

Is Eleven Really a Lucky Number?

Measuring Trust and the Problem of Clumping*

Eric M. Uslaner

Department of Government and Politics

University of Maryland–College Park

College Park, MD 20742

euslaner@gvpt.umd.edu

ABSTRACT

Measuring trust has long been controversial. Recently several studies have shifted from the standard dichotomous measure to an 11-point scale and advocates of this new scale argue that it is superior and that it leads to more convincing results. They do not consider problems with the 11-point scale, notably "clumping," the tendency of people to give responses clustering around the middle. I show that clumping is pervasive across 11-point scales and even among 5 and 7 point scales, that respondents often assign the middle categories somewhat randomly. There is also evidence that clumping is patterned and I estimate models for clumping patterns showing that better connected people are more likely to clump on generalized trust and less likely to clump on institutionalized trust.

Ever since generalized trust has entered the discourse of survey research—long before anyone was talking about or writing about “social capital”—its measurement has stoked controversy. Until recently, measures of trust from the early surveys in National Opinion Research Center surveys in the 1940s and 1950s to what is now known as the “standard question” formulated by Elisabeth Noelle-Neumann in Germany in 1948, have all been dichotomies. This measure has been used in a large number of surveys, most notably the Civic Culture surveys, the American National Election Studies from 1964 to 1976 and from 1992 to the present, the General Social Survey from 1972 to the present, the World Values Surveys (WVS) since 1980, and many other surveys. In the late 1990s, the Swiss Household Panel and the Citizenship, Involvement, Democracy (CID) cross-national surveys in Europe¹ shifted to an 11-point scale. Shortly thereafter the new European Social Survey (ESS) adopted the 11-point scale as well.² The American National Election Studies (ANES) in 2008 has adopted a five point scale in addition to the standard question as part of its collaboration of the National Longitudinal Study of Youth, which long has used the latter measure.

Is the 11-point scale an advance over the dichotomous measure? Or are the benefits oversold? I argue the latter. The arguments for greater precision may be compelling, but the larger number of alternatives may induce confusion in the minds of respondents. Using the ESS and the CID surveys for the United States and for Romania, Moldova, and Spain (which included both the 11 point scale and the dichotomy), I show that respondents tend to “clump” in the midpoints of the scale (at 4, 5, and 6) for both generalized trust and confidence in political institutions, that these clumping responses are essentially random, and that the 11-point scale is only moderately related to the dichotomy. The presumed benefits of the 11-point scale are, in

Samuel Johnson's description of second marriages, "the triumph of hope over experience."

Miller (1956) argued many years ago that a seven point scale is the most that people can handle in a survey question. Beyond seven options, people lose the ability to discriminate among alternatives. Students of psychology learn about the problem of too many choices: Too many options simply confuse people and they might make suboptimal choices (Schwartz, 2003). An 11-point scale may simply demand too much of respondents. Yet, seven-point or even five-point scales may not resolve the problem of clumping. I argue below that the ANES five point scale showed severe clumping in the 2006 Pilot study—and that there is evidence of clumping in seven-point scales in the 2004 ANES. Even seven-point scales are not immune from clumping.

Measuring Trust

I focus primarily on the "new" 11-point scale used in the Swiss Household Panel and the CID and ESS surveys compared to the more commonly used dichotomy. Scherpenzeel (n.d., 1) argued:

Scales with relatively few response alternatives force respondents to categorize their reaction towards an attitude object instead of directly mapping it onto the response continuum, thus causing information loss....respondents differentiate more between objects when offered response scales with greater numbers of categories...

Scherpenzeel and Saris (1995) that the 11-point scale demonstrated the greatest reliability and validity of any measure of satisfaction in a survey across 10 nations.

Zmerli and Newton (2008) and Hooghe and Reeskens (2007) argue that the 11-point scale marks a great improvement over the dichotomous measure of generalized trust. Zmerli and

Newton (2008) show that the 11-point scales for generalized trust and trust in government lead to much more robust correlations between the two measures than does the simple dichotomy for generalized trust and a four-category trust in government measure, as in the WVS. They argue: “Most of the studies finding no significant correlations are based on World Values or Eurobarometer surveys, which use single questions for social and for political trust, and either a dichotomised (Yes/No) response option, or else a four-point scale.” When they use the 11 point scales in the ESS, they find for all 24 countries in the sample “strong and statistically significant four-cornered set of associations between generalised social trust [and] confidence in domestic political institutions...after controlling for a set of eight independent variables that have commonly been found to be associated with social and political trust...” Moreover, these results stand even when they replace the three variable misanthropy scale with the single measure for generalized trust using the 11-point scale.

Zmerli and Newton (2008) dichotomize the 11-point scale in their tests for a dichotomous measure. They also collapse the trust in government measure to a four-point scale, as is standard in the ANES and WVS. Hooghe and Reeskens (2007) are better able to compare the dichotomous and 11-point scales, since their Belgian Youth Survey of 2006 included *both measures*. They argue (19-20) that the 11-point measure outperforms the dichotomy in a regression analysis of generalized trust. Gender, religious involvement, membership in associations, and immigrant status are all significant predictors of the “continuous” measure, but not of the dichotomy.³

Case made? Not quite. There are two key issues here. One is theoretical, the other methodological: First the theoretical concerns: Why should there be strong relations between

trust in other people and confidence in government? Why should we expect gender or membership in associations to be related to generalized trust—or religious involvement to lead to *more trust in other people*? The literature on generalized trust is badly split on whether we should expect ties between generalized trust and confidence in government. Zmerli and Newton (2008) now boldly claim that “generalised social trust, political confidence in domestic and international institutions, and satisfaction with democracy are indeed tied together in a tight four-cornered syndrome” (cf. Putnam, 1993, 169-171, 180). Yet, earlier Newton (1999, 179-181) proclaimed, based upon findings from the WVS, that “[p]ersonal trust belongs to the private sphere; political trust belongs to the public political sphere....political trust is usually of a thinner kind than social trust than social trust...” and “social trust tends to be high among those who hold a central position in society...political trust is often randomly distributed throughout social groups and types...” (cf. Uslaner, 2002, ch. 5). Politics is about winners and losers and is inherently polarizing (Anderson et al., 2005 ; Newton, 1999, 181; Uslaner, 2002, 120), while generalized trust is all about cooperation among people with diverse preferences (Putnam, 1993, 169; Uslaner, 2002, ch. 2).

There is also little reason to believe that gender should affect generalized trust or that religious involvement, especially in a country that is overwhelmingly Catholic such as Belgium, would be associated with higher trust. Putnam (1993, 107) argues that Catholicism, as a hierarchical religion, should lead to *lower trust*. The connection between trust and associational membership is more contested, but the bulk of the evidence suggests that joining voluntary associations *does not* lead people to become more trusting (Claibourne and Martin, 2004; Stolle, 1998; Uslaner, 2002, chs. 4-5).

If a “better” measure of social trust leads to findings that are counterintuitive rather than grounded in strong theory, we might be more skeptical than positive about the new indicator. Gender has rarely been significant in the now burgeoning literature on trust; when Catholicism is related to trust, it is mostly negative. The connections between generalized trust and both confidence in government and joining voluntary associations are largely weak—and support for stronger links should rest at least as much on better theoretical linkages than on new measurements. Stronger relationships are not always signs of a methodological advance.

The second concern is methodological: Does the 11-point scale represent an advancement over the dichotomous measure? The core of my argument is not theoretical even as I am skeptical that Zmerli and Newton and Hooghe and Reeskens have made their case based upon the findings they report. Rather, I focus on the properties of the 11-point scales and later seven point scales. I turn now to the problem of clumping.

The Center Always Holds

No one would argue that trust is like a light switch, always on or off. People may be more or less trusting of others. But that is not how we have measured it. We have traditionally given people a simple choice of whether most people can be trusted or should we be more careful in dealing with others. An attempt to refine the measure of trust—to give people greater opportunities to express their *level of trust*—should surely be welcome. If, as Scherpenzeel (n.d.) argues, the 11-point scale has been tested and is ready, then it should be a significant improvement over the dichotomy.

Yet, there is a problem lurking behind this claim. Does the 11-point scale offer the appropriate number of alternatives to provide reliable estimates of trust? If 11 options are too

many for people to handle, how might people respond? The hopeful response from Scherpenzeel (n.d., 2) is that people will focus on the two reference points, 0 and 10. But this makes little sense for a measure such as trust. Few people would say that they are completely trusting of others: We can always think of reasons why we shouldn't give complete faith to others. Few would admit to being the ultimate Scrooge, who trusts nobody at any time. An alternative reference point is psychologically appealing: The midpoint of five. If people are confused by the 11-point scale they should cluster around the mean—a psychological regression to the mean. But we might not expect people to focus only on one value. If they are confused—or simply uncertain—about the scale, they may select values nearby the mean such as four or six. This is what I call “clumping,” a focusing on the center of the distribution. If survey respondents were to “clump” at 4, 5, and 6 for one question or in just one survey, we might attribute this to question wording or order or to some fluke. However, if clumping is systematic—occurring across a large range of questions in multiple surveys, this would constitute evidence that the 11-point scale may lead people to focus on the center. This is what we see across trust questions in the ESS and the United States CID.

I present histograms of the various trust measures in the first two waves of the ESS in Figures 1 and 2. The picture is clear: There is substantial clumping occurring across all questions. There is a skewness toward greater trust for police and the United Nations and toward less confidence for politicians (see Figure 3). For every measure except trust in the police—where 38 percent of respondents chose 4, 5, or 6—clumping respondents constitute over 40 percent of all answers to these questions. This is evidence of a flight to the middle.

Clumping on generalized trust falls below 40 percent only where we know trust to be

high— in the Nordic countries (Denmark, Finland, Norway, and Sweden) and in two other countries (Greece and Ireland). The pattern is largely the same for trust in the police. Yet trust in parliament is uniformly high—below 40 percent in two countries with extraordinarily low or high levels of confidence. Denmark and the Czech Republic rank 1.75 standard deviations above and below the mean on confidence in parliament and Poland falls 2 standard deviations below the mean. Denmark falls just below the mark of 40 percent clumping, while only one-third of respondents in the two transition countries give responses in the middle.

It is hard to escape the conclusion that there is a flight to the center across all six questions for most countries. Across the six questions and 21 countries, 41 percent of all responses to trust questions were 4, 5, or 6. Only Denmark (33 percent) and Finland (35 percent) fell below 37 percent; Switzerland, Belgium, Italy, France, Spain, and the Netherlands were all above 45 percent.

Clumping is not restricted to the ESS or to Europe. The United States CID survey⁴ includes a wide range of trust questions—about both other people and institutions. I present the histograms for these measures in Figures 4 and 5 and a bar chart of clumping responses in Figure 6.

[Figures 4, 5, and 6 about here](#)

There is considerable variability in the level of clumping. Trusting neighbors and fellow club members have fewer than 30 percent clumping, but 12 of the 19 measures have between 40 and 50 percent clumping. Clumping seems least common among people you know well, most common among national political institutions. Even though there is considerable variation

across the 19 measures, the standard deviation of clumping responses is rather low— at .07, which is slightly larger than the .04 for seven measures in the ESS. The mean clumping shares are identical at .41. For both the ESS and the US CID, seeking the middle ground is ubiquitous.

There is a logic behind the patterns of clumping. There is a moderate *negative* relationship between aggregate levels of clumping on generalized trust and the mean level of trust across countries ($r^2 = .317$, see Figure 7). But there is a much stronger *positive* relationship between aggregate levels of trust in politicians and the mean of confidence in these elites ($r^2 = .829$). It may seem that this reflects either randomness or, worse, a schizophrenic pattern. These graphs suggest that people focus on the midpoints rather than give what they may think are unacceptably high or low values to a survey question. Clumping is most common on generalized trust in countries with low levels of faith in people. So people who may not trust others may seek a more “socially acceptable” middle ground rather than give responses at the bottom of the scale. Clumping is most common on trust in politicians in countries where confidence is *highest*. Yet skepticism of politicians, rather than confidence, is the norm. The mean score on the 11-point scale across the 21 countries for trust in politicians is 3.8. Only in Denmark is the mean score above 5 (5.52).

Figure 7 about here

Similarly, for the United States CID, there is a strong negative correlation across the 19 measures in the survey between the aggregate shares of clumping and the mean levels of trust. The simple correlation is $-.70$ ($r^2 = .49$). The least clumping occurs where trust is highest—for trust in neighbors and fellow club members—the only indicators with clumping below 30 percent.

The six measures with clumping below 40 percent have a mean score of 6.0, the seven measures with 46 percent or more clumping have a mean rating of 4.5.

The Romanian CID survey presents a unique opportunity to examine this account by comparing responses to the dichotomy and the 11-point scale for the same sample.⁵ The two measures are hardly identical: The r^2 between them is a mere .164 and we see in Figure 8 that the mean score on the 11-point scale is 3.5 for mistrusters (on the dichotomy) and 5.8 for trusters. Romania is a mistrusting society: 32 percent chose the trusting option in the dichotomous measure, which is higher than in either the 1990 or 1995 wave of the World Values Survey (16 percent and 18 percent, respectively). As a Romanian woman named Sanda described her country to a young British couple walking across Central and Eastern Europe in 1993: “...nobody trusts nobody” (Goodwin, 2000, 152). The low mean score on the 11-point scale for trusters (in the dichotomy) is evidence of a low trusting society.

Figure 8 about here

There is evidence for the “socially acceptable” argument in the Romanian CID sample (see Table 1): Thirty six percent of respondents who answered both questions gave clumping responses. Seventy-six percent of Romanians scoring 4 on the 11-point measure are mistrusters on the dichotomous measure, as are 69 percent directly in the middle at 5, and even 58 percent who scored a 6 on the expanded measure. Mistrusters do not hesitate to give low scores on the 11 point scale: 87 percent of mistrusters give responses of either 0 or 1 on the 11-point scale. Only 6.8 percent of respondents in Table 1 gave responses of either 9 or 10—and almost one third (32 percent) of them said that they were mistrusters in the dichotomy. Looking at the data

percentagized by the dichotomous question, only 63 percent who said that you can't be too careful in the dichotomous question give responses of 0-4 on the 11-point scale and just 56 percent who believe that most people can be trusted give responses of 6-10 (see Table 2). There is thus some evidence, if hardly conclusive, that mistrusting people might give a more socially acceptable neutral response than one at the bottom of the scale.⁶

The problem is not unique to Romania. Both versions of the trust question were asked in Moldova, a former Soviet republic, and in Spain. In Moldova, where trust is even lower than in Romania, over half (50.8% of 1219) of the respondents gave responses of 4, 5, or 6 to the generalized trust question: 79.4 percent of "clumpers" said that "you can't be too careful" in the dichotomous question and 60 percent of respondents at 9 or 10 also said that "you can't be too careful." The r^2 between the two measures is only .061. Clumping is also prevalent in Spain: 50.7 percent of respondents gave scores of 4, 5, or 6 (2058 of 4061)—with almost a quarter picking the option directly in the middle; 77 percent of those scoring 5 said that "you can't be too careful" in the dichotomy.

There is far greater coherence in Spain: The r^2 between the two measures is .266, healthier but far from robust. Mistrusters are consistent in Spain: 92 percent of respondents scoring between 0 and 4 say that "you can't be too careful." Trusters are less consistent: Only two-thirds of respondents scoring between 6 and 10 say that "most people can be trusted," and barely more than half (50.9 percent) scoring 6 were trusters on the dichotomy. Twenty-six percent of respondents scoring at 9 or 10 also said that "you can't be too careful" in the dichotomy. Seventy percent saying that "most people can be trusted" score 6-10, but only 57 percent saying that "you can't be too careful" score 0-4. In all three countries in the CID in

which both versions of the trust question were asked, there 11-point scale did not seem to discriminate as well as its defenders might hope.

Tables 1, 2 about here

The Center versus the Extremes

Can we attribute the higher correlations between generalized trust and institutional trust that Zmerli and Newton found in the ESS to clumping? If clumping were the culprit, we should see that respondents who gave middling scores on one trust measure would do the same for other indicators of confidence. There is no support for such a claim. Instead, there seems to be the “opposite” problem: Scores across the clumping measures across ESS respondents seem almost (though not quite) random. The correlations among clumping scores are rather low, suggesting that respondents do not discriminate well among scores of 4, 5, or 6.

In Table 3, I present average correlations for each measure with each of the other trust variables for “clumping” and “non-clumping” samples. A sample correlation for this table would be for respondents who had *only* clumping (column 1) or *only* non-clumping responses (column 2) for a particular pair of variables (e.g., generalized trust and trust in parliament). I used the actual values of the trust measures rather than the dummy variables for clumping or non-clumping to compute the correlations. So the generalized trust-trust in parliament correlation for clumping would be computed only for respondents who rated *both* variables at 4, 5, or 6. For non-clumping, the correlation is computed for respondents who rated *both* variables at 0, 1, 2, 3, 7, 8, or 9.

Table 3 about here

The correlations for the “not clumping” samples are robust, especially for survey measures. This is hardly surprising. Trust in institutions largely reflects the performance of those institutions (Citrin, 1974; also as argued by Newton, 1999). When things are going well, people will have faith in a wide range of institutions. When things are not going well, confidence will lag across the board. The strong average correlations—from .576 for the somewhat distant United Nations⁷ to .670 for the more proximate Parliament—all support this argument. Even generalized trust has a moderate average correlation with institutional trust ($r = .410$). If we exclude generalized trust from the average trust in Table 3 for the “not clumping” samples, the average correlation rises to over .75.

For the “clumping” sample, the average correlations are much lower. This clearly exonerates clumping as the cause of the stronger relations that Zmerli and Newton report. However, it poses its own puzzle. The correlations among responses of 4, 5, or 6 for the “clumping” sample suggest that there is not much coherence in these rankings. A four or a five or a six is largely the same score among most of these trust measures. Even the “high” correlations—370 for “clumping respondents” for the European Parliament and the United Nations—are modest. Trust in politicians are the parliament only correlate at $r = .298$ for the “clumping” sample.

The connection between social trust and institutional confidence is essentially random (mean $r = .066$). The highest correlation between generalized trust and institutional trust for this sample is .078 (for confidence in the police).

The comparison between the clumping and non-clumping responses may be somewhat

unfair, since the range of variation in the latter (three points on the 11-point scale) is far greater than the range for the latter (the remaining eight points). The low correlations for the clumping responses are not generally attributable to the lack of variation. I coded each trust measure into two extremity measures: extreme scores (0, 1, 9, 10) and very extreme scores (0, 10). The average correlation for the extreme scores is .276 (.336 without generalized trust). For the very extreme scores, with only two values, the average correlation is .362 (.451 without generalized trust).

These results are disheartening for advocates of the 11-point scale. They call into question respondents' ability to discriminate among responses coded 4, 5, and 6. This is not a trivial problem since on average respondents to the ESS gave clumping responses over 40 percent of the time—and one-third of respondents “clumped” on half or more of the questions. While the higher correlations observed by Zmerli and Newton do not stem from the clumping problem, the inability of respondents to discriminate among some of the most common response categories is troublesome.

Clumping and Social Desirability

Without more detailed knowledge of the motivations of people responding to the trust questions, it is impossible to say anything definitive about my social desirability thesis—that people may be more willing to give higher scores on the 11-point scale to show that they are less mistrusting of other people but lower scores to indicate that they are not quite so negative about their leaders and institutions.

Some supporting, even if indirect, evidence for this claim comes from analyses of clumping and extreme values for generalized trust and confidence in institutions. I assume, from

the literature on both social and political trust (Citrin, 1974; Uslaner, 2002), that people better situated in their social and political worlds would be more likely to have faith in both other people and political institutions. If clumping is highest for people who are low on generalized trust and for those who are higher on institutional trust, a social desirability model would predict:

- People who are less well connected to their social and political worlds would be more likely to clump on generalized trust.
- People who are better connected to their social and political worlds would be more likely to clump on institutional confidence.

My argument stems from the results in Figure 7: Clumping on generalized trust is higher in countries with low levels of trust; clumping on institutional trust is higher in countries with higher levels of confidence in institutions

I test these claims by estimating a probit analysis for clumping values for generalized trust (Table 4) and a negative binomial regression for institutional trust clumping from the two-wave ESS. The clumping measure for generalized trust is a dichotomy, as above. The total clumping measure for institutional trust is the sum, across the six institutions, of the clumping responses.

These models are *not* meant to be explanations of either type of trust. As I have argued above, a more theoretically appealing argument is that generalized and institutional trust have different roots. Rather, they are designed to provide some evidence, however indirect, that my social desirability argument has some support.

How do I measure connections to one's social and political world? I am constrained by the variables available in the ESS so I focus on the following measures that seem best to tap a

either a positive orientation toward their social and political world—or at least a willingness to engage in the political world. People who withdraw from these worlds might be more alienated from both other people and their political institutions. Alienated people would be more likely to give middling (clumping) values rather than more positive evaluations for generalized trust but more likely to give middling rather than low evaluations for institutional trust.

The variables in the models are: life satisfaction, household income, education (in years), age, hours spent daily reading newspapers, frequency of Internet usage, whether the respondent voted in the last election, the level of political interest, and whether the respondent is from a transition country. People who are satisfied with their life or who have high incomes should be better connected to their social worlds—and less likely to give middling responses. More highly educated and older respondents should also be more likely to be supportive or critical responses and not to seek refuge in clumping responses.

People who read newspapers daily, who use the Internet frequently, who vote, and who express an interest in politics are better connected to their political worlds—and thus should be less likely to give middle responses to generalized trust, but more likely to clump than to give extreme (especially negative) scores on institutional confidence. Finally, people in transition countries have lower levels of trust in both other people (Uslaner, 2003) and in government, so I would expect that the clumping responses of 4, 5, or 6 might be considered “high” scores in these countries (see Table 2 on generalized trust in Romania). I present the analyses in Table 4. Since the coefficients in probit, in particular, have no straightforward interpretation, I focus on the “effects,” the changes in the probability from the minimum to the maximum values of each predictor.⁸

Table 4 about here

People very satisfied with their lives are five percent less likely to give clumping responses to the generalized trust question, though the coefficient is barely significant at $p < .10$. The biggest effect comes from age: Young people are substantially more likely to give middle responses. So are less highly educated people. These variables have the greatest effects in the model. Young people are generally less well connected to their social and political worlds and thus may be more ambivalent. More educated people, as expected, also are more likely to choose more discriminating responses. People who use the Internet frequently are four percent less likely to clump, while there is a more modest effect for voting in the last election (two percent). The other variables are all not significant—and the model does not do a very good job in predicting who will clump and who will not.

There is more support for the social desirability argument in the regression for total institutional trust clumping. The effects indicate the changes from zero to six instances of clumping across the institutional variables. Younger people, people with lower income, those who had not voted in the last election and have little political interest, and respondents from transition countries are *less* likely to clump. Overall, connection to one's political world—through voting and interest in politics—leads people to be more willing to give middle scores—presumably rather than lower values. Younger people, who traditionally participate in political life much less than the elders, are substantially more likely to give clumping responses. Younger people may also be less knowledgeable about politics—and scores of 4, 5, or 6 may reflect uncertainty about judging political leaders and institutions. Older people have greater exposure to politics and are

likely to be more discerning in their scores for the confidence measures. People in transition countries rate their institutions much less positively: They are substantially more likely to rate political institutions at either 0 or 1 and much less prone to give the top scores of 9 or 10. People in transition countries may thus see scores of 4, 5, or 6 as relatively high—and shy away from them.

More for institutions than for generalized trust there are clear patterns for clumping. Rating institutions in the middle seems to be an alternative to assigning low scores. The correlation between the total number of clumping scores across the six institutions and the total number of scores of 0 or 1 is $-.36$ for Northern Europeans, who express substantially more confidence in institutions than other respondents; the correlations for other respondents is $-.25$.

Whither Eleven? Whither Two?

The ubiquity of clumping should be troublesome for those who advocate the replacement of the dichotomous measure with 11-point scales. The problem is not confined to 11-point scales. Clumping is common on five- and seven-point scales.

A five- point scale measuring whether “how often do you trust other people” in the 2006 ANES Pilot survey led to even more extreme clumping: Only 2.8 percent (nine) of 319 respondents replied either always or never, with the remainder in the three middle categories. Sixty percent of people who said that the trusted people half of the time said that “you can’t be too careful in dealing with people” in the 2004 ANES.⁹ And the new measure did not fare as well as the standard question in predicting (or being predicted by) charitable giving, volunteering, anti-authoritarianism, and religious fundamentalism. The five-point measure “excelled” mostly in correlations with trust in government (Uslaner, in press).

The seven-point scale, which Miller (1956) argued, marks the high end of comprehensibility for the survey respondent, has been used for many years in the ANES. The 2004 ANES included seven-point scales for preferences for less spending versus more services (for the respondent and both House candidates), for defense spending, for ideology (respondents and House candidates), and ethnic stereotypes—intelligent, hardworking, trustworthy—for whites, Asians, Hispanics, and African-Americans. Across the 20 seven-point scales, the clumping share ranged from 55 percent on the ideological placement of Republican House candidates to 82 percent for Hispanics’ intelligence. All of the stereotype measures had between 60 percent and 82 percent clumping and clumping for the four trustworthy measures ranged from 71 percent (white) to 80 percent (for the other groups). Two-thirds of white respondents clumped on eight or more of the 12 stereotype questions. Thirty seven percent of whites responded with scores of 4 on half or more of the stereotype questions.

Before we indict the seven- and eleven-point scales, we should consider two possible objections to the argument I have laid out. First, the institutional trust questions might seem more distant to people than other issues—and thus should be more prone to clumping. There is some support for this claim, but it is modest. In the ESS, the “satisfaction” questions—with life, one’s education and health, the economy, the government, and democracy are all 11-point scales. Most people are satisfied with their lives (with a mean score of 7.1 out of 10) and only 24 percent of respondents give responses of 4, 5, or 6. Yet, 38 percent clump on satisfaction with health and 41 percent on satisfaction with their education, compared to 40 percent for satisfaction with the economy, 41 percent for democracy, and 44 percent for the government—and means around 40 percent for most institutional trust measures. Your health and your education are hardly more

“distant” than your government, yet the clumping shares are similar.

Second, clumpers may truly be people in the middle. They might be moderates, who give centrist responses to all sorts of questions. There is little support for this argument either. The simple correlations for the clumping on trust in the ESS and a folded measure of ideology (with strong ideologues on the right or left at zero and centrists at five) range from .027 for trust in the police to .065 for trust in parliament. For the 2004 ANES, there are, of course, substantial correlations between a folded measure of ideological strength and clumping on ideology. There are moderate correlations (about .2) for spending measures. However, the stereotype measures had correlations of virtually zero with the stereotype questions: The summary measure of clumping on stereotypes (ranging from zero to 12) has a correlation of .016 with the measure of ideological moderation. Neither proximity nor moderation explains clumping.

Reprise

Alwin and Krosnick (1991, 164) investigated the reliability of scales with between two and nine response categories. While in general they found that reliability increased up to seven points—with no significant differences between seven- and nine-point measures—they found a notable exception:

The 2-category scales are a major exception to this pattern, as they have relatively reliable responses....2-category questions unambiguously measure the direction of attitudes only, with no pretense of measuring intensity, whereas 4 and more category response scales presumably are intended to measure both direction and intensity. The direction of attitude responses may...be more reliably assessed than the intensity of responses....

Intensity may be easier to measure on ideology or policy alternatives, where people may have clear and fixed views. However, where attitudes such as confidence in institutions may change overtime depending upon the state of the economy, the quality of leaders, or who is in power (Anderson et al., 2005; Citrin, 1974), intensity may be more difficult to estimate precisely.

Generalized trust is more stable. Yet the calculation of how widespread trust is—especially when the zero and 10 reference points are both so extreme—may demand more precision than people can or are willing to make.

The standard question as a dichotomy is far from perfect. Its problems have been discussed at length (see Smith, 1997, for one of the best critiques, but also see Uslaner, 2002, 68-74 for a response). Yet the superiority of an 11-point scale has not been established. Despite criticisms, the dichotomous measure is well understood when people are asked what it means and it is stable both over time when asked in panels and from parents to children (Uslaner, 2002, 68-74). The search for an ideal measure of trust goes on, but with the caveat that any new measure will sacrifice comparability with almost four decades of using the standard question.

FIGURE 1

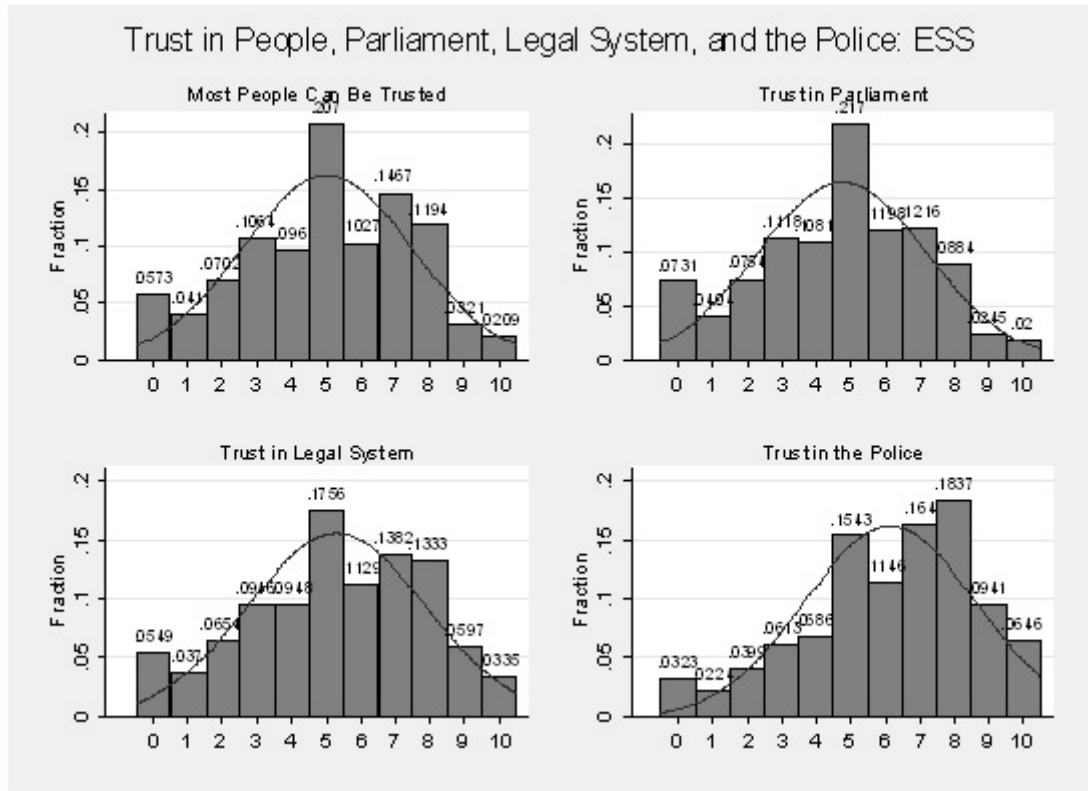


FIGURE 2

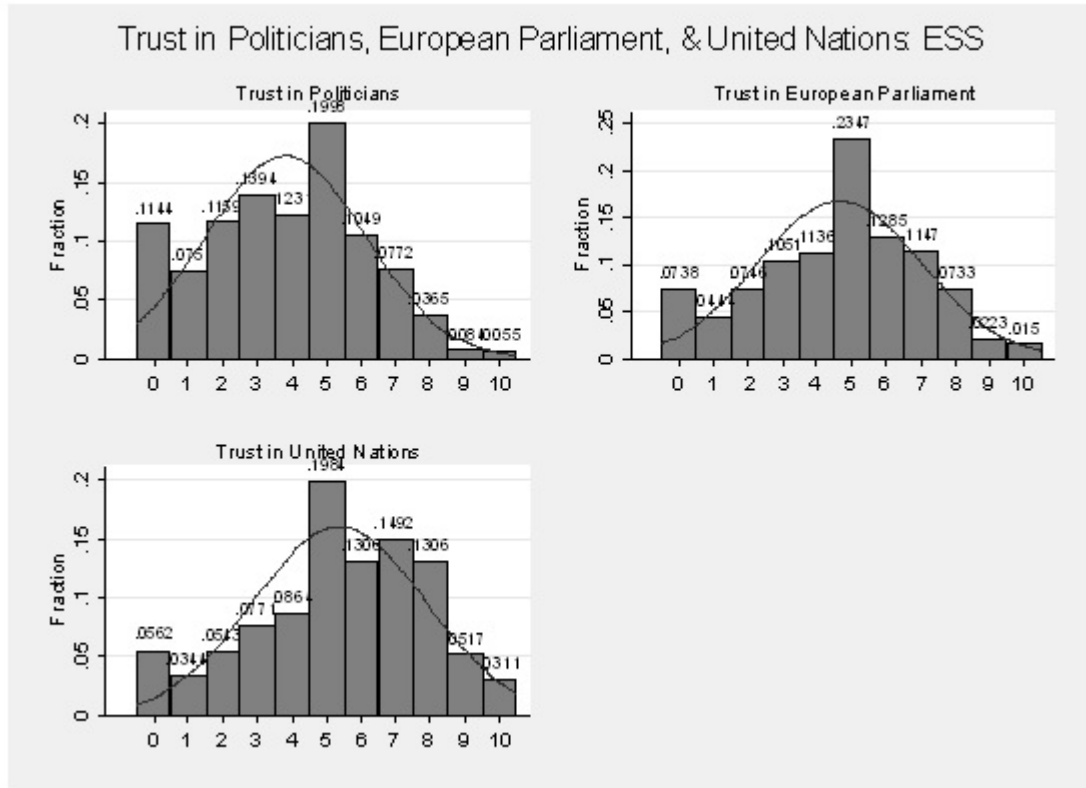


FIGURE 3

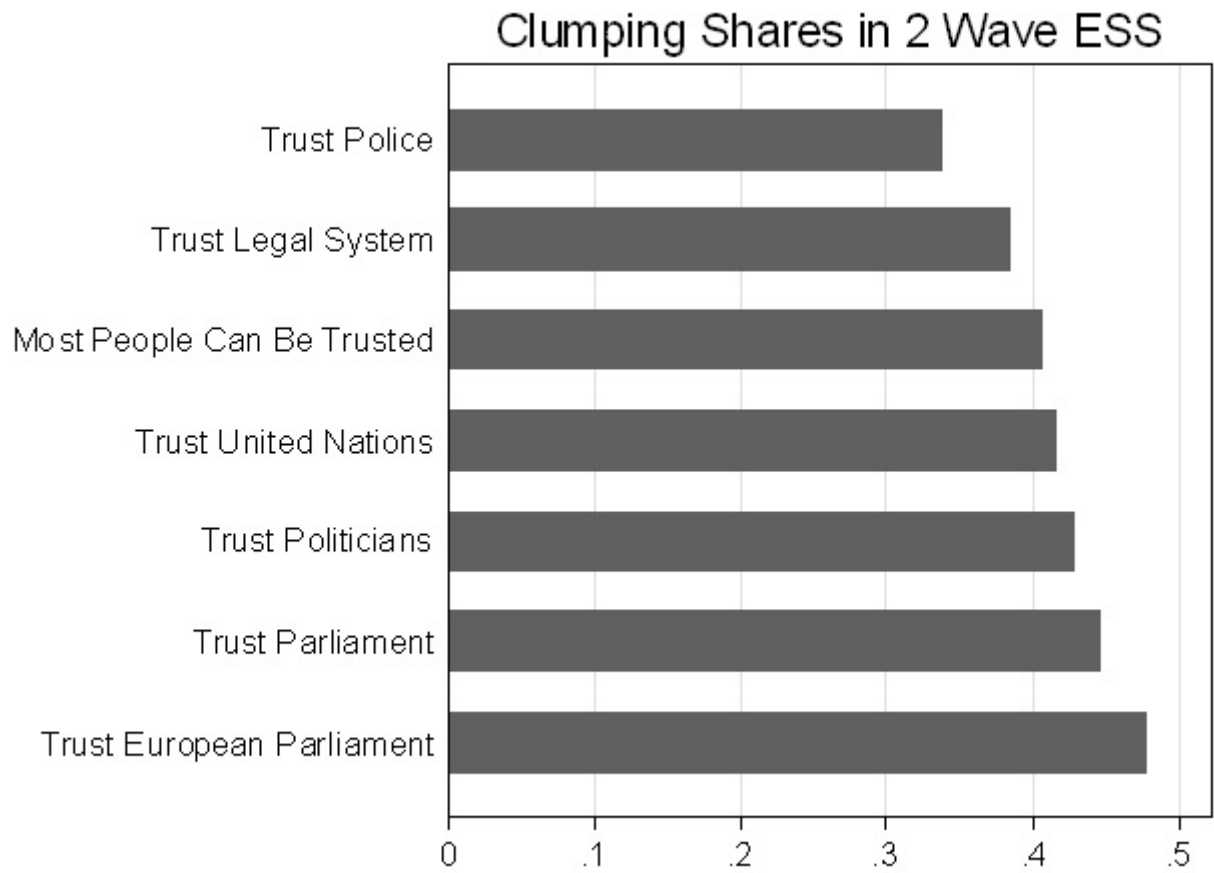


FIGURE 4

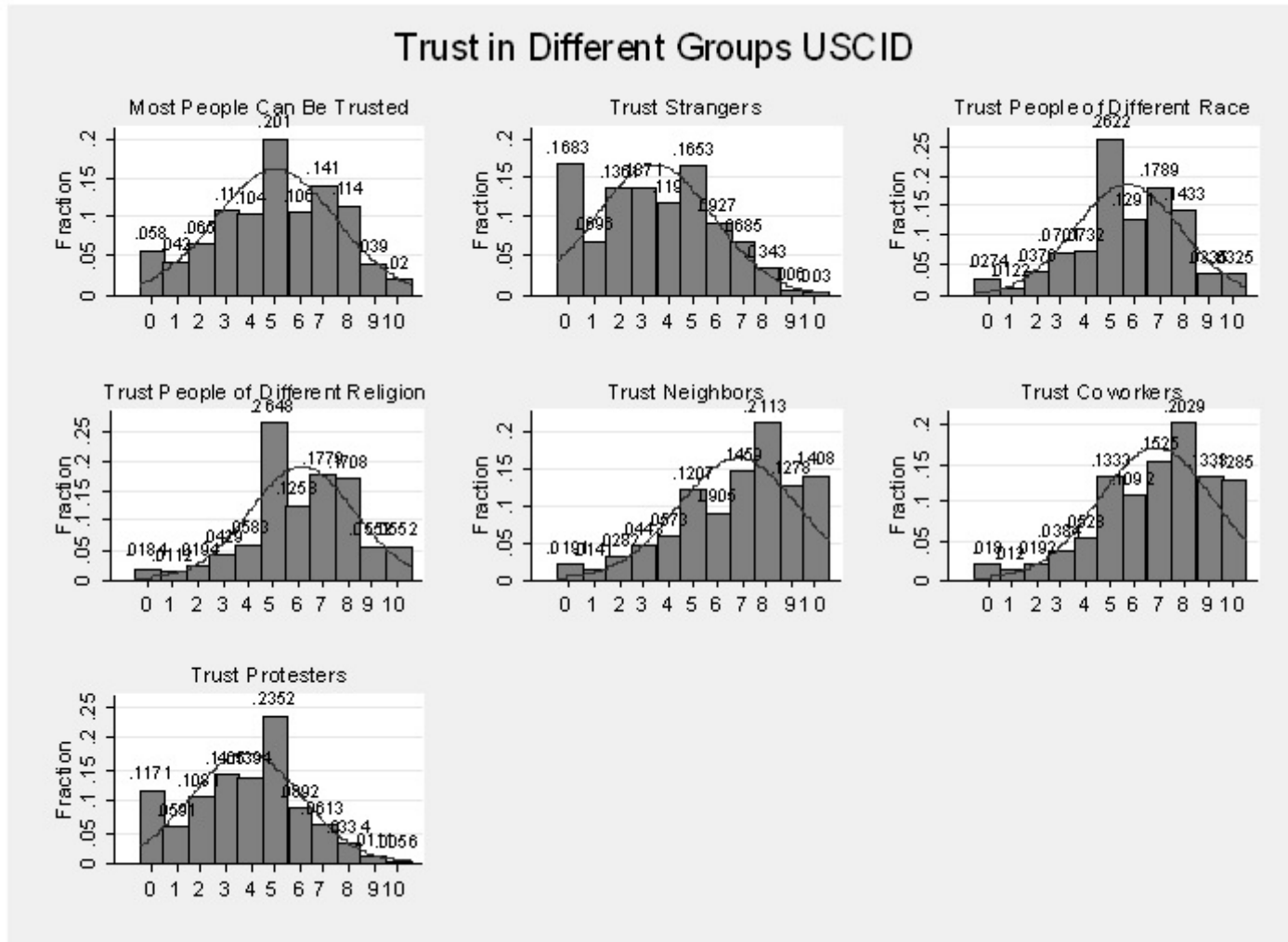


FIGURE 5

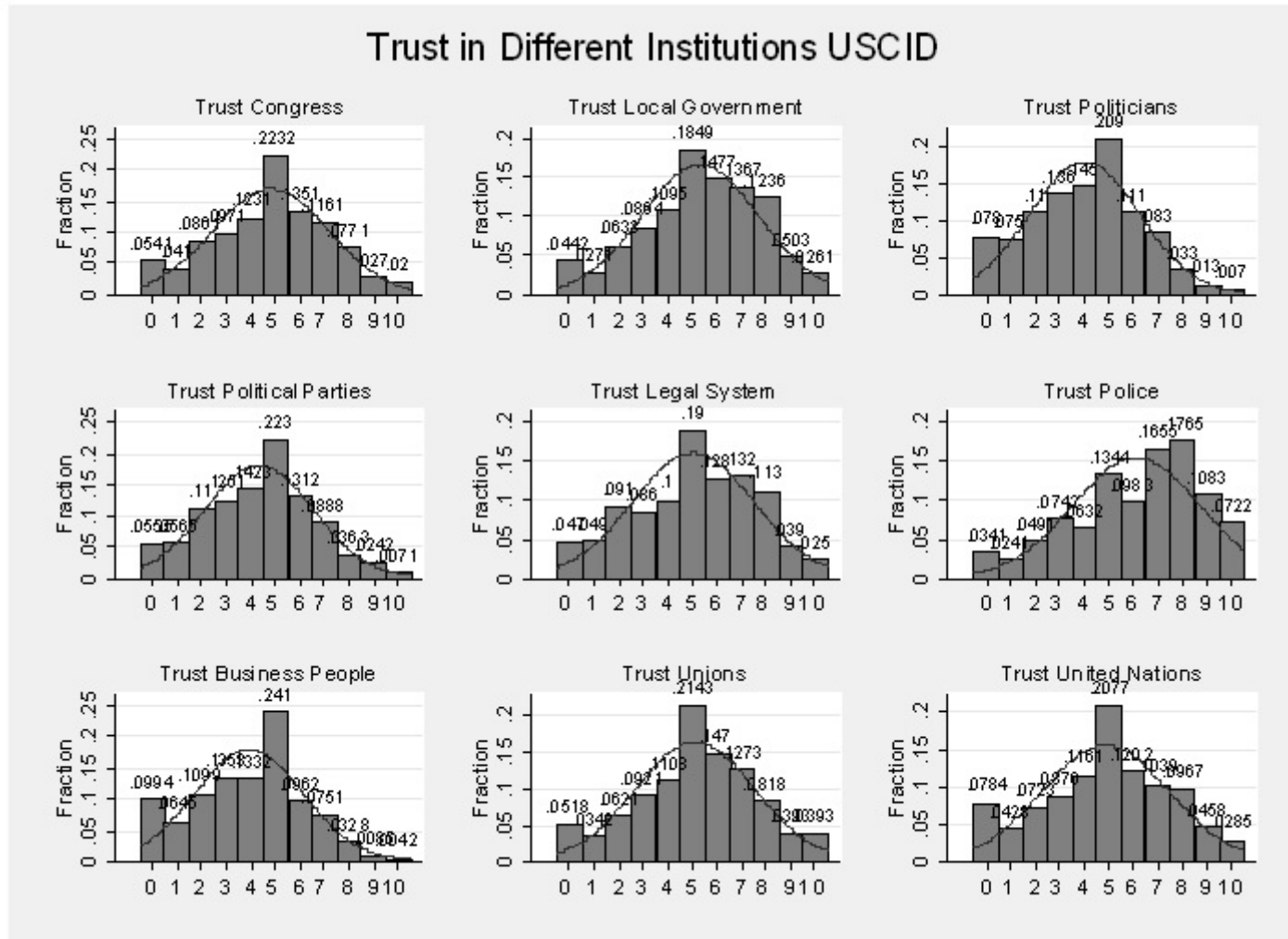


FIGURE 6

Clumping Proportions in US Citizenship, Involvement, Democracy Survey



FIGURE 7

Clumping in European Samples: Generalized Trust and Trust in Politicians

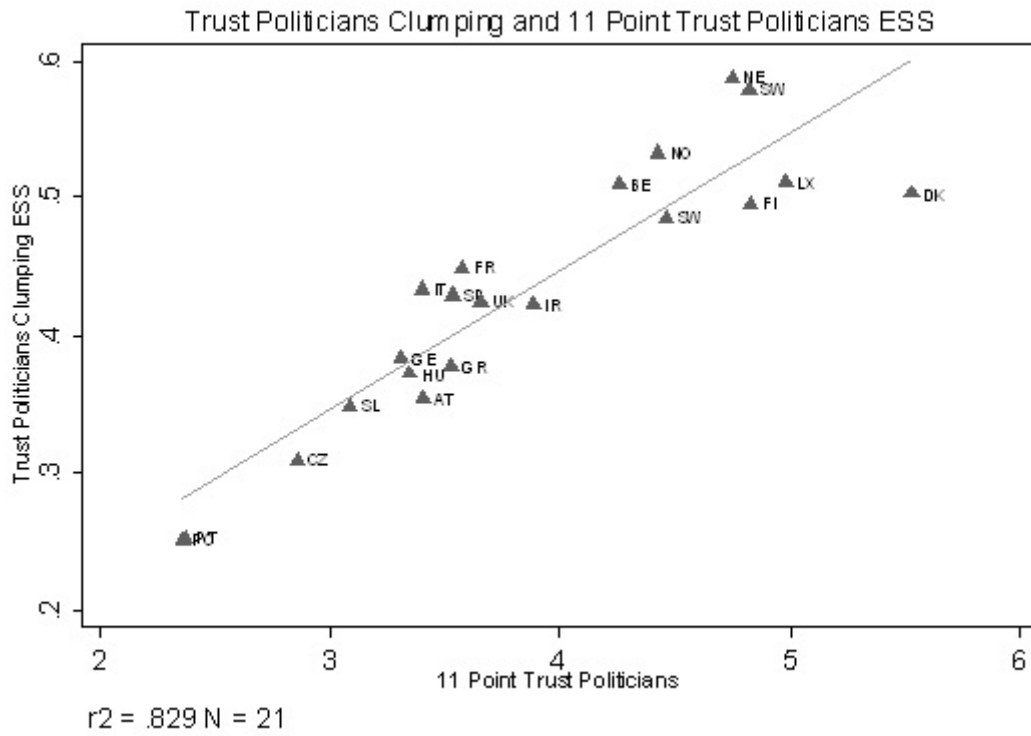
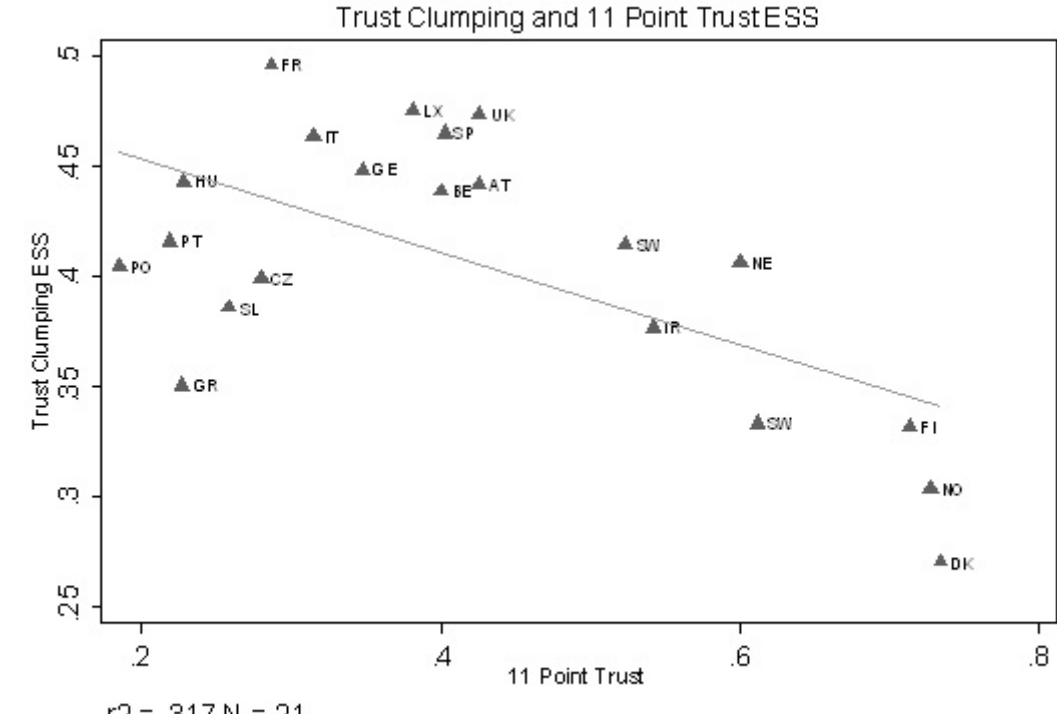


FIGURE 8

11 Point Trust Versus Dichotomy: Romanian CID

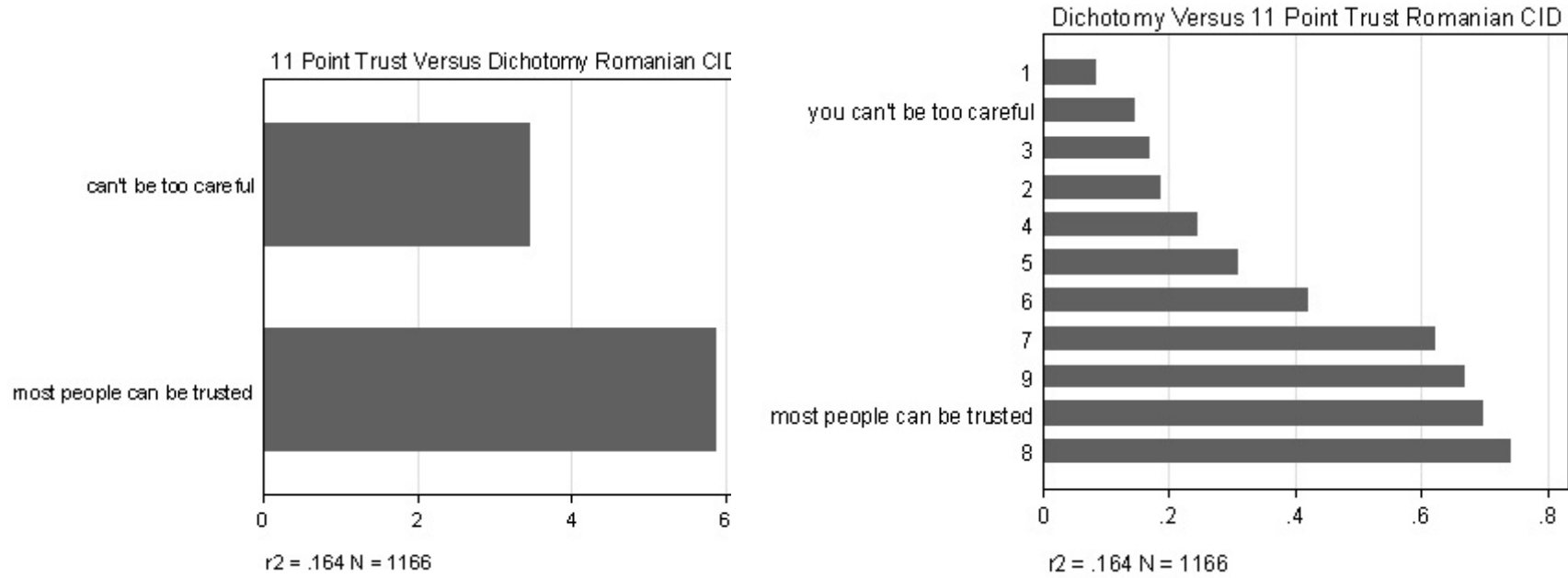


TABLE 1

The Dichotomy and the 11-Point Scale in the Romanian CID (By Row)

	11-Point Scale											Total
	0	1	2	3	4	5	6	7	8	9	10	
Can't be too careful	85.6	91.6	81.6	83.2	75.8	69.3	58.0	37.8	26.0	33.3	30.4	783
	(137)	(65)	(71)	(129)	(91)	(156)	(47)	(42)	(20)	(11)	(14)	(67.2)
Most people can be Trusted	14.4	8.5	18.4	16.8	24.2	30.7	42.0	62.2	74.0	66.7	69.6	32.9
	(23)	(6)	(16)	(26)	(29)	(69)	(34)	(69)	(57)	(22)	(32)	(383)
N	160	71	87	155	120	225	81	111	77	33	46	1166

TABLE 2

The Dichotomy and the 11-Point Scale in the Romanian CID (By Column)

Trust Scale	Careful	Trusted	Total
0	17.5 (137)	6.0 (23)	13.7 (160)
1	8.3 (65)	1.6 (6)	6.1 (71)
2	9.1 (71)	4.2 (16)	7.5 (87)
3	16.5 (129)	6.8 (26)	13.3 (155)
4	11.6 (91)	7.6 (29)	10.3 (120)
5	19.9 (156)	18.0 (69)	19.3 (225)
6	6.0 (47)	8.9 (34)	7.0 (81)
7	5.4 (42)	18.0 (69)	9.5 (111)
8	2.6 (20)	14.9 (57)	6.6 (77)
9	1.4 (11)	5.7 (22)	2.8 (33)
10	1.8 (14)	8.4 (32)	4.0 (46)
Total	783	383	1166

TABLE 3

Average Correlations Among Trust Measures by Clumping: 2 Wave ESS

Variable	Average correlation: Clumping	Average correlation: Not Clumping
Generalized Trust	.066	.410
(minimum, maximum)*	(.054, .078)	(.286, .488)
Trust in Parliament	.167	.670
(minimum, maximum)	(.063, .298)	(.488, .822)
Trust in Politicians	.175	.614
(minimum, maximum)	(.076, .298)	(.447, .822)
Trust in Legal System	.164	.637
(minimum, maximum)	(.062, .307)	(.458, .772)
Trust in Police	.145	.528
(minimum, maximum)	(.078, .307)	(.382, .744)
Trust in European Parliament	.176	.584
(minimum, maximum)	(.054, .370)	(.287, .762)
Trust in United Nations	.165	.576
(minimum, maximum)	(.060, .370)	(.399, .762)
N	12,170 - 20,587	22,698 - 32,844

Minimum correlation for institutional measures always with generalized trust.

TABLE 4

Seemingly Unrelated Probit of Clumping for Generalized Trust and Ordinary Least Squares Regression
for Institutional Trust Clumping Totals : ESS 2 Wave

Variable	Generalized Trust Clumping Probit				Institutional Trust Total Clumping: OLS		
	Coefficient	S.E.	z	Effect	Coefficient	S.E.	t
Life Satisfaction	-.013*	.010	-1.30	-.049	.002	.014	.17
Household income	-.006	.006	-.93	-.025	.020**	.011	1.86
Education	-.007**	.004	-1.70	-.111	.003	.005	.62
Age	-.004*****	.001	-7.62	-.125	-.005*****	.001	-5.79
Hours reading paper	.014	.008	1.89	.039	.013	.012	1.06
Internet frequency	-.013***	.005	-2.78	-.036	-.001	.006	-.27
Vote	-.050*	.030	-1.64	-.019	.182*****	.048	3.76
Interest in politics	.020	.013	1.60	.024	.051***	.019	2.68
Transition country	-.016	.040	-.40	-.006	-.162**	.081	-1.99
Constant	.262**	.128	2.04		2.422*****	.145	16.64
N				50107			50107
-2 * Log Likelihood				66751.086			R ² = .012 F = 29.94 RMSE = 1.6164

Generalized trust and trust in parliament equations estimated by seemingly unrelated probit; effects estimated from simple probits; standard errors clustered by country for both probit and OLS.

For generalized trust probit: McKelvey-Zavoina R² = .080, Percent predicted correctly: 61 (model), 61 (mean)

* p < .10 ** p < .05 *** p < .01 **** p < .0001 (all tests one tailed except for constants)

REFERENCES

- Alwin, Duane .F. and Jon A. Krosnick. (1991). "The Reliability of Survey Attitude Measurement: The Influence of Question and Respondent Attributes," Sociological Methods and Research, 20:139-181.
- Anderson, Christopher, Andre Blais, Shaun Bowler, Todd Donovan, and Ola Listhaug. 2005. Losers' Consent: Elections and Democratic Legitimacy. New York: Oxford University Press.
- Citrin, Jack. 1974. "Comment: The Political Relevance of Trust in Government," American Political Science Review, 68:973-88.
- Claibourn, Michele P. and Paul S. Martin. 2004. "Trusting and Joining? An Empirical Test of the Reciprocal Nature of Social Capital," Political Behavior, 22:267-291.
- Goodwin, Jason. 2000. On Foot to the Golden Horn. New York: Henry Holt.
- Hooghe, Marc and Tim Reeskens. 2007. "To Trust or Not To Trust: That's The Question: The Measurement of Generalized Trust as a Dichotomous or Continuous Variable." Presented at the 2nd Conference of the European Survey Research Association, Prague, June.
- Miller, George A. 1956. "The Magic Number 7, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Reports, 63:81-97, available at <http://www.well.com/user/smalin/miller.html>.
- Newton, Kenneth. 1999. "Social and Political Trust." In Pippa Norris, ed., Global Citizens: Global Support for Democratic Government. New York: Oxford University Press.
- Putnam, Robert D. 1993. Making Democracy Work: Civic Traditions in Modern Italy.

- Princeton: Princeton University Press.
- Scherpenzeel, Annette. N.d. "Why use 11-point scales?", at <http://www.swisspanel.ch/file/doc/faq/11pointscales.pdf>
- Scherpenzeel, Annette.C. and W.E. Saris. 1995. "The Quality of Indicators of Satisfaction Across Europe: A Meta-analysis of Multitrait-multimethod Studies." In Annette C. Scherpenzeel, A Question of Quality. Evaluating Survey Questions by Multitrait-multimethod Studies. Dissertation. Royal PTT Nederland NV, KPN Research.
- Schwartz, Barry. 2003. The Paradox of Choice: Why More Is Less. New York: Ecco.
- Smith, Tom W. 1997. "Factors Relating to Misanthropy in Contemporary American Society," Social Science Research, 26:170-96.
- Stolle, Dietlind. 1998. "Bowling Together, Bowling Alone: The Development of Generalized Trust in Voluntary Associations," Political Psychology, 19:497-526.
- Uslaner, Eric M. 2002. The Moral Foundations of Trust. New York: Cambridge University Press.
- _____. 2003. "Trust and Civic Engagement in East and West." In Gabriel Badescu and Eric M. Uslaner, eds., Social Capital and the Transition to Democracy. London: Routledge.
- _____. In press. "The Generalized Trust Questions in the 2006 ANES Pilot Study." In Kathleen M. McGraw and John Aldrich, eds., The Book of Ideas.
- van Deth, Jan W., José Ramón Montero, and Anders Westholm. 2008. Citizenship and Involvement in European Democracies: A Comparative Analysis. London: Routledge.
- Zmerli, Sonja and Ken Newton. 2008. "Social Trust and Attitudes Towards Democracy," Public

Opinion Quarterly, 72:706-724.

ENDNOTES

- * I am grateful for the support of the General Research Board of the University of Maryland–College Park and for helpful discussions with and comments on an earlier version by Mike Hanmer, Marc Hooghe, Kerem Ozan Kalkan, Geoff Layman, Ken Newton (who raised the two key objections discussed), and Tim Reeskens. I am also grateful to my colleagues on the Romanian CID team, Gabriel Badescu and Paul Sum and to the Starr Foundation through the Institutional Researches and Exchanges Board of the United States Department of State for a grant under the IREX Caspian and Black Sea Collaborative Program (2001) to collect the data. I am also indebted to Marc Morje Howard of Georgetown University for the opportunity to work with the United States CID data.
1. For the Swiss Household Panel, see <http://www.swisspanel.ch/>. For the history of the CID project, see <http://www.mzes.uni-mannheim.de/projekte/cid/> and <http://www.mzes.uni-mannheim.de/projekte/cid/background.htm> . See Van Deth, Montero, and Westholm (2008) for essays based upon the public surveys.
 2. See <http://www.europeansocialsurvey.org/>.
 3. Tables in the Hooghe and Reeskens (2007) paper indicate that they are using two-tailed tests of significance. One-tailed tests are more appropriate, so immigrant status would be a significant predictor of dichotomous trust.
 4. Data and documentation are available at <http://www.uscidsurvey.org/>.
 5. I was part of the Romanian CID team and wanted to make sure that we had data using the standard question. As one might expect for a low trusting society, clumping is modest in

Romania ranging from 25 percent for trusting politicians to 32 percent for confidence in service providers. The Romanian survey was conducted in 2001 by Gabriel Badescu at Babes-Bolyai University with 1217 respondents. The dichotomous trust question was the first question asked in the survey, with the 11-point scale asked shortly thereafter in accordance with the CID protocol (see n. 1). The Romanian team was also responsible for the Moldovan survey (Moldovians are mostly ethnic Romanians). I learned of the Spanish survey from Ken Newton. The data have not been publicly released but I have access to them as a member of the Romania-Moldova team.

6. There are no measures of trust in government other than the 11-point scales in the Romanian CID.
7. In the sense that most people don't have contact with it or follow it.
8. For probit, I estimated the effects by manipulating each predictor in turn while the remaining variables retained their original values.
9. The 2006 Pilot was a panel with the 2004 ANES. Jon Krosnick and Arthur Lupia, the principal investigators of the ANES, contacted me about the forthcoming collaboration with the National Longitudinal Survey of Youth (NLSY) that will link young people's responses to the trust question with their parents' responses (cf. Uslander, 2002, 98-107, 162-165). The NLSY uses the five-point scale and the ANES was preparing to switch. I argued for the standard question and was asked to prepare the experiment reported in Uslander (in press) in which both the dichotomy and the five-point scale were asked of 319 respondents.