.058</td>
<td>0.006</td>
<td>0.249</td>
</tr>
<tr>
<td>W. SOU. CENTRAL</td>
<td>0.082</td>
<td>0.084</td>
<td>0.002</td>
<td>0.650</td>
</tr>
<tr>
<td>mountain</td>
<td>0.044</td>
<td>0.045</td>
<td>0.002</td>
<td>0.538</td>
</tr>
<tr>
<td>pacific</td>
<td>0.131</td>
<td>0.133</td>
<td>0.003</td>
<td>0.551</td>
</tr>
<tr>
<td>Additional variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex (1 male, 2 female)</td>
<td>1.55</td>
<td>1.55</td>
<td>-0.004</td>
<td>0.494</td>
</tr>
<tr>
<td>year of birth</td>
<td>1962.5</td>
<td>1962.2</td>
<td>-0.311</td>
<td>0.049</td>
</tr>
<tr>
<td>mother's education (years)</td>
<td>11.86</td>
<td>11.78</td>
<td>-0.082</td>
<td>0.428</td>
</tr>
<tr>
<td>father's education (years)</td>
<td>11.68</td>
<td>11.73</td>
<td>0.054</td>
<td>0.645</td>
</tr>
<tr>
<td>number of siblings</td>
<td>3.55</td>
<td>3.45</td>
<td>-0.106</td>
<td>0.176</td>
</tr>
</tbody>
</table>
### Appendix Table 9: Regression Estimates Giving Rise to Table 2

<table>
<thead>
<tr>
<th></th>
<th>Wave 1 to Wave 3 identity shifts</th>
<th>$t - 1$ to $t$ identity shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>logit OLS</td>
<td>logit OLS</td>
</tr>
<tr>
<td>lagged identification as a $j$</td>
<td>3.801*** 0.690***</td>
<td>3.943*** 0.710***</td>
</tr>
<tr>
<td></td>
<td>[0.042] [0.005]</td>
<td>[0.033] [0.004]</td>
</tr>
<tr>
<td>dissimilarity index, $j$</td>
<td>3.797 0.09</td>
<td>-1.237 0.032</td>
</tr>
<tr>
<td></td>
<td>[6.839] [0.139]</td>
<td>[4.860] [0.081]</td>
</tr>
<tr>
<td>lagged liberalism</td>
<td>-0.014 -0.007</td>
<td>0.007 -0.003</td>
</tr>
<tr>
<td></td>
<td>[0.040] [0.004]</td>
<td>[0.025] [0.002]</td>
</tr>
<tr>
<td>lagged conservatism</td>
<td>0.077* 0.012***</td>
<td>0.095*** 0.012***</td>
</tr>
<tr>
<td></td>
<td>[0.038] [0.003]</td>
<td>[0.027] [0.002]</td>
</tr>
<tr>
<td>lagged Democrat</td>
<td>0.033 0.000</td>
<td>0.090*** 0.005**</td>
</tr>
<tr>
<td></td>
<td>[0.032] [0.003]</td>
<td>[0.019] [0.002]</td>
</tr>
<tr>
<td>lagged Republican</td>
<td>-0.179*** 0.001</td>
<td>-0.122*** 0.007***</td>
</tr>
<tr>
<td></td>
<td>[0.037] [0.003]</td>
<td>[0.025] [0.002]</td>
</tr>
<tr>
<td>dissimilarity index, $j$ x…</td>
<td></td>
<td></td>
</tr>
<tr>
<td>…lagged liberalism</td>
<td>-0.691** -0.057**</td>
<td>-0.401* -0.035**</td>
</tr>
<tr>
<td></td>
<td>[0.236] [0.020]</td>
<td>[0.160] [0.013]</td>
</tr>
<tr>
<td>…lagged conservatism</td>
<td>0.759** 0.056**</td>
<td>0.544** 0.043**</td>
</tr>
<tr>
<td></td>
<td>[0.283] [0.019]</td>
<td>[0.200] [0.013]</td>
</tr>
<tr>
<td>…lagged Democrat</td>
<td>-0.764*** -0.050**</td>
<td>-0.554*** -0.032**</td>
</tr>
<tr>
<td></td>
<td>[0.194] [0.016]</td>
<td>[0.133] [0.011]</td>
</tr>
<tr>
<td>…lagged Republican</td>
<td>1.755*** 0.073***</td>
<td>2.436*** 0.112***</td>
</tr>
<tr>
<td></td>
<td>[0.293] [0.018]</td>
<td>[0.202] [0.011]</td>
</tr>
<tr>
<td>age</td>
<td>-0.001* 0.000</td>
<td>-0.002*** -0.000***</td>
</tr>
<tr>
<td></td>
<td>[0.001] [0.000]</td>
<td>[0.000] [0.000]</td>
</tr>
<tr>
<td>education</td>
<td>-0.020*** -0.002***</td>
<td>-0.022*** -0.002***</td>
</tr>
<tr>
<td></td>
<td>[0.003] [0.000]</td>
<td>[0.002] [0.000]</td>
</tr>
<tr>
<td>sex</td>
<td>0.03 0.003*</td>
<td>0.015 0.001</td>
</tr>
<tr>
<td></td>
<td>[0.019] [0.001]</td>
<td>[0.012] [0.001]</td>
</tr>
<tr>
<td>Wave 1 in 2008</td>
<td>0.005 0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.024] [0.002]</td>
<td></td>
</tr>
<tr>
<td>Wave 1 in 2010</td>
<td>-0.003 0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.023] [0.002]</td>
<td></td>
</tr>
<tr>
<td>year of survey: 2010</td>
<td></td>
<td>-0.002 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.020] [0.001]</td>
</tr>
<tr>
<td>year of survey: 2012</td>
<td>0.03 0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.020] [0.001]</td>
<td></td>
</tr>
<tr>
<td>year of survey: 2014</td>
<td>-0.024 -0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.022] [0.001]</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.423*** 0.036*</td>
<td>-3.899*** 0.033***</td>
</tr>
<tr>
<td></td>
<td>[0.725] [0.015]</td>
<td>[0.527] [0.009]</td>
</tr>
<tr>
<td>N</td>
<td>3,856 4,637</td>
<td></td>
</tr>
<tr>
<td>p, joint significance of shaded coefficients above</td>
<td>&lt;.001 &lt;.001</td>
<td>&lt;.001 &lt;.001</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests).
## Model Accounting for Selection

<table>
<thead>
<tr>
<th></th>
<th>Logit</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liberalism</td>
<td>-0.067***</td>
<td>-0.011***</td>
</tr>
<tr>
<td></td>
<td>[0.018]</td>
<td>[0.003]</td>
</tr>
<tr>
<td>Conservatism</td>
<td>0.036</td>
<td>0.017***</td>
</tr>
<tr>
<td></td>
<td>[0.019]</td>
<td>[0.003]</td>
</tr>
<tr>
<td>Democrat</td>
<td>0.132***</td>
<td>0.019***</td>
</tr>
<tr>
<td></td>
<td>[0.018]</td>
<td>[0.003]</td>
</tr>
<tr>
<td>Republican</td>
<td>-0.116***</td>
<td>0.017***</td>
</tr>
<tr>
<td></td>
<td>[0.023]</td>
<td>[0.003]</td>
</tr>
<tr>
<td>Dissimilarity index, j</td>
<td>-22.962</td>
<td>-3.61</td>
</tr>
<tr>
<td></td>
<td>[123.042]</td>
<td>[16.657]</td>
</tr>
<tr>
<td>Dissimilarity index, j x...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...Liberalism</td>
<td>0.094</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>[0.171]</td>
<td>[0.027]</td>
</tr>
<tr>
<td>...Conservatism</td>
<td>0.926***</td>
<td>0.115***</td>
</tr>
<tr>
<td></td>
<td>[0.198]</td>
<td>[0.025]</td>
</tr>
<tr>
<td>...Democrat</td>
<td>-0.263</td>
<td>-0.037</td>
</tr>
<tr>
<td></td>
<td>[0.166]</td>
<td>[0.026]</td>
</tr>
<tr>
<td>...Republican</td>
<td>2.282***</td>
<td>0.208***</td>
</tr>
<tr>
<td></td>
<td>[0.242]</td>
<td>[0.027]</td>
</tr>
<tr>
<td>Age</td>
<td>-0.002***</td>
<td>-0.000***</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
</tr>
<tr>
<td>Education</td>
<td>-0.030***</td>
<td>-0.004***</td>
</tr>
<tr>
<td></td>
<td>[0.002]</td>
<td>[0.000]</td>
</tr>
<tr>
<td>Sex</td>
<td>0.040***</td>
<td>0.005***</td>
</tr>
<tr>
<td></td>
<td>[0.011]</td>
<td>[0.002]</td>
</tr>
<tr>
<td>Year of survey</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>[0.006]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>Ever Democrat</td>
<td>-4.399</td>
<td>-0.616</td>
</tr>
<tr>
<td></td>
<td>[10.637]</td>
<td>[1.807]</td>
</tr>
<tr>
<td>Ever Republican</td>
<td>-0.291</td>
<td>0.930</td>
</tr>
<tr>
<td></td>
<td>[10.593]</td>
<td>[1.826]</td>
</tr>
<tr>
<td>Ever Liberal</td>
<td>-1.107</td>
<td>0.454</td>
</tr>
<tr>
<td></td>
<td>[8.972]</td>
<td>[1.525]</td>
</tr>
<tr>
<td>Ever Conservative</td>
<td>0.512</td>
<td>-0.439</td>
</tr>
<tr>
<td></td>
<td>[8.690]</td>
<td>[1.510]</td>
</tr>
</tbody>
</table>

(continued)
year x…

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ever Democrat</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>[0.005]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>ever Republican</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>[0.005]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>ever liberal</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>[0.004]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>ever conservative</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>[0.004]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>dissimilarity index, j</td>
<td>0.011</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>[0.061]</td>
<td>[0.008]</td>
</tr>
</tbody>
</table>

dissimilarity index, j x year x…

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ever Democrat</td>
<td>7.692</td>
<td>0.343</td>
</tr>
<tr>
<td></td>
<td>[113.760]</td>
<td>[15.234]</td>
</tr>
<tr>
<td>ever Republican</td>
<td>74.331</td>
<td>9.981</td>
</tr>
<tr>
<td></td>
<td>[114.217]</td>
<td>[15.444]</td>
</tr>
<tr>
<td>ever liberal</td>
<td>61.983</td>
<td>10.167</td>
</tr>
<tr>
<td></td>
<td>[90.992]</td>
<td>[13.185]</td>
</tr>
<tr>
<td>ever conservative</td>
<td>-32.714</td>
<td>-4.417</td>
</tr>
<tr>
<td></td>
<td>[88.203]</td>
<td>[12.681]</td>
</tr>
</tbody>
</table>

dissimilarity index, j x year x…

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ever Democrat</td>
<td>-0.004</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>[0.057]</td>
<td>[0.008]</td>
</tr>
<tr>
<td>ever Republican</td>
<td>-0.037</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>[0.057]</td>
<td>[0.008]</td>
</tr>
<tr>
<td>ever liberal</td>
<td>-0.031</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>[0.045]</td>
<td>[0.007]</td>
</tr>
<tr>
<td>ever conservative</td>
<td>0.016</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>[0.044]</td>
<td>[0.006]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>3,872</th>
</tr>
</thead>
<tbody>
<tr>
<td>p, joint significance of shaded coefficients above</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001 (two-tailed tests).