

**A Method of Linking Surveys Using Affective “Signatures” with an Application to
Racial/Ethnic Groups in the U.S.**

April 28, 2008

Marisa A. Abrajano
Keith T. Poole
Department of Political Science
University of California, San Diego

Abstract

This paper addresses a concern often faced by social scientists who study subgroups within a given population, as they are frequently limited in the scope and breadth of their research questions due to the quality of available survey data (i.e. inadequate sample size or lack of comprehensive questions). To address this problem, we develop a procedure for linking respondents from different surveys based on their internal (subjective) utility for political stimuli, which we capture by using an individual's responses to a set of feeling thermometer questions. Feeling thermometer questions, as demonstrated in previous research, are an accurate measure of an individual's *subjective utility* because they are measures of *affect*. We apply this technique to the 2004 National Annenberg Election survey and the 2004 American National Election Studies survey. Linking survey respondents based on their thermometer scores not only recovers the distributions on group demographics such as race/ethnicity, gender, and education but it also recovers the distributions of these groups' preferences across a wide array of issues and policies as well.

Introduction

The main problem that we address in this paper is of concern to social scientists in general but particularly for those whose area of research focuses on sub-groups within a given population (e.g., combinations of gender, race, and ethnicity). What happens when researchers are interested in understanding sub-group behavior, but data to do so is limited? What happens when researchers interested in studying sub-groups have adequate sample sizes but are lacking in the appropriate questions to test their arguments?

For example, consider the 2004 National Annenberg Election Survey, which is a companion survey to the 2000 survey conducted by the Annenberg team of scholars at the University of Pennsylvania (Romer et al., 2006). This survey is highly desirable for most social scientists, as it contains more than 150 questions pertaining to an individual's political attitudes, behaviors and perceptions. Moreover, it interviews an extremely large number of individuals, more than 80,000; thus the number of sub-group populations captured is also sizeable – approximately 5,000 Latinos and 7,000 Blacks. But a major drawback to this survey is that key questions pertaining to issue attitudes are simply yes/no. More fine grained measures such as 7-point issue scale questions are not available in the Annenberg survey.

On the other hand, the preeminent survey on American political attitudes and behavior, the American National Election Survey (ANES), contains numerous 7-point issue scales, as well as a number of detailed issue questions and feeling thermometer questions. The main problem with this data is its small sample size ($N = 1,212$), and in turn, a limited sample of sub-group populations. For instance, the 2004 NES survey interviewed 81 Latinos, 180 Blacks, and 876 Whites.

This paper resolves these two problems by developing a method that combines the desirable qualities of separate surveys by linking survey respondents based on their reported internal (subjective) utility for political stimuli. In our application we combine the large sample size of the Annenberg survey and the detailed issue questions provided by the NES survey. By doing so, we are able to overcome the “small sample size” and “detailed question” problems encountered by researchers studying sub-group populations.

More specifically, our method links survey respondents from the NES and the Annenberg surveys based on their responses to a set of 10 feeling thermometer questions asked about politicians/parties that the surveys have in common. We use feeling thermometer questions since, as we will discuss in more detail below, they are likely to be the most accurate indicator of an individual’s *subjective utility* because they are measures of *affect*. Our method is quite simple -- we pair each Annenberg respondent to the NES respondent who most closely resembles his/her set of thermometer scores. This means that individuals are being linked to one another based on a vector of candidate affect. When these individuals are paired based on these vectors of affect, we find that different groups (e.g. race/ethnicity, gender) possess distinctive affective signatures or patterns. Thus we can recover group characteristics quite accurately based solely on their affective signatures. We primarily focus on sub-group populations based on race and ethnicity, though this technique can certainly be applied to other sub-group populations.

We make several contributions to the existing literature. First, this method allows researchers interested in studying sub-group populations to attain a more in-depth and nuanced understanding of their political attitudes and preferences. Moreover, our technique makes it possible to recover a larger sample of groups, particularly for ethnic and racial minority groups

that are often underrepresented in cross-national surveys. The same is true for scholars who study state and local politics, as they often face sampling issues in national surveys.

In the next section, we discuss the literature on the various strands of research focusing on feeling thermometer questions. We then present our data and methods, followed by a discussion of our main findings. In the final section, we summarize our results and discuss further avenues for research.

Using Feeling Thermometers to Capture Affective Signatures

Feeling thermometer questions were originally developed for group evaluations by Aage Clausen and were first used in the American National Election Survey (ANES) in 1964. The group feeling thermometer questions were for Protestants, Catholics, Jews, Blacks, Whites, Southerners, big business, labor unions, liberals, and conservatives. Herbert Weisberg and Jerrold Rusk added feeling thermometer questions for individuals (either prominent politicians or candidates) in the 1968 ANES. A “feeling” thermometer asks respondents to respond to a set of political stimuli (individuals or groups) based on their subjective views of warmth towards each of these groups/politicians. The thermometer ranges from 0 to 100 degrees with 100 indicating warm and very favorable feeling, 50 indicating neutrality towards the group/politician, and 0 indicating that the respondent feels cold and very unfavorable towards the group/politician.

Since its inception in the 1964 ANES, feeling thermometers have remained a constant not only in this preeminent survey on American political behavior and attitudes, but also in other fields (e.g. psychology). Feeling thermometers emerged as a standard tool in survey-based political research for several reasons. As Weisberg and Rusk (1970) note, feeling

thermometers allow respondents to evaluate candidates on “those dimensions which come naturally to them, [those] which are [their] normal guidelines for thinking about candidates.” Since feeling thermometers do not impose any types of frames on respondents, they can tap into those evaluative dimensions that they consider most important to them. Feeling thermometers have also been shown to accurately capture an individual’s affective sentiments (Weisberg and Rusk 1970). As such, we expect the responses from a set of feeling thermometers to be an excellent proxy for an individual’s internal subjective utility.

We assume that an individual’s reported “feeling” for a politician or group is generated by the individual’s subjective utility function over the relevant issue/policy space as well as all non-policy attributes related to the individuals' psychological makeup. That is:

$$\text{Thermometer Score} = f[U_i(X, Z)]$$

where f is a simple mapping function that takes the subjective utility and translates it into the 0 – 100 scale, U_i is the utility function for individual i , X are the relevant issue/policy dimensions, and Z are dimensions such as "likeability", "leadership", and for racial/ethnic minorities, possibly a dimension pertaining to "ethnic group identity." The combination of X and Z is in part determined by the standard demographic characteristics that we are concerned with as social scientists. With respect to the X dimensions, we assume that, consistent with a standard spatial model of choice (Downs, 1957; Enelow and Hinich, 1984), the individual has an ideal point (or most preferred point) on each dimension. The Z dimensions are best thought of as *valence* dimensions (Enelow and Hinich, 1984); that is, either the politician/group has the attribute or not – likeable, not likeable; honest, corrupt; etc. Here we assume that individuals prefer the positive side of the valence dimension and politicians/groups that have the attributes have higher subjective utility.

When we pair respondents based upon sets of thermometers we are actually linking people with similar internal utility functions. If this logic holds, pairing respondents based on their feeling thermometer scores should be more accurate than pairing respondents based on demographics if what we are interested in is distributions of sub-populations over political issues.

Feeling thermometer questions sparked a great deal of research in the 1970s and early 1980s -- Weisberg and Rusk (1970), Wang, et al. (1975), Rabinowitz (1976), Cahoon, et al. (1978), Poole and Rosenthal (1984), Poole (1984, 1990) – with the main focus on modeling the latent dimensions underlying the thermometers as well as testing theories of spatial voting. Other scholars, such as Knight (1984), Giles and Evans (1986) and Wilcox, et al. (1989), explored the reasons behind the variations in feeling thermometer responses, and cautioned in the interpretation of the responses to these feeling thermometer questions. This is because individuals can vary in their interpretation of the 0-100 scale; while some may use the entire scale, others may restrict themselves to only a certain part of the scale (Wilcox, et al. 1989). As such, Knight (1984) recommends adjusting thermometer ratings for groups by subtracting the average score for an individual's set of responses from the score for the group of interest. Giles and Evans (1986) also suggest accounting for both the mean and standard deviation of the thermometer scores. However, since this burst of activity thermometers have been relatively understudied. In part, we hope to reintroduce the usefulness of feeling thermometers to researchers who are not only interested in understanding voter attitudes and perceptions, but also for those who wish to study new methodological techniques.

While our method is somewhat similar in spirit to the increasingly popular method known as matching, our procedure differs from this technique in several major ways.

Essentially, what matching seeks to do is to compare individuals in a treatment group with similar individuals in a comparison or control group. The logic is that, after matching individuals from both groups based on specific background characteristics, then any difference that arises between these two groups can be attributed to the treatment being applied. For example, political scientists studying political behavior have long been interested in understanding whether voter mobilization efforts, such as being contacted by a campaign or receiving mailers, increase turnout (Arceneaux et al., 2006, Imai, 2005, Gerber and Green, 2000, 2005). One way of assessing the impact of voter contact on turnout is to match a treated group of individuals (those who were asked to vote) with a control group of individuals (those who were not asked to vote) based on background variables such as their age, levels of education, income, etc. Matching on these demographic characteristics would control for other factors that may influence their rates of turnout and thus, any differences in turnout could be attributed to mobilization efforts. In this regard, this procedure mirrors a controlled trial design experiment used by medical researchers.

While our method also “matches” individuals based on shared characteristics, which in our case is their responses to a set of feeling thermometer questions, our goal is not to identify a specific causal mechanism between a treated and untreated group. Instead, we “link” individuals based on their affective signatures as a way to predict their political attitudes and opinions. Thus this procedure is particularly useful for researchers who are interested in understanding the issue attitudes and viewpoints of subgroups within the U.S. who are oftentimes under-sampled in many of the major public opinion surveys.

Another factor that distinguishes our technique from matching is that most, if not all, of the research using matching methods have done so by matching on observed data such as an

individual's background characteristics (e.g. Greiner, 2006, Nickerson, 2005). Greiner (2006) examines a variety of civil rights legislation (e.g. employment discrimination, death penalty, and redistricting) by matching on the group's covariates. Likewise, Imai (2004) also matches individuals based on their background characteristics in his reanalysis of Gerber and Green's well-known 2000 experiment on voter mobilization. On the same topic, Arceneaux et al. (2006) use matching methods in a voter mobilization experiment, and matches on covariates pertaining to an individual's age, gender, household size, whether or not he/she is a newly registered voter and past voting rates. Finally, Ho et al. (2007) developed a software application in R (*MatchIt*) that involves nonparametric preprocessing of the data, and matches on the control and treatment groups' background characteristics. While nonparametric preprocessing is desirable because it can reduce bias and inefficiency, matching on a group's background characteristics is not the only observed data available to researchers. Thus, we expect that pairing survey respondents on observed variables beyond background characteristics is a realistic assumption.

Data and Methods

As we discussed earlier, we use two datasets—the 2004 National Annenberg Election Survey and the 2004 NES. The 2004 NES interviewed 1,212 individuals. Respondents were asked to give thermometer ratings to fourteen political figures; George W. Bush, John Kerry, Ralph Nader, Richard Cheney, John Edwards, Laura Bush, Hillary Clinton, Bill Clinton, Colin Powell, John McCain, John Ashcroft, the Democratic Party, the Republican Party and Ronald Reagan.

The 2004 National Annenberg Election survey was designed as a rolling cross-sectional that was in the field from October 27, 2003 to November 16, 2004. The survey was conducted

by Daniel Romer, Kate Kenski, Kenneth Winneg, Christopher Adasiewicz and Kathleen Hall Jamieson of the Annenberg Public Policy Center of the University of Pennsylvania (Romer et al., 2006). There were 81,422 individuals who were randomly selected and then interviewed this time period. Given the nature of the survey design, an average of 150-300 interviews were conducted on a daily basis.

Altogether, twenty thermometer questions were asked in the NAES. Respondents were asked to evaluate the following political figures: George W. Bush, John Kerry, Dick Cheney, John Edwards, Ralph Nader, Wesley Clark, Howard Dean, Richard Gephardt, Joe Lieberman, John Ashcroft, Laura Bush, Bill Clinton, Hillary Clinton, Rudy Giuliani, Al Gore, Teresa Heinz Kerry, Rush Limbaugh, John McCain, Condoleezza Rice, and Arnold Schwarzenegger. Unfortunately, respondents were not asked to evaluate all of these individuals for each wave of the cross-sectional survey. Moreover, while some overlap exists in the thermometer questions used in the NES, they are not identical. Thus we only link respondents based on the ten feeling thermometer questions that were common to both data sets (Bush, Kerry, Cheney, Edwards, Nader, Laura Bush, Bill Clinton, Hillary Clinton, Ashcroft, and McCain). In the Annenberg, respondents answered anywhere between 4-7 questions. And in the NES, they were asked all ten feeling thermometer questions.

Our formula for pairing respondents is quite straightforward. For each respondent in the larger yet less comprehensive sample (Annenberg), we search for the respondent in the smaller and more comprehensive sample (NES) who has the closest set of thermometer scores for a given set of political stimuli. We identify the respondent whose *link score* minimizes the following expression:

$$Link\ Score = \frac{\sum_{k=1}^K |r_i - r_j|}{K}$$

where K denotes the number of political stimuli in common, r_i is the i th respondent in one of the surveys and r_j is the j th respondent in the other survey. If a respondent pairs perfectly to all of the stimuli, then his/her match score would be 0.

For example, suppose the Annenberg respondent answers five thermometer questions—he gives Bush a score of 100, Kerry 0, Cheney a score of 60, Edwards a score 40 and Bill Clinton a score 0. Our method then finds an NES respondent with the closest scores for all the stimuli. Thus, suppose an NES respondent gives Bush a score of 100, Kerry a score of 0, Cheney a score of 50, Edwards a score of 50, and Bill Clinton a score 0. Of these five answers, the Annenberg and NES respondent only differ on his/her scores for Cheney and Edwards. For this respondent, his/her link score would be:

$$Link\ Score = \frac{(|100-100|) + (|0-0|) + (|60-50|) + (|40-50|) + (|0-0|)}{5} = 4$$

If this is the lowest possible link score, then this is the NES respondent who is closest to this particular Annenberg respondent. Figure 1 presents the distribution of the link score, which ranges from a minimum of 0, which indicates a perfect score, to a maximum of 25.83. The average link score is 4.37, with a standard deviation of 3.64. Considering that an individual responded to an average number of approximately 4.5 stimuli (with a standard deviation of .67), this means that our score is off on average by only 1 unit on the 0-100 scale. Moreover, of the 74,011 Annenberg respondents with four or more thermometer scores, the

algorithm recovered 10,113 (13.7%) NES respondents with identical sets of thermometer scores.

[Figure 1 goes here]

Given that the Annenberg is much larger than the NES, certain NES respondents pair up with the Annenberg respondents more so than others. The frequency of these pairings range from 2-897. There were two NES respondents who paired with Annenberg respondents more than 1,000 times, but this was due to the fact that their responses were largely indifferent (most of their scores were 50). As such, we drop these observations from our data set. This leaves us with 69,011 in the linked data set.

Since we are interested in subgroups, especially those that pertain to race/ethnicity, it would also be interesting to know of the Latino, Black and White samples in the linked data, what is the percentage of Latinos, Blacks and Whites that linked from the NES data? That is, did Blacks from the NES mostly pair up with Blacks in the Annenberg? Did the majority of Latinos from the linked data map onto the Latino NES respondents?

The breakdown of the three largest ethnic/racial groups in the linked data is the following: 50,546 Whites, 8,442 Blacks and 4,277 Latinos. For Latinos in the linked data, the 147 NES Latinos paired up with them, and the number of NES Black respondents who were paired with Latinos was 435. This means that the majority of Latino respondents from the linked dataset were actually Whites from the NES (3,626). The same is true for the Black sample in our linked dataset; 6,435 were paired with White respondents from the NES, 1,424 with Blacks and 422 with Latinos. Finally, for the White sample in our linked data, the majority (45,354) mapped onto the affective signatures of White NES respondents, 2,961 Black NES respondents, and 1,500 NES Latino respondents. Given that the distribution in the

NES is heavily skewed towards Whites, it is understandable why the majority of Blacks and Latinos from the linked data mapped onto the subjective utilities of Whites. These distributions suggest that demographic characteristics do not necessarily predict an individual's subjective internal utility function.

Findings

How well does our procedure recover respondents' demographic traits, such as gender, race/ethnicity and education level, when compared to the original NES and Annenberg data? We first present these simple comparisons in order to determine the accuracy of our method. In the analyses we present below, we are interested in knowing how closely the distributions of the linked Annenberg respondents (who, because of our procedure, now have responses to all of the questions from the NES survey) compare with the distributions from those respondents who were actually interviewed by the NES. By doing so, we can evaluate the accuracy and effectiveness of our procedure. And in some of our analyses, we also present the distributions of the respondents from the Annenberg survey.

Table 1 offers some comparisons of demographic distributions for our linked sample, the original Annenberg survey, and the original NES survey. The gender distribution in the linked data is 47.6% male and 52.4% female, which is very close to the NES breakdown – 46.7% male and 53.5% female. The breakdown for a respondent's education level is also quite close across the three sets of data, though it is not as precise as the gender breakdowns. In terms of race/ethnicity, the percentage of Blacks appears to be overrepresented in our linked sample (12.5%) relative to both the Annenberg (8%) and NES sample (9.9%). Some Blacks in the NES sample had affective “signatures” that linked many non-Blacks (mostly Whites) in the Annenberg sample. The percentage of Whites in the linked sample is lower than it is for the

other two datasets. The Latino sample is much closer, with 6.3% in the linked data, 7.5% in the Annenberg and 6.7% in the NES. And finally for Asians, the percentage in each dataset is very similar. The final demographic variable that we consider is age. Here, the linked, Annenberg and NES data are nearly identical, with an average age of 48 in both the Annenberg and the linked data and 47.3 in the NES. Overall, this initial check makes us reasonably confident that our technique is doing an adequate job of recovering the distributions of demographic characteristics.

[Table 1 goes here]

Some additional checks are shown in Table 2 where we compare distributions on vote choice based on the race/ethnicity. Overall, the vote choice distributions for all three datasets are quite comparable. In the linked data, 60.3% of Latinos supported Kerry and 60.5% of Latinos in the NES sample also voted for Kerry. In the Annenberg, Latino support for Kerry was 57.8%. The distribution for Whites in the linked data seems to underestimate their support for Kerry, relative to the other two datasets, by approximately 5-8 percentage points. For Blacks, the percentage of support for Kerry (86.3) is located in between the estimates from the Annenberg and the NES, 91.5% and 84.5%, respectively.

[Table 2 goes here]

Another way to test how well the thermometer sores capture an individuals' internal utility function is to look at the distributions of the ethnic/racial group thermometers in the NES for Latinos, Blacks and Whites. These thermometer questions simply ask respondents how they feel towards "Whites", "Blacks", and "Hispanics". If these feeling thermometers were really tapping into an individual's preferences, we would expect that for each ethnic group, on average, they would feel best about their own group. We present these distributions

in Table 3 for both the linked data and for the original NES data. First, we see that consistent with our expectations, each group evaluates their own group the highest. For example, when evaluating their own group, Black respondents have an average score of 88.9 in the linked data and 87 in the NES data. Blacks then feel warmest towards Whites, followed by Latinos. Likewise, Latinos rate themselves the highest with a mean score of 82.9. But unlike Blacks, after their own group, Latinos then feel warmest towards Blacks and then Whites. For Whites, they too rate themselves the highest, followed by Blacks and then Latinos. Notice, though, that across these three racial/ethnic groups, it is Blacks who evaluate their own group with the highest score (88.9), followed by Latinos (82.9) and then Whites (74.8). Blacks may feel “warmest” towards their own group due to their shared historical experiences of discrimination in the U.S., which as Dawson (1995) argues, has created a very powerful and cohesive black group identity. On the other hand, given that the term Latino is a panethnic label that encompasses individuals from various Spanish-speaking countries of origin, their level of group cohesiveness and identity may not be as strong as it is for Blacks. These distributions also show that the linked data provide very similar distributions to the ones from the NES data. Again, this gives us additional confidence in the effectiveness of our procedure.

[Table 3 goes here]

All of these reality checks provide us with enough reassurance to proceed with addressing the second part of the question that we posed at the onset of the paper—what happens when the appropriate question does not exist in a particular data set? Recall that the Annenberg data does not contain any 7-point issue scale questions, which are extremely valuable in understanding voter preferences, since they ask individuals to place themselves on a 1-7 point scale on a number of different of issues, ranging from the U.S. intervention in Iraq

to government aid in assisting Blacks and Hispanics.¹ The end points of the 7-point scales are labeled and respondents are told these (usually) polar opposite positions. For example, for the “government services” scale the questions are phrased in the following manner: “Some people think the government should provide fewer services even in areas such as health and education in order to reduce spending. Suppose these people are at one end of a scale, at point 1. Other people feel it is important for the government to provide many more services even if it means an increase in spending. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5 or 6.”² To compare how well our matching procedure predicts responses to these NES questions Table 4 shows the average response for our linked data set and the NES.

[Table 4 goes here]

Altogether we examine nine 7-point issue scale questions, and we compare the mean responses of the linked and the NES respondents based on one’s race/ethnicity, gender, and vote choice (Bush or Kerry). Again, recall that the row entries in Table 4 are Annenberg survey respondents who have now “answered” the NES 7-point issue scale questions as a result of our linking procedure. Across these distributions, the mean responses appear to be quite similar across these different sub-group populations. Responses by Latinos, Blacks, and Whites in the linked data are nearly identical to those in the NES for the scaling questions pertaining to government services, defense spending, jobs, aid to blacks, the environment, and aid to Hispanics. Likewise, the distributions of the mean responses to the other demographic

¹ Respondents are also asked to place candidates on these 7-point issue scale questions, making it possible to calculate the distance between an individual’s position on an issue from their placement of the candidate’s position on that issue.

² The appendix contains the complete wording for all the 7-point issue scale questions that we examine.

sub-group that we examine, gender, are comparable in both sets of data. For example, in the scaling question that asks about women's role in society, the mean response in the linked dataset for women is 1.90, and the average response of women from the NES is 1.88. The mean response by men in the linked data is 1.93 and in the NES, men's mean response is 1.96. In fact, the largest discrepancy in the NES and the linked distributions, based on gender, is only .25. The other sub-group that we examine is those who supported Kerry versus those who supported Bush in 2004. Here, we once again find the means from the two datasets are quite similar to one another.

Next, we compare the distributions to an issue question common to both the NES and the Annenberg survey. This is a particularly rigorous way to test the validity of our procedure because if the distribution on the linked respondents, who are in fact the Annenberg respondents answering the NES question, reproduces a similar distribution from the Annenberg respondents, then we have every reason to believe that linking individuals based on subjective utility can recover groups' distributions on policy preferences and attitudes. Thus, in Table 5, we present the distributions to this common issue question, which asks respondents whether they approved or disapproved of the way George W. Bush is handling the economy. We also include the distributions from the NES survey, in order to check whether the linked distributions reflect the Annenberg distribution more so than the NES. As in our presentation of previous distributions, we look at groups' opinions (race/ethnicity and gender).

[Table 5 goes here]

What is striking about these distributions is how well the linked data recovers the gender and racial/ethnic group distributions of the Annenberg survey. For instance, the distributions by race/ethnicity in the linked dataset are 48.8% for Latinos, 51.3% for Whites

and 15.3% for Blacks, and in the Annenberg data, this breakdown is 43.9%, 48.6% and 13.7%, respectively. Moreover, the distributions by gender are nearly identical in the linked and in the Annenberg data – 47.8% of men in the linked data approved the President’s performance on the economy while 48.5% of men in the Annenberg data held this opinion. The comparison for women is also fairly similar – 45.6% in the linked data and 42% in the Annenberg.

Given these reality checks, we now turn to the distributions of our linked sample across the NES 7-point scales. In Table 6 we present the distributions by ethnicity of the linked and NES samples over four of the 7-point issue scale questions – health insurance, government spending, U.S. intervention and government aid to blacks. Across these three racial/ethnic groups, the overall distributions from the linked sample (recall that these are Annenberg respondents “answering” NES questions) reproduce the NES distributions reasonably well. Take for instance the distribution of White opinions on health insurance. In the linked data and the NES they are comparable at almost every point on the scale, with the largest difference being 3.4%. The distribution of Black and Latino opinions on this issue in the linked and the NES data are also quite similar, though at the tail ends of the scale we see some differences. For instance, 42.6% of Blacks place themselves at “1” in the linked data, but only 31.5% of Blacks from the NES do so. Moreover, 2.6% of Latinos in the linked sample place themselves at “2” on the health insurance scale, though a much larger percentage of Latinos, 9.1%, place themselves on this same value in the NES. But at every other point on this scale, the distributions on Black and Latino opinions towards the provision of health insurance in both the linked and the NES data are comparable. For the other three 7-point issue scale distributions, we see a similar pattern as well. Thus with relatively few exceptions, the linked distributions reproduce the NES distributions quite closely on a question that the Annenberg

respondents “hypothetically” answered as a result of our technique. Because of the small samples of Blacks and Latinos in the 2004 NES we think that the distribution of our much larger linked samples over the 7-point scales more accurately represent the distribution of Blacks and Latinos in the general population. We recognize that that this depends upon our claim that our linking procedure can ignore ethnicity by matching on subjective utility thereby assigning attitudes on policy issues by Whites in the NES to Blacks or Latinos in the Annenberg survey. However, our claim is not that radical nor is it a big leap of faith. Note that it boils down to saying that many Whites share the opinions of Blacks and Latinos (note the distributions of the NES respondents from the three groups greatly overlap in Table 6) and our subjective utility method simply matches some Whites with some Blacks and some Latinos who share the same opinions.

[Table 6 goes here]

Our final test compares issue distributions from our linked data with a completely independent data source. By doing so, we can compare how well our linking procedure recovers the issue preferences and attitudes of racial/ethnic groups in the U.S. We examine two sets of data from the Pew Research Center for the People and the Press, one taken in 2004 and the other conducted in 2006.³ The 2004 Pew survey is ideal since it was administered in the same year as the Annenberg survey, although the sample size of Blacks and Latinos is rather small (N <200). As a result, we also look at data from the 2006 Pew Immigration Survey, as it interviewed a much larger number of Blacks and Latinos. The issue question that was common to all three surveys pertains to presidential approval ratings. Table 7 presents these distributions, by a respondent’s race/ethnicity.

³ Both of these datasets are available at: <http://people-press.org/dataarchive/#2006>

[Table 7 goes here]

Overall, the distributions on presidential approval in the linked data are very similar to the distributions on presidential approval from 2004 Pew survey data. And relative to the 2006 Pew data, the distributions are also quite comparable to one another. For instance, in the linked data, we see that 85% of Blacks disapproved of Bush, while 83.1% of Blacks in the 2004 Pew and 88.5% of Blacks in the 2006 Pew responded in the same fashion. For Latinos, 54.1% from the linked data disapproved of Bush, 55.9% from the 2004 Pew, and 63.8% from the 2006 Pew. Finally, 60.6% of Whites approved of Bush in the linked data, 56.8% in the 2004 Pew data, and 55.6% in the 2006 Pew data. Across these three racial/ethnic groups, the difference in the distributions between the 2004 Pew and the linked data ranges from 1.8%-3.8%. This level of precision helps to validate our procedure and provides us with considerable reassurance that, by linking individuals based on their affective signatures, their political attitudes can be accurately recovered.

Conclusion

In this paper, we have outlined a method for linking respondents from different surveys based on their subjective *utility*, which we capture by using an individual's responses to a set of feeling thermometer questions. Such questions, as demonstrated in previous research as well as in this paper, are an accurate measure of an individual's *subjective utility* since they are measures of *affect*. In applying our technique to the 2004 Annenberg survey and the 2004 NES survey, we find that linking survey respondents based on their thermometer scores not only recovers the distributions on group demographics such as race/ethnicity, gender, and education extremely well but it also recovers the distributions of these groups' preferences across an array of issues and policies.

For researchers who are interested in sub-group populations, such as state politics and racial politics scholars, our technique offers them the opportunity to use survey data that may have been previously difficult to use due to sample size issues or lack of a comprehensive set of questions. For example, much of the empirics testing the spatial theories of voting have largely been applied to Whites, as they are the predominant racial/ethnic group in the U.S. But for researchers interested in applying the spatial theories of voting to other racial/ethnic groups, the capability to do so is now available, as one could use our procedure and apply it to existing datasets that contain the necessary questions to that test these theories. Moreover, this new technique enriches and expands the work on matching and survey research more broadly, as we demonstrate that linking individuals based on their affective signatures can predict their political attitudes quite accurately.

References

- Arceneaux, Kevin, Alan S. Gerber and Donald P. Green. 2006. "Comparing Experimental and Matching Methods Using a Large-Scale Voter Mobilization Experiment." *Political Analysis*, 14: 37-62.
- Cahoon, Lawrence S., Melvin J. Hinich, and Peter C. Ordeshook. 1978. "A Statistical Multidimensional Scaling Method based on the Spatial Theory of Voting." In *Graphical Representation of Multivariate Data*, ed. by P.C. Wang. New York, Academic Press.
- Dawson, Michael C. 1995. *Behind the Mule: Race and Class in African-American Politics*. Princeton, NJ, Princeton University Press.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Enelow, James M. and Melvin J. Hinich. 1984. *The Spatial Theory of Voting*. New York: Cambridge University Press.
- Gerber, Alan S., and Donald P. Green. 2000. "The Effects of Personal Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review*, 94(3): 653–664.
- Gerber, Alan S. and Donald P. Green. 2005. "Do Phone Calls Increase Voter Turnout? An Update." *Annals of the American Academy of Political and Social Science*, 601:142-154.
- Giles, Michael, and Arthur Evans. 1986. "The Power Approach to Intergroup Hostility." *Journal of Conflict Resolution*. 30: 469-486.
- Greiner, D. James. 2006. "Causal Inference in Civil Rights Legislation." Working Paper.

- Ho, Daniel E., Kosuke Imai, Gary King and Elizabeth A. Stewart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15:199-236.
- Imai, Kosuke. 2005. "Do Get-Out-the-Vote Calls Reduce Turnout? The Importance of Statistical Methods for Field Experiments." *American Political Science Review*, 99(2):283–300.
- Knight, Kathleen. 1984. "The dimensionality of partisan and ideological affect." *American Politics Quarterly*. 12: 305-334.
- Nickerson, David W. 2005. "Scalable Protocols Offer Efficient Design for Field Experiments." *Political Analysis*. 13: 233-252.
- Poole, Keith T. 1984. "Least Squares Metric, Unidimensional Unfolding." *Psychometrika*, 49(3): 311-323.
- Poole, Keith T. 1990. "Least Squares Metric, Unidimensional Scaling of Multivariate Linear Models", *Psychometrika*, 55: 123-149.
- Poole, Keith T. and Howard Rosenthal. 1984. "U.S. Presidential Elections 1968-1980: A Spatial Analysis." *American Journal of Political Science*, 28:282-312.
- Rabinowitz, George. 1976. "A Procedure for Ordering Object Pairs Consistent with the Multidimensional Unfolding Model", *Psychometrika*, 45:349-373.
- Romer, Daniel, Kate Kenski, Kenneth Winneg, Christopher Adasiewicz and Kathleen Hall Jamieson. 2006. *Capturing Campaign Dynamics 2000 & 2004*. Philadelphia, PA, University of Pennsylvania Press.

Weisberg, Herbert F. and Jerrold G. Rusk. 1970. "Dimensions of Candidate Evaluation."

American Political Science Review, 64:1167-1185.

Wilcox, Clyde, Lee Sigelman and Elizabeth Cook. 1989. "Some Like It Hot: Individual

Differences in Responses to Group Feeling Thermometers." *The Public Opinion*

Quarterly. 53(2):246-257.

Zeng, Langche. 2007 "Matching and Causal Inference", discussion notes, 2007 Summer

Workshop, Department of Political Science, University of California San Diego.

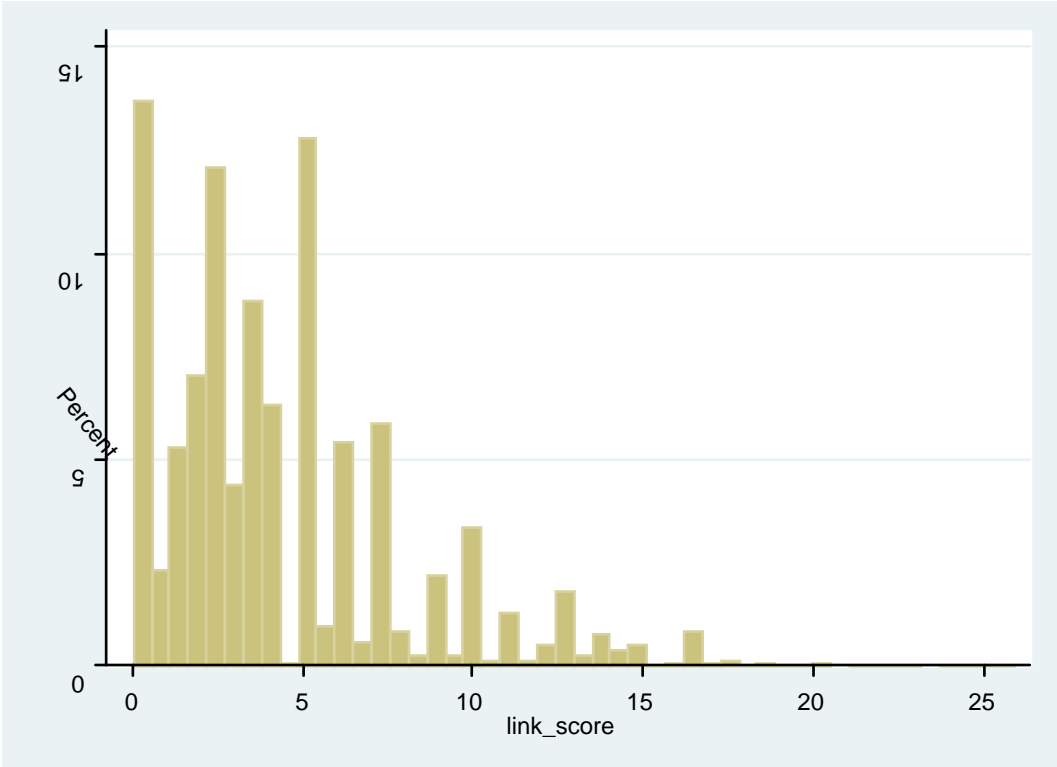


Figure 1: Link Score Distribution

Table 1: Comparison of Demographic Variables

	Linked	Annenberg	NES
Gender (%)			
Male	47.6	44.7	46.7
Female	52.4	55.3	53.5
Education (%)			
8th or less	2.4	2.0	3.1
9-11th grade	5.5	5.3	6
HS degree	28.6	25.6	29.3
Some college	24.5	17.7	21.8
2yr college	8.1	7.9	9.9
BA	17.6	20.0	18.4
Advanced	13.3	14.4	11.5
Race/Ethnicity (%)			
Black	12.5	8.0	9.9
Asian	1.7	1.5	1.6
White	74.8	83.3	78.0
Latino	6.3	7.5	6.7
Age (mean)	48.0	48.0	47.3
N	69,011	81,422	1,212

Table 2: Vote Choice Distributions, by ethnic/racial group

Race/Ethnicity	Vote Choice (Linked)		Vote Choice (NES)		Vote Choice (Annenberg)	
	Kerry	Bush	Kerry	Bush	Kerry	Bush
Latino	60.3	37.9	60.5	34.9	57.8	26.6
Black	86.3	12.4	84.5	12.1	91.5	7.3
White	35.0	62.8	40.0	57.6	43.5	55.0
<i>N</i>	21214	25980	399	412	1036	1113

Table 3: Group Thermometer Evaluations, by Ethnic/Racial Group

	Group Thermometer Scores Towards...							
R's Race	<i>Latinos</i>		<i>Blacks</i>		<i>Whites</i>		<i>N</i>	
	Linked	NES	Linked	NES	Linked	NES	Linked	NES
Latino	82.9 (14.0)	82.7 (15.5)	76.3 (20.1)	75.8 (18.7)	71.7 (19.6)	74.2 (18.3)	3162	66
Black	67.48 (17.5)	68.8 (18.2)	88.9 (14.7)	87.0 (15.5)	70.9 (23.0)	72.3 (20.0)	7364	154
White	66.6 (19.1)	66.6 (19.3)	68.60 (18.8)	69.2 (18.4)	74.8 (19.2)	73.8 (19.2)	45813	763

Entries in parenthesis denote the standard deviation and the entries not in parenthesis are the mean thermometer scores for each respondent's racial/ethnic group.

Table 4: Comparison of Mean Responses to 7-point Issue Scales, Linked vs. NES Data

	Gov't services 7pt scale		Defense Spending 7pt scale		Gov't Jobs 7pt scale		Gov't Aid to Blacks 7pt scale		Env't vs. Jobs 7pt Scale		Women's Role 7pt scale		Govt vs. Private Health Insur	
Race	Linked	NES	Linked	NES	Linked	NES	Linked	NES	Linked	NES	Linked	NES	Linked	NES
Latino	4.44	4.57	4.45	4.49	4.49	4.28	4.26	4.28	3.96	3.82	1.57	1.53	3.75	3.42
Black	5.41	5.25	4.37	4.30	3.19	3.31	3.02	3.31	3.81	3.71	1.57	2.01	2.66	3.31
White	4.28	4.36	4.78	4.65	4.49	4.82	4.93	4.82	3.65	3.58	1.99	1.93	3.91	3.78
Gender														
Men	4.20	4.32	4.87	4.74	4.44	4.61	4.88	4.61	3.73	3.52	1.93	1.96	3.83	3.79
Women	4.61	4.69	4.54	4.41	4.25	4.48	4.50	4.48	3.60	3.66	1.90	1.88	3.67	3.54
Vote Choice														
Bush	3.58	3.71	5.38	5.19	5.23	5.17	5.43	5.30	4.25	4.04	2.33	2.17	4.55	4.41
Kerry	5.14	5.07	4.02	3.89	3.53	3.49	3.71	3.88	2.93	3.02	1.83	1.72	2.92	3.06
<i>N</i>	60964	1060	61090	1061	64572	1103	61460	1073	58934	1019	50598	1157	62388	1112

	Gov't aid to Hispanics 7pt scale		U.S. Intervention 7-pt scale	
Race	Linked	NES	Linked	NES
Latino	3.85	3.68	4.14	3.65
Black	3.95	3.84	3.05	3.19
White	4.96	4.92	4.01	3.93
Gender				
Men	4.93	4.71	4.23	3.98
Women	4.62	4.62	3.54	3.55
Vote Choice				
Bush	5.20	5.15	4.68	4.67
Kerry	4.05	4.17	2.68	2.81
<i>N</i>	55015	937	60757	1041

Table 5: Distribution of Linked, Annenberg and NES Data on a common issue question

	<i>Approve of the way the President is handling the economy (% Approving)</i>		
	Linked	NES	Annenberg
Race			
Latino	48.8	35.8	43.9
Black	15.3	12.8	13.7
White	51.3	47.3	48.6
Gender			
Men	47.8	43.1	48.5
Women	45.6	38.1	42.0
Aggregate	46.7	40.4	44.9
<i>N</i>	69,011	1121	84,122

Table 6: Distribution of 7-point scales, by race/ethnicity

<i>7-pt issue scale</i>	Blacks		Latinos		Whites	
	Linked	NES	Linked	NES	Linked	NES
<i>Health Insurance</i>						
1 (<i>Gov't</i>)	42.6	31.5	21.6	19.7	17.7	16.6
2	8.9	9.4	2.6	9.1	10.4	13.0
3	8.9	9.4	15.3	22.7	9.8	13.7
4	17.1	22.0	24.9	22.7	22.7	19.3
5	10.5	11.3	15.8	15.2	17.3	16.0
6	4.2	6.9	13.7	4.6	12.6	12.8
7 (<i>Individual/Private</i>)	7.7	9.4	6.1	6.1	9.6	8.7
N	7735	159	2930	66	47,527	820
<i>Gov't Services</i>						
1 (<i>Decrease</i>)	2.3	3.4	7.1	6.0	7.1	4.9
2	1.0	1.3	1.0	4.5	9.9	8.1
3	4.7	4.0	9.3	10.5	14.1	13.4
4	22.2	25.5	44.4	26.9	23.8	26.8
5	19.7	20.1	14.9	22.4	20.5	24.2
6	15.1	15.4	12.4	17.9	12.2	12.0
7 (<i>Increase</i>)	35.1	30.2	10.9	11.9	12.3	10.6
N	6784	149	3857	67	45,831	776
<i>U.S. Intervention</i>						
1 (<i>Diplomacy</i>)	32.8	29.0	15.9	19.7	13.9	12.8
2	6.8	9.2	5.5	13.6	7.6	9.6
3	14.7	13.7	3.8	6.1	13.4	14.3
4	33.1	29.0	36.4	27.3	24.4	25.6
5	5.6	7.6	18.4	18.2	21.7	19.3
6	2.2	3.8	7.8	6.1	8.3	9.1
7 (<i>Military Force</i>)	4.9	7.6	12.3	9.1	10.7	9.2
N	6376	131	3907	66	46,534	781
<i>Aid to Blacks</i>						
1 (<i>Gov' Aid</i>)	31.8	28.2	13.4	7.5	5.4	4.7
2	9.7	9.8	4.5	9.0	2.9	4.5
3	7.4	10.4	11.7	11.9	5.4	8.4
4	33.1	27.6	19.8	28.4	27.2	25.5
5	12.3	12.3	24.8	16.4	14.8	17.8
6	0.2	1.2	13.6	16.4	22.0	21.2
7 (<i>No Aid</i>)	5.5	10.4	12.1	10.5	22.1	17.9
N	7999	163	2993	67	46,760	782

Table 7: Presidential Approval Ratings, by Race/Ethnicity

	Blacks			Latinos			Whites		
<i>Presidential Approval Rating</i>	Pew '06	Pew '04	Linked	Pew '06	Pew '04	Linked	Pew '06	Pew '04	Linked
Approve	11.5	16.9	15.0	36.2	44.1	45.9	55.6	56.8	60.6
Disapprove	88.5	83.1	85.0	63.8	55.9	54.1	44.4	43.2	39.5
<i>N</i>	686	260	8,383	698	179	4,391	4,514	2218	51,413

Source: Pew Research Center for the People and the Press, Biennial Media Consumption Survey (June 2004) and Pew Research Center for the People and the Press & Pew Hispanic Center, 2006 Immigration Survey.

Appendix

7-point issue scale questions

Intervention-diplomacy in Iraq: “Some people believe the United States should solve international problems by using diplomacy and other forms of international pressure and use military force only if absolutely necessary. Suppose we put such people at "1" on this scale. Others believe diplomacy and pressure often fail and the US must be ready to use military force. Suppose we put them at number 7. And of course others fall in positions in-between, at points 2, 3, 4, 5, and 6.”

Government spending: “Some people think the government should provide fewer services even in areas such as health and education in order to reduce spending. Suppose these people are at one end of a scale, at point 1. Other people feel it is important for the government to provide many more services even if it means an increase in spending. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5 or 6.”

Defense Spending: “Some people believe that we should spend much less money for defense. Suppose these people are at one end of a scale, at point 1. Others feel that defense spending should be greatly increased. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5 or 6.”

Government vs. private health insurance: “There is much concern about the rapid rise in medical and hospital costs. Some people feel there should be a government insurance plan which would cover all medical and hospital expenses for everyone. Suppose these people are at one end of a scale, at point 1. Others feel that all medical expenses should be paid by individuals through private insurance plans like Blue Cross or other company paid plans. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6.”

Jobs: “Some people feel the government in Washington should see to it that every person has a job and a good standard of living. Suppose these people are at one end of a scale, at point 1. Others think the government should just let each person get ahead on their own. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6.”

Aid to Blacks: “Some people feel that the government in Washington should make every effort to improve the social and economic position of blacks. (Suppose these people are at one end of a scale, at point 1.) Others feel that the government should not make any special effort to help blacks because they should help themselves. (Suppose these people are at the other end, at point 7.) And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6.”

Aid to Hispanics: “Some people feel that the government in Washington should make every effort to improve the social and economic position of Hispanics. (Suppose these people are at one end of a scale, at point 1.) Others feel that the government should not make any special effort to help Hispanics because they should help themselves. (Suppose these people are at the other end, at point 7.) And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6.”

Environment vs. Jobs: “Some people think it is important to protect the environment even if it costs some jobs or otherwise reduces our standard of living. (Suppose these people are at one end of the scale, at point number 1) Other people think that protecting the environment is not as important as maintaining jobs and our standard of living. (Suppose these people are at the other end of the scale, at point number 7.) And of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6.”

Women’s role in society: Recently there has been a lot of talk about women's rights. Some people feel that women should have an equal role with men in running business, industry, and government. (Suppose these people are at one end of a scale, at point 1.) Others feel that a woman's place is in the home. (Suppose these people are at the other end, at point 7.) And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5 or 6.