Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology?

Scott Clifford

Ryan M. Jewell

Philip D. Waggoner

Abstract

Amazon's Mechanical Turk (MTurk) is an increasingly popular tool for the recruitment of research subjects. While there has been much focus on demographic differences between MTurk samples and the national public, we know little about whether liberals and conservatives recruited from MTurk share the same psychological dispositions as their counterparts in the mass public. In the absence of such evidence, some have argued that the selection process involved in joining MTurk invalidates the subject pool for studying questions central to political science. In this paper, we evaluate this claim by comparing a large MTurk sample to two high-quality national samples – one conducted online and one conducted face-to-face. We examine the personality and value-based motivations of political attitudes and identities across the three samples. All three samples produce substantively identical results with only minor variation in effect sizes. In short, liberals and conservatives in our MTurk sample closely mirror the psychological divisions of liberals and conservatives in the mass public.

Researchers are increasingly turning to Amazon's Mechanical Turk (MTurk) to recruit subjects for public opinion research (e.g., Ahler 2014; Arceneaux 2012; Clifford 2014; Gerber et al. 2012; Grimmer, Messing, and Westwood 2012; Huber and Paris 2013; Johnston, Newman, and Velez 2015). MTurk allows the rapid recruitment of a diverse sample of subjects at a dramatically lower cost than professional online panels (Berinsky et al. 2012). A variety of experiments conducted on nationally representative samples have been replicated on MTurk, providing evidence for its validity as a recruitment tool for experimental subjects (Berinsky et al. 2012; Weinberg et al. 2014).

While there are clear benefits to the use of MTurk, researchers have raised a number of concerns about its use (e.g., Chandler, Mueller, and Paolacci 2014; Krupnikov and Levine 2014). Much of the debate over the validity of MTurk as a recruitment tool has followed the discipline's "near obsession" with the external validity of a sample (McDermott 2002, 334). Researchers consistently find that MTurk samples tend to be more politically liberal, younger, less religious, and less racially diverse than the population (Berinsky et al. 2012; Huff & Tingley 2014).

While scholars have documented many of the differences and similarities between MTurk samples and the national public, some critics maintain that the sample is invalid for political research.[1] According to this view, the selection process that produces disproportionately liberal samples implies that the conservatives who opt into MTurk differ from other conservatives in psychological dispositions central to their identities. If this claim were true, it may render MTurk samples invalid for studying political and ideological divides. This would be particularly worrisome for research using ideology or partisanship as a moderator of

---

[1] For example, see recent discussion at the Cultural Cognition Project and Washington Post's *Monkey Cage* (Searles & Ryan 2015; Kahan 2013).

experimental treatment effects (e.g., Bullock 2011; Druckman, Peterson, and Slothuus 2013; Nyhan and Reifler 2010) or examining psychological differences between liberals and conservatives (e.g., Feldman and Johnston 2014). Yet, we have little systematic evidence for whether the selection process generates samples that are valid for studying ideological divisions in the mass public.

In this paper, we first review the literature on the psychological motivations behind participation in online panels and political identity. We then compare psychological models of political ideology across American National Election Studies (ANES) and MTurk samples. We find that conservatives recruited from MTurk look nearly identical to their counterparts in national samples, though MTurk liberals tend to have more liberal dispositions than ANES liberals. We next model political ideology as a function of personality and values and find strikingly similar results across samples. Overall, our findings suggest that research into the psychological dispositions behind political ideology would reach largely the same conclusions regardless of the sample chosen.

Psychological Differences between Liberals and Conservatives

Liberals and conservatives differ in a variety of psychological dispositions, most notably in their personality traits and value orientations. Decades of research has established the Big Five personality traits – Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability – as a useful framework for understanding stable dispositions that underlie individual behavior (Costa & McCrae 1992). Two of the Big Five traits, in particular, are consistently strong predictors of ideology. Liberals tend to score higher in Openness, demonstrating greater interest in new information and ideas (Gerber et al. 2013). Conservatives

tend to score higher in Conscientiousness, demonstrating greater dependability and self-discipline (Gerber et al. 2013). Findings for other traits are less consistent, but there is some evidence that Extraversion is associated with conservatism (e.g., Gerber et al. 2010; Mondak and Halperin 2008), and that Agreeableness predicts liberalism (Mondak 2010).

Liberals and conservatives differ even more strongly in their values. We focus on four values commonly used in research on political attitudes and identities. The first three values represent the preference for equal opportunity (egalitarianism), traditional family structures (moral traditionalism), and conformity and obedience (authoritarianism). Last, racial resentment represents a blend of racial prejudice and the values of individualism and self-reliance (Henry & Sears 2002).

Psychological Motivations for Selection into Online Panels

The literature on the psychological motivations for participating in online panels is still nascent, but there is some evidence that some of the same dispositions that predict political liberalism also predict participation in online panels. Agreeableness consistently predicts participation in web panels and online surveys, while Conscientiousness, Openness, and Extraversion have also been found to predict survey participation (Brüggen & Dholakia 2010; Rogelberg et al. 2003; Dollinger & Leong 1993).

Though we are aware of no study analyzing the determinants of participation on MTurk, MTurk samples tend to differ from other samples in personality and other dispositions. Compared to an adult community sample, MTurk workers have lower self-esteem, lower levels of Extraversion, Emotional Stability, and Openness (Goodman et al. 2013). MTurk subjects also tend to score higher in Need for Cognition and Need to Evaluate than national samples (Berinsky

4

et al. 2012). In short, there is some evidence that MTurk samples differ from other populations on personality traits and other psychological dispositions.

We have little evidence, however, regarding how the self-selection process might affect political divides on MTurk. Scherer, Windschitl, and Graham (2014) find the same partisan differences in System Justification in an MTurk sample and a nationally representative sample. Grimmer, Messing, and Solomon (2012) find that the associations between partisanship, ideology, and feelings towards Obama are substantively identical across MTurk and nationally representative samples. However, no work has systematically investigated whether liberals and conservatives on MTurk share the same personality traits and values as their counterparts in the mass public.

Data and Methods

The data for our analyses consists of the ANES 2012 Time Series Study conducted before and after the 2012 United States presidential election, and a survey fielded on MTurk in June 2015. The ANES 2012 Time Series Study recruited 1,413 respondents for face-to-face interviews (FTF) and 3,860 respondents for a web-based survey (Web). The FTF sample was collected using computer assisted self-interviewing and relied on an address-based sampling frame. The Web sample was recruited from GfK Knowledge Networks' address-based sampling frame. Our MTurk survey was posted online and described as a "Short Political Survey." The HIT was made available to 1,500 U.S. residents with an acceptance rate of at least 95%. Subjects were paid $0.40 for completion.[2] Demographics are shown in Table A1 in the Appendix.

---

[2] This study was approved by [university omitted] Committee for the Protection of Human Subjects (Application #15560-EX – (6556)).

Measures

All of our measures are drawn directly from the ANES, so we discuss them only briefly here (see Appendix for full question wording).[3] Political ideology and partisanship are both measured using standard seven-point scales. Social and economic ideology are each measured as indices of policy attitudes following the approach of Feldman and Johnston (2014). All dependent variables are coded such that higher values are more conservative.

We measure the Big Five personality traits using the Ten-Item Personality Inventory (TIPI; Gosling, Rentfrow, and Swann 2003). Moral traditionalism, racial resentment, and egalitarianism are measured as the average agreement with 4-6 statements. Authoritarianism is measured using the child traits battery (Feldman & Stenner 1997). Our demographic controls include age, education, gender, income, and religiosity (following Feldman and Johnston 2014).
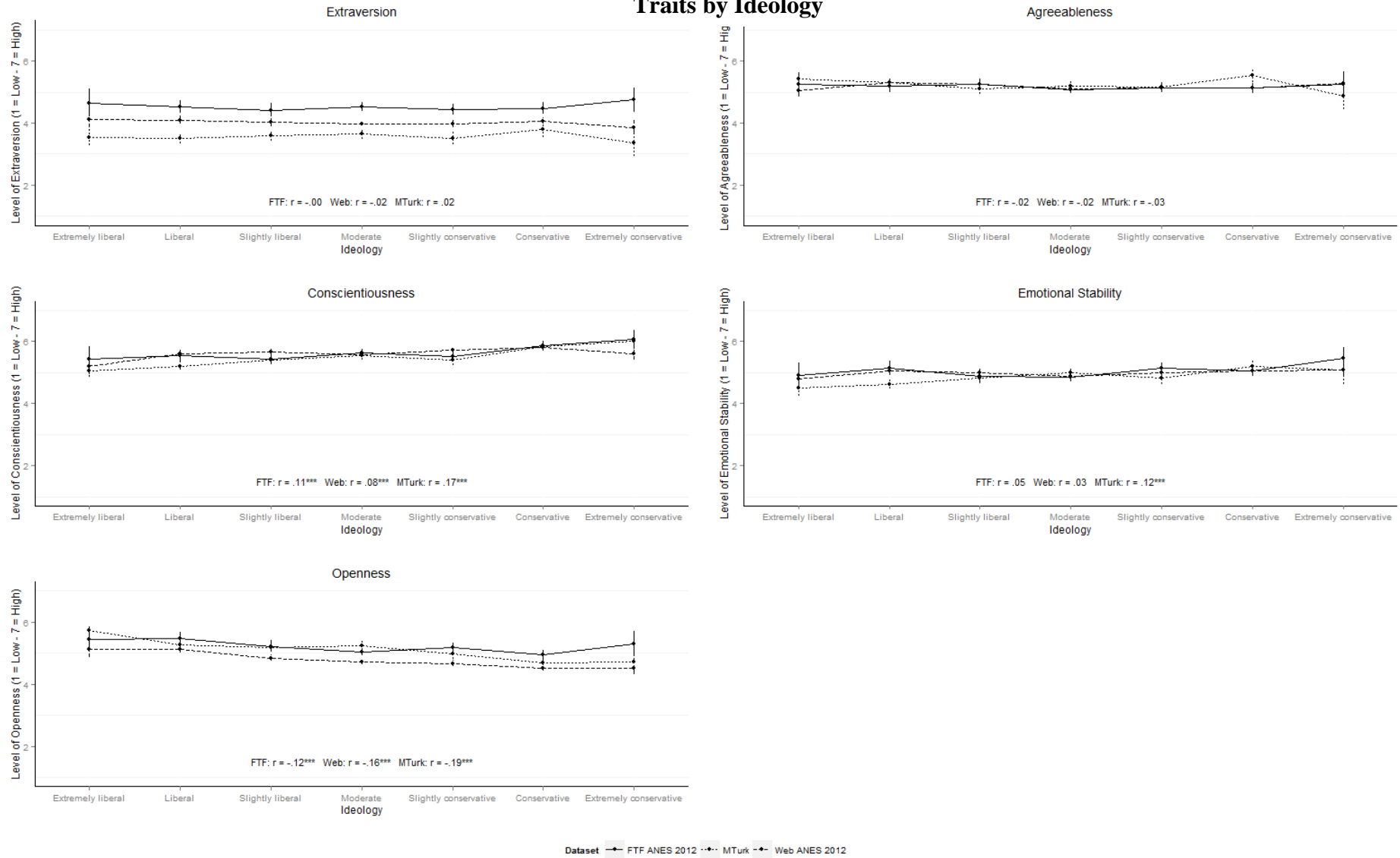
Results

We begin by graphically displaying the levels of each independent variable across each measure of ideology. Results for partisanship were substantively identical and are shown in the Appendix (Figure A1-3). Figure 1 displays the average level of each personality trait across self-placement ideology for each sample (ANES FTF, ANES Web, and MTurk). Means and sample differences are shown in Table A2 in the Appendix.

Starting with the top left panel, Extraversion is uncorrelated with ideology across all samples. However, ANES FTF respondents are significantly more extraverted than both ANES Web and MTurk samples, and MTurk subjects are less extraverted than the FTF sample (Cohen's $d = .59$) and the Web sample ($d = .30$). Turning to the top right panel, Agreeableness is not significantly related to political ideology in any of the samples, nor are there large

---

[3] We edited wording of some questions to be appropriate for the survey mode.

## Figure 1: Sample Differences in Personality Traits by Ideology

Extraversion

FTF: r = -.00   Web: r = -.02   MTurk: r = .02

Agreeableness

FTF: r = -.02   Web: r = -.02   MTurk: r = -.03

Conscientiousness

FTF: r = .11***   Web: r = .08***   MTurk: r = .17***

Emotional Stability

FTF: r = .05   Web: r = .03   MTurk: r = .12***

Openness

FTF: r = -.12***   Web: r = -.16***   MTurk: r = -.19***

Dataset → FTF ANES 2012 ···•··· MTurk -•- Web ANES 2012

*p<0.05; **p<0.01; ***p<.001

7

differences in means between the samples. Conscientiousness is positively correlated with conservatism across all three samples (all ps < .01). Emotional Stability is positively correlated with ideology in all samples, but this relationship is statistically significant only in the MTurk sample (p < .01). Finally, Openness is negatively correlated with ideology in all samples (all ps < .01). The MTurk sample scores notably higher in Openness than the Web sample (d = .39), but similar to the FTF sample (d = .09). Overall, correlations between the Big Five traits and political ideology are substantively identical across samples with the exception of Emotional Stability. The MTurk sample is also lower in Extraversion than both ANES samples.

Figure 2 displays the mean level of each value orientation across self-reported ideology. Authoritarianism, racial resentment, and moral traditionalism are positively associated with conservatism across all three samples (all ps < .01). While conservatives are nearly indistinguishable across samples, MTurk liberals score notably lower in each value than ANES liberals. These differences may be due, in part, to lower levels of religiosity (see Table A3 in the Appendix). Egalitarianism is negatively correlated with ideology across all samples (ps < .01), though MTurk liberals score slightly higher than ANES liberals. Overall, we find the same relationships between values and political ideology, and conservatives are nearly indistinguishable across samples. However, we find a stronger relationship between values and political ideology in the MTurk sample, with liberals taking more characteristically liberal positions.
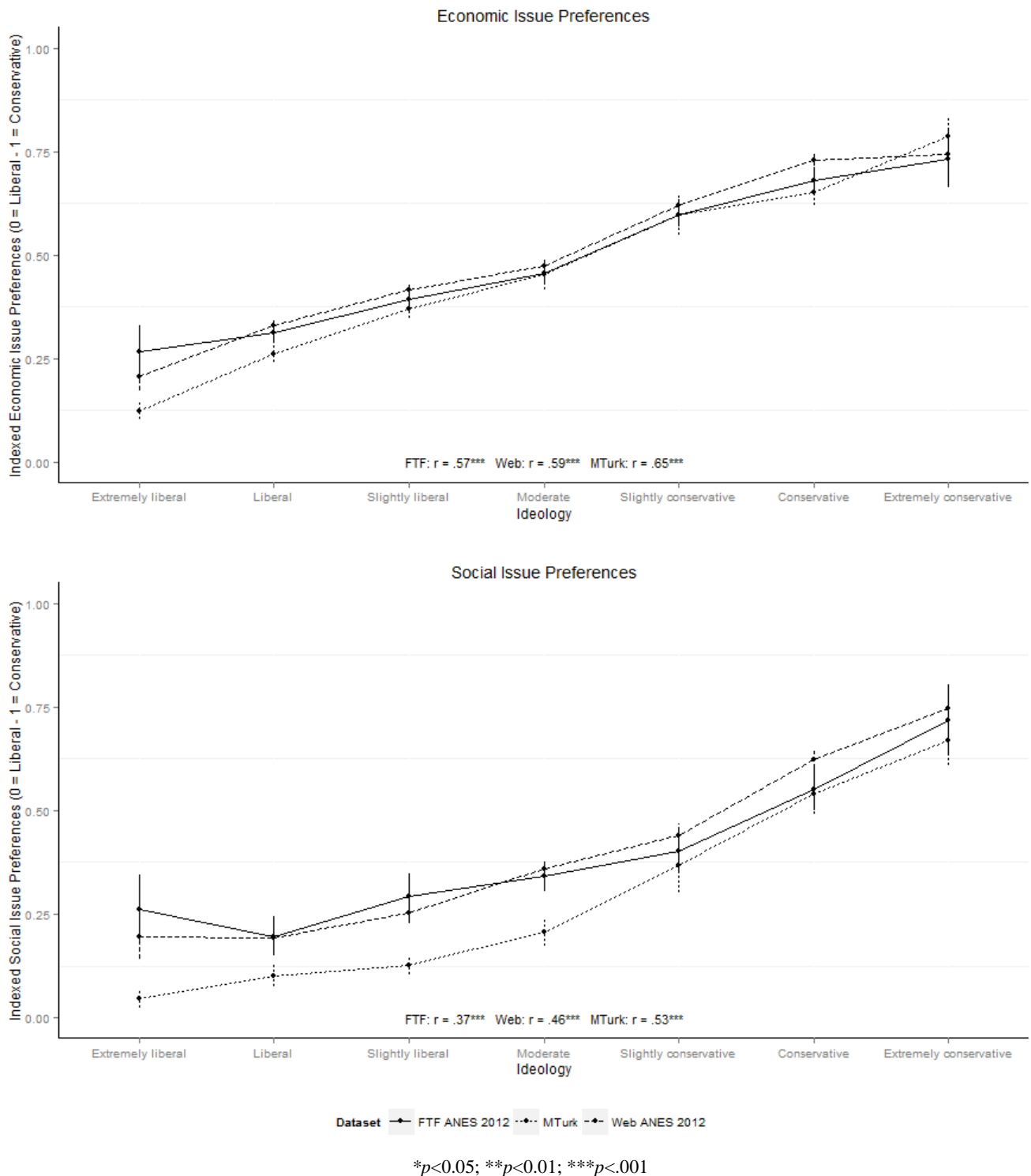
Figure 3 displays social and economic ideology as a function of self-reported ideology. Ideological subgroups are highly similar across samples, though MTurk liberals have more liberal social preferences and slightly more liberal economic preferences, consistent with the results above.

Figure 2: Sample Differences in Values by Ideology
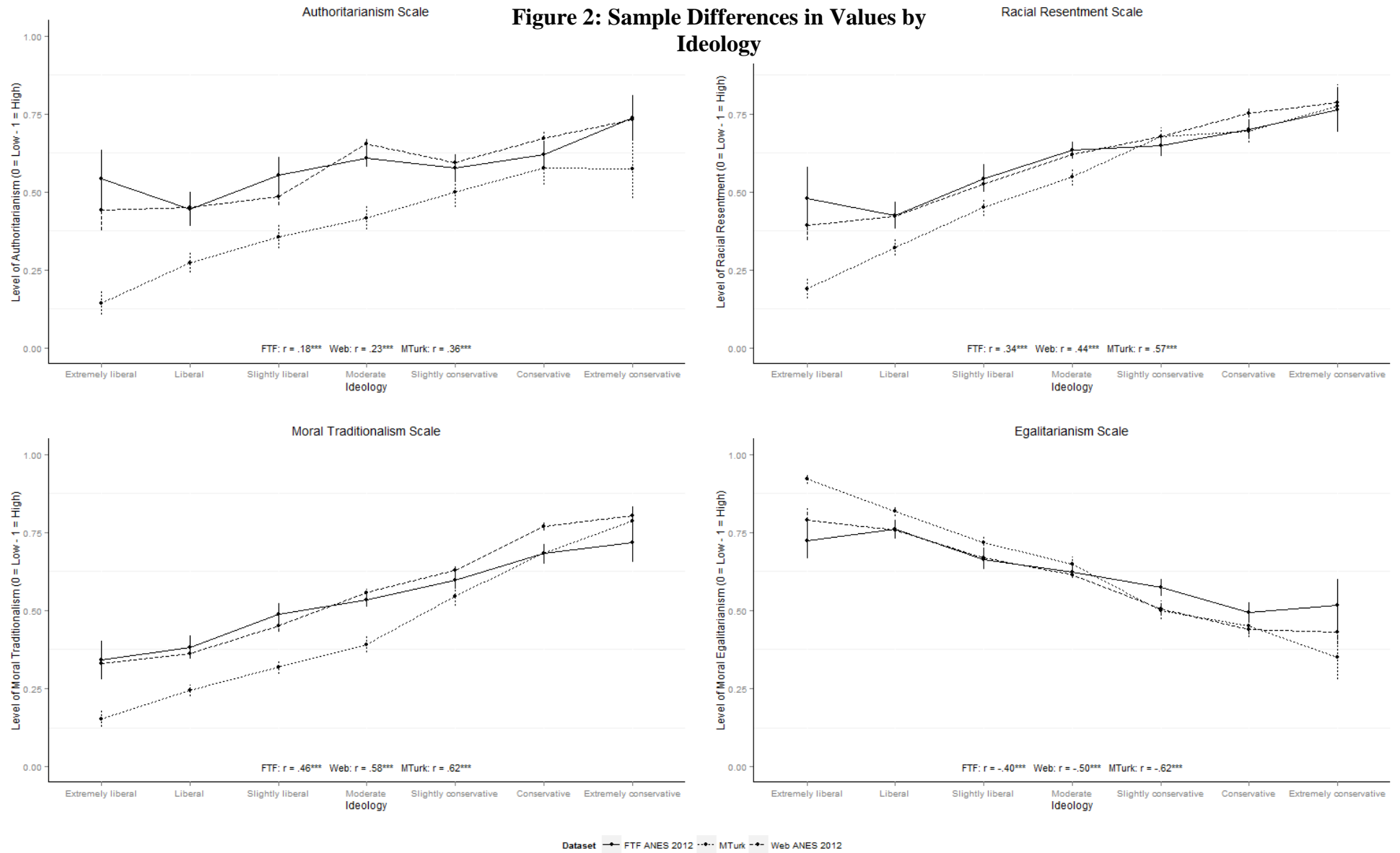
*p<0.05; **p<0.01; ***p<.001

Figure 3: Sample Differences in Economic and Social Issue Preferences by Ideology

Economic Issue Preferences

FTF: r = .57***   Web: r = .59***   MTurk: r = .65***

Social Issue Preferences

FTF: r = .37***   Web: r = .46***   MTurk: r = .53***

Dataset ─●─ FTF ANES 2012 ·····●····· MTurk ─●─ Web ANES 2012

*p<0.05; **p<0.01; ***p<.001

Psychological Models of Political Ideology

We now move away from descriptive statistics and ask whether researchers investigating

the psychological antecedents of political ideology would reach the same conclusions if they

relied on an MTurk sample rather than an ANES sample. Below we present six OLS models,

each predicting one of three dependent variables (self-reported ideology, social ideology, and

economic ideology) using one of two sets of independent variables (personality traits or values)

and demographic controls. For similar approaches, see Gerber et al. (2010), Gerber et al. (2012),

Mondak and Halperin (2008), and Feldman and Johnston (2014). For ease of comparison, we

plot the coefficients from the models and exclude demographic variables (see Tables A4-7 in the

Appendix for full model results).

The left column of Figure 4 displays the coefficients for personality traits predicting self-

reported ideology. Starting at the top, Openness is a significant predictor of liberal ideology

across all samples (ps < .001). The coefficients on Emotional Stability are positive in all samples

and similar in magnitude, but statistically significant only in the MTurk sample ($p < .05$).

Conscientiousness is a strong predictor of conservatism across all samples (ps < .001).

Agreeableness is a modest predictor of liberalism in only the Web and MTurk samples (ps <

.001). Finally, Extraversion does not significantly predict ideology in any of the samples.

Overall, the results are nearly identical across samples.

The middle left panel of Figure 4 displays the results of a similar analysis predicting

economic ideology. Openness significantly predicts liberalism in all samples (ps < .01).

Emotional Stability is positive across all samples but only statistically significant in the Web

sample ($p < .001$). Conscientiousness predicts conservatism in all samples (ps < .05).

Agreeableness predicts liberalism in all samples (ps < .001), though this effect falls just short of

statistical significance in the Web sample. Extraversion is positive in all samples, but only statistically significant in the Web sample. Overall, the results are again highly similar across samples, though the Web sample showed two minor deviations from the other samples.

The bottom left panel of Figure 4 shows the results for social ideology. Openness significantly predicts liberalism for all samples ($ps < .001$). Emotional Stability has a null effect in all samples. Conscientiousness is positive for all samples, but is statistically significant only in the ANES samples ($ps < .01$). Agreeableness significantly predicts liberalism in the Web and MTurk samples ($ps < .05$), but is null in the FTF sample. Finally, Extraversion is null for all samples. Overall, there was some variation in the results across samples, though the ANES samples seemed to disagree with each other about as often as with the MTurk sample.

We now conduct the same three analyses using values as our independent variables. The top right panel of Figure 4 displays the results for self-reported ideology. Authoritarianism does not significantly predict ideology in any of the samples. Racial resentment and moral traditionalism both significantly predict conservatism across all samples ($ps < .001$). Lastly, egalitarianism is a significant predictor of liberalism across all samples ($ps < .001$). Overall, the results are substantively identical across samples.

The middle right panel of Figure 4 displays the results predicting economic ideology. Authoritarianism predicts more liberal economic ideology across all samples ($ps < .05$). Racial resentment and moral traditionalism both predict conservatism across all samples ($ps < .001$). Lastly, egalitarianism is a strong predictor of liberalism for all samples ($ps < .001$). Again, we find no substantive differences between samples.

Finally, the bottom right panel of Figure 4 shows the results predicting social ideology. Authoritarianism significantly predicts conservatism in both ANES samples ($ps < .001$), but is

Figure 4: The Effects of Personality and Values on Political Ideology
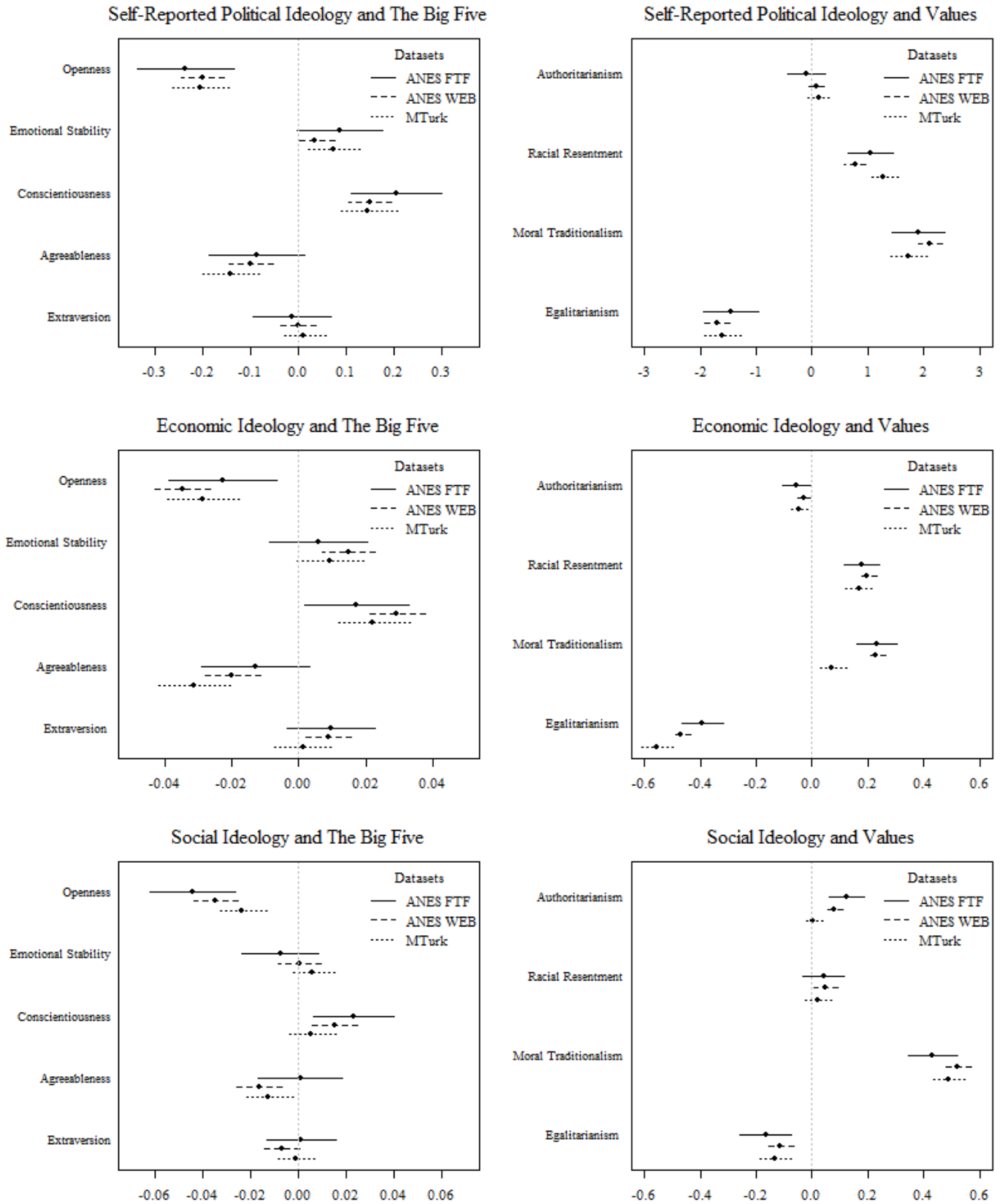
null in the MTurk sample. Racial resentment is positive in all samples, but only statistically

significant in the Web sample ($p < .05$). Moral traditionalism strongly predicts social

conservatism for all samples and equalitarianism predicts liberalism across all samples ($ps <$

.001). Overall, the results are highly similar with two exceptions in which the Web sample and

MTurk sample disagreed with the other samples.

Do Coefficient Estimates Significantly Differ Across Samples?

While the results above are highly similar across samples, there were some apparent

differences in effect sizes. To test whether these effects are significantly different from each

other, we estimated a series of models pooling two of the three samples together at a time (e.g.,

Web vs. MTurk). Each model includes a dummy variable indicating the sample (e.g., Web = 0,

MTurk = 1) and interactions between the dummy variable and each of our independent variables

(excluding demographic variables). The interaction terms in each model indicate whether the

coefficients are significantly different in magnitude across samples. We repeated this approach

for each of our four dependent variables (self-reported ideology, partisanship, economic

ideology, and social ideology). Full model results are shown in Tables A8-14. For each of the

three sample comparisons, this amounts to 36 tests (for a total of 108), raising concerns about

multiple comparisons. We address this concern by controlling the false discovery rate

(Benjamini & Yekutieli 2001; Benjamini & Hochberg 1995), which we set to .05 (e.g.,

Battaglini, Morton, and Palfrey 2007).

Beginning with the comparison between the ANES FTF and Web samples as a

benchmark, we find one significant difference in effect size out of 36 tests. However, after

controlling for the discovery rate (FDR), we find no significant differences. Comparing MTurk

to the FTF sample, we find five significant differences, though only one remains after controlling

the FDR (authoritarianism predicting social ideology). Comparing MTurk to the Web, we find six significant differences, but after controlling the FDR, only three out of 36 remain. These three differences are in moral traditionalism predicting economic ideology, racial resentment predicting partisanship, and authoritarianism predicting social ideology. The only apparent pattern in these few differences is that all three independent variables are values endorsed more by conservatives. Overall, however, our MTurk sample generated results that are substantively nearly identical to the ANES and effect sizes that are indistinguishable in the vast majority of cases.

Conclusion

While much research has documented the characteristics of MTurk workers, there has been little research on whether the selection process of joining MTurk results in conservatives who are psychologically different from the conservative population. Our results here suggest that MTurk conservatives share the same personality traits and values as conservatives drawn from high-quality national samples. Indeed, our conservatives were largely indistinguishable across samples. The differences that we did find were largely among the liberal subjects. MTurk liberals appear to hold more characteristically liberal values and political attitudes. As for personality traits, we found few differences across samples with the exception of Extraversion. Here, the differences were primarily in the average level of the variable (rather than its relationship with ideology). The FTF sample was the most extraverted, while our MTurk sample was less extraverted than both samples.

Following the approach of recent research in political psychology, we also examined the psychological predictors of political ideology. The results were substantively nearly identical

across samples, indicating that a researcher would draw nearly identical conclusions if she chose to sample from MTurk, rather than rely on the ANES. MTurk effect sizes significantly differed from the ANES in only four tests out of seventy-two. Most of these cases involved the predictive power of conservative values, such as moral traditionalism. It is unclear why these variables might have weaker predictive power among the MTurk sample, but it might indicate shifting political divisions among a younger and less religious population.

Finally, we must note that our research does not provide blanket support for the use of MTurk samples for two reasons. First, our MTurk sample was not a random sample of the MTurk workforce. As a result, we cannot be confident that our results apply to any sample drawn from MTurk. However, samples drawn from MTurk consistently diverge from the national population on several key characteristics, such as age, ideology, and religiosity. The persistence of these findings suggests that the demographics of MTurk are fairly stable (Berinsky et al. 2012; Krupnikov & Levine 2014; Buhrmester et al. 2011; Huff & Tingley 2014; Shapiro et al. 2013). Second, concerns remain about the effects of subject "savviness" or "non-naïveté" (Krupnikov & Levine 2014; Chandler et al. 2014), cross-talk (Chandler et al. 2014) and other issues (for discussion, see Goodman, Cryder, and Cheema 2013; Shapiro, Chandler, and Mueller 2013). However, our research does suggest that MTurk workers largely share the psychological dispositions of their ideological counterparts in the mass public.

References

Ahler, D.J., 2014. Self-Fulfilling Misperceptions of Public Polarization. *The Journal of Politics*, 76(03), pp.607–620. Available at: http://www.journals.cambridge.org/abstract_S0022381614000085 [Accessed November 28, 2014].

Arceneaux, K., 2012. Cognitive Biases and the Strength of Political Arguments. *American Journal of Political Science*, 56(2), pp.271–285. Available at: http://doi.wiley.com/10.1111/j.1540-5907.2011.00573.x [Accessed December 27, 2014].

Battaglini, M., Morton, R.B. & Palfrey, T., 2007. Efficiency, Equity, and Timing of Voting Mechanisms. *American Political Science Review*, 101(03), p.409. Available at: http://journals.cambridge.org/abstract_S0003055407070281 [Accessed July 30, 2015].

Benjamini, Y. & Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, 57(1), pp.289–300.

Benjamini, Y. & Yekutieli, D., 2001. The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics*, 29(4), pp.1165–1188.

Berinsky, A.J., Huber, G.A. & Lenz, G.S., 2012. Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3), pp.351–368. Available at: http://pan.oxfordjournals.org/cgi/doi/10.1093/pan/mpr057 [Accessed February 21, 2014].

Brüggen, E. & Dholakia, U.M., 2010. Determinants of Participation and Response Effort in Web Panel Surveys. *Journal of Interactive Marketing*, 24(3), pp.239–250. Available at: http://www.sciencedirect.com/science/article/pii/S1094996810000277 [Accessed May 30, 2015].

Buhrmester, M., Kwang, T. & Gosling, S.D., 2011. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, 6(1), pp.3–5. Available at: http://pps.sagepub.com/content/6/1/3.short [Accessed July 11, 2014].

Bullock, J.G., 2011. Elite Influence on Public Opinion in an Informed Electorate. *American Political Science Review*, 105(03), pp.496–515. Available at: http://journals.cambridge.org/abstract_S0003055411000165 [Accessed May 6, 2015].

Chandler, J., Mueller, P. & Paolacci, G., 2014. Nonnaïveté among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers. *Behavior research methods*, 46(1), pp.112–30. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23835650 [Accessed July 10, 2014].

Clifford, S., 2014. Linking Issue Stances and Trait Inferences: A Theory of Moral Exemplification. *The Journal of Politics*, 76(03), pp.698–710. Available at: http://journals.cambridge.org/abstract_S0022381614000176 [Accessed September 2, 2014].

Costa, P.T. & McCrae, R.R., 1992. Four ways five factors are basic. *Personality and Individual Differences*, 13(6), pp.653–665. Available at: http://www.sciencedirect.com/science/article/pii/019188699290236I [Accessed January 17, 2015].

Dollinger, S.J. & Leong, F.T.L., 1993. Volunteer Bias and the Five-Factor Model. *The Journal of Psychology*, 127(1), pp.29–36. Available at: http://www.tandfonline.com/doi/abs/10.1080/00223980.1993.9915540#.Vb_BLvnGp64 [Accessed August 3, 2015].

Druckman, J.N., Peterson, E. & Slothuus, R., 2013. How Elite Partisan Polarization Affects Public Opinion Formation. *American Political Science Review*, 107(01), pp.57–79. Available at: http://www.journals.cambridge.org/abstract_S0003055412000500 [Accessed March 12, 2014].

Feldman, S. & Johnston, C., 2014. Understanding the Determinants of Political Ideology: Implications of Structural Complexity. *Political Psychology*, 35(3), pp.337–358. Available at: http://doi.wiley.com/10.1111/pops.12055 [Accessed July 30, 2015].

Feldman, S. & Stenner, K., 1997. Perceived Threat and Authoritarianism. *Political Psychology*, 18(4), pp.741–770. Available at: http://doi.wiley.com/10.1111/0162-895X.00077 [Accessed August 2, 2015].

Gerber, A.S. et al., 2013. Big Five Personality Traits and Responses to Persuasive Appeals: Results from Voter Turnout Experiments. *Political Behavior*, 35(4), pp.687–728. Available at: http://link.springer.com/10.1007/s11109-012-9216-y [Accessed May 31, 2015].

Gerber, A.S. et al., 2010. Personality and Political Attitudes: Relationships across Issue Domains and Political Contexts. *American Political Science Review*, 104(1), pp.111–133.

Goodman, J.K., Cryder, C.E. & Cheema, A., 2013. Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision Making*, 26(3), pp.213–224. Available at: http://doi.wiley.com/10.1002/bdm.1753 [Accessed May 6, 2015].

Gosling, S.D., Rentfrow, P.J. & Swann, W.B., 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), pp.504–528. Available at: http://www.sciencedirect.com/science/article/pii/S0092656603000461 [Accessed July 9, 2014].

Grimmer, J., Messing, S. & Westwood, S.J., 2012. How Words and Money Cultivate a Personal Vote: The Effect of Legislator Credit Claiming on Constituent Credit Allocation. *American*

Political Science Review, 106(04), pp.703–719. Available at: http://journals.cambridge.org/abstract_S0003055412000457 [Accessed July 14, 2015].

Henry, P.J. & Sears, D.O., 2002. The Symbolic Racism 2000 Scale. *Political Psychology*, 23(2), pp.253–283. Available at: http://doi.wiley.com/10.1111/0162-895X.00281 [Accessed August 3, 2015].

Huber, G.A. & Paris, C., 2013. Assessing the Programmatic Equivalence Assumption in Question Wording Experiments: Understanding Why Americans Like Assistance to the Poor More Than Welfare. *Public Opinion Quarterly*, 77(1), pp.385–397. Available at: http://poq.oxfordjournals.org/content/early/2013/01/28/poq.nfs054.short [Accessed August 3, 2015].

Huff, C. & Tingley, D., 2014. *"Who are These People": Evaluating the Demographic Characteristics and Political Preferences of MTurk Survey Respondents*,

Johnston, C.D., Newman, B.J. & Velez, Y., 2015. Ethnic Change, Personality, and Polarization Over Immigration in the American Public. *Public Opinion Quarterly*, p.nfv022–. Available at: http://poq.oxfordjournals.org/content/early/2015/05/22/poq.nfv022.short [Accessed June 18, 2015].

Kahan, D.M., 2013. What's a "valid" sample? Problems with Mechanical Turk study samples, part 1. *Cultural Cognition Project Blog*. Available at: http://www.culturalcognition.net/blog/2013/7/8/whats-a-valid-sample-problems-with-mechanical-turk-study-sam.html.

Krupnikov, Y. & Levine, A.S., 2014. Cross-Sample Comparisons and External Validity. *Journal of Experimental Political Science*, 1(01), pp.59–80. Available at: http://journals.cambridge.org/abstract_S2052263014000074 [Accessed May 8, 2015].

McDermott, R., 2002. Experimental Methodology in Political Science. *Political Analysis*, 10(4), pp.325–342. Available at: http://pan.oxfordjournals.org/content/10/4/325\nhttp://pan.oxfordjournals.org/content/10/4/325.abstract\nhttp://pan.oxfordjournals.org/content/10/4/325.full.pdf.

Mondak, J.J., 2010. *Personality and the Foundations of Political Behavior*, Cambridge University Press. Available at: https://books.google.com/books?hl=en&lr=&id=J6ggAwAAQBAJ&pgis=1 [Accessed June 23, 2015].

Mondak, J.J. & Halperin, K.D., 2008. A Framework for the Study of Personality and Political Behaviour. *British Journal of Political Science*, 38(02), pp.335–362. Available at: http://journals.cambridge.org/abstract_S0007123408000173 [Accessed November 3, 2014].

Nyhan, B. & Reifler, J., 2010. When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*, 32(2), pp.303–330. Available at: http://link.springer.com/10.1007/s11109-010-9112-2 [Accessed October 2, 2014].

Rogelberg, S.G. et al., 2003. Profiling Active and Passive Nonrespondents to an Organizational Survey. *Journal of Applied Psychology*, 88(6), pp.1104–1114.

Searles, K. & Ryan, J.B., 2015. Researchers are rushing to Amazon's Mechanical Turk. Should they? *Washington Post*. Available at: http://www.washingtonpost.com/blogs/monkey-cage/wp/2015/05/04/researchers-are-rushing-to-amazons-mechanical-turk-should-they/.

Shapiro, D.N., Chandler, J. & Mueller, P.A., 2013. Using Mechanical Turk to Study Clinical Populations. *Clinical Psychological Science*, 1(2), pp.213–220. Available at: http://cpx.sagepub.com/content/early/2013/01/31/2167702612469015.full [Accessed May 13, 2015].

Weinberg, J.D., Freese, J. & McElhattan, D., 2014. Comparing Data Characteristics and Results of an Online Factorial Survey between a Population-Based and Crowdsource-Recruited Sample. *Sociological Science*, 1, pp.292–310.