Personality Stability and Politics: TIPI Variability*

Joshua Boston†
Washington University in St Louis

Jonathan Homola‡
Washington University in St Louis

Betsy Sinclair§
Washington University in St Louis

Michelle Torres¶
Washington University in St Louis

Patrick Tucker‖
Washington University in St Louis

October 4, 2014

*We thank the Weidenbaum Center for generously providing data from The American Voter Panel Survey for this project. We also thank Steve Smith, Michelle Margolis, Jon Rogowski, Jim Gibson and Chris Classen for their thoughtful comments that have significantly improved this manuscript.

†Department of Political Science, 1 Brookings Drive, St Louis, MO 63130

‡Department of Political Science, 1 Brookings Drive, St Louis, MO 63130

§Corresponding author: Department of Political Science, 1 Brookings Drive, St Louis, MO 63130; bsinclai@wustl.edu.

¶Department of Political Science, 1 Brookings Drive, St Louis, MO 63130

‖Department of Political Science, 1 Brookings Drive, St Louis, MO 63130

1 Abstract

Researchers frequently claim that personality traits, as measured using the Big Five personality through the TIPI (Ten Item Personality Inventory) battery, affect Americans' political attitudes and behaviors. Such studies often depend on two key assumptions: personality measurements display stability over time and variability in such measurements predates political behaviors of interest. In this paper we employ new panel survey data to test these assumptions. First, we find high levels of variability in responses to TIPI. Second, we associate this variability with not only socioeconomic and demographic characteristics, but also, and more concerning, political attitudes. The variability and associations of the instrument suggest that the relationship between personality and politics may be weaker than indicated by previous scholars and moreover should not be employed as a variable that predates political behavior. Finally, we address the consequences of these tests by applying our findings to previous research that hinges on the relationship between personality and political behavior. Ignoring the dynamic nature of (measured) personality alters the interpretation of its relationship to political attitudes.

2 Introduction

Whether political attitudes and behaviors are stable or fluid over time greatly impacts our ability to understand and predict the political world. Scholars have long been concerned with the stability and durability of many political phenomena, including partisan identification (e.g. Clarke & McCutcheon 2009; Green, Palmquist, & Schickler 2004), policy preferences (Highton 2012), and authoritarianism (Feldman 1988; Goren 2005). As such, many political scientists have turned to studying personality – as a stable, latent trait of the individual – and how it influences political attitudes and behavior (Carney, Jost, Gosling, & Potter 2008; Gerber, Huber, Doherty, & Dowling 2011, 2012, 2013; Gerber, Huber, Doherty, Dowling, & Ha 2010; Mondak & Halperin 2008).

In doing so, scholars regularly use different quantified measures of personality to explain citizens' political attitudes (Carney et al. 2008; Gerber et al. 2010; Mondak & Halperin 2008), strength of partisan attachments (Gerber et al. 2012), U.S. state legislators' behavior (Dietrich, Lasley, Mondak, Remmel, & Turner 2012), presidential approval (Mondak & Halperin 2008), and other political phenomena in the U.S. and comparative contexts (Caprara, Barbaranelli, & Zimbardo 2002; Gerber, Huber, Doherty, Dowling, Raso, & Ha 2011; Ha, Kim, & Jo 2013). These studies most frequently rely upon a survey instrument referred to as the Ten Item Personality Inventory (TIPI), which collapses personality into five personality dimensions (the "Big Five": openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism/emotional stability) (Gosling et al. 2003).

Given the frequency of causal claims regarding the relationship between personality and political attitudes, we are concerned with the potential dynamic nature of personality traits. Individual-level stability across the Big Five personality dimensions is of critical importance for research on politics and personality. When arguing that personality consistently influences political behaviors, social scientists implicitly assume that personality is stable across time and does not vary based upon exposure to political phenomena. However, this assumption has not been rigorously tested.

Therefore, this study investigates the stability of the TIPI measurement of personality. We are not only interested in whether the measurement of individual-level personality varies over time, but also whether variability is associated with particular types of individuals or with political behavior. By analyzing whether political outcomes and individual characteristics affect the TIPI

measurement of these five traits, we are effectively evaluating whether the Big Five traits as captured by TIPI are appropriate explanatory variables in understanding how an individual forms and shapes her political responses. If the measurement of an individual's personality varies over time, then scholars should be more uncertain about the explanatory power of these personality traits as set forth by TIPI. Additionally, if the measurement of an individual's personality varies on the basis of political attitudes or opinions (or those covariates which are associated with political attitude and opinion), then it may simply be that the dimensions covered by TIPI are not appropriate tools to explain response to a political world.

We analyze TIPI data from six waves of the nationally-representative American Panel Survey (TAPS) sample taken over the course of almost two years – a uniquely extensive sample for the TIPI personality battery that provides unprecedented precision in evaluating the stability of this personality battery. Our results help adjudicate previous tension in the literature regarding stability. Research in psychology establishes that socio-environmental and contextual factors influence personality over time (Caspi, Roberts, & Shiner 2005; Srivastava, John, Gosling, & Potter 2003). This situational perspective, where an individual reacts to external circumstances rather than merely relying on a fixed internal disposition (Digman 1990), might pose a problem for reliably measuring personality. On the other hand, Gerber, Huber, Doherty, & Dowling (2013), relying on a two-wave national panel survey, find personality to be stable and unaffected by political events. Beyond this, little political science research exists on individual-level variability of the Big Five personality traits.

This article will proceed in four parts: first, we develop a theoretical explanation for how individual-level political and social factors influence changes in personality, as measured through the TIPI, which is a mechanism for capturing the Big Five personality dimensions. Second, we discuss individual-level variability in the TIPI personality measurements from panel survey data provided by a nationally-representative survey, The American Panel Survey (TAPS). Third, we examine how personality variations across panel waves relate to static socio-demographic and political characteristics, as well as dynamic measures of political attitudes and preferences. Finally, we consider the implications of our findings for the study of political behavior and the impact our research has for future studies.

3 Personality and Politics

Social scientific narratives of political behavior consist of various pre-existing characteristics that affect an individual's decision-making processes. This idea of temporal dependence is present in many studies of political phenomena. For instance, previous analyses have found that socialization predates party identification (Campbell, Converse, Miller, & Stokes 1960), economic conditions predate evaluations of the president (MacKuen, Erikson, & Stimson 1992), and information about candidates or issues predates vote choice (Lupia 1994). Each of these antecedent or "pre-treatment" variables helps, at least in part, to explain the resulting political behavior.

In turn, any study seeking to determine what influences individual-level political behavior would need to know certain things about each respondent that precede the respective behavior. Due to the considerable number of endogeneity issues in social science research, where variables cause each other simultaneously, it can be quite difficult to identify true causes of behavior. For example, does an individual's ideology cause them to choose certain media sources, or do the media sources influence the individual's ideology (Stroud 2008)? Therefore, scholars have, over time, winnowed the field of variables to those that are the most fundamental and exogenous to political behavior.

Chief among these variables is personality, which scholars have argued is exogenous, in that it originates prior to individual behavior. As a fundamental and intrinsic characteristic, personality has been labeled "the psychology of individual differences" (Wiggins 1996). Gerber, Huber, Doherty, & Dowling (2011) note that research outside political science has drawn associations between personality and alcohol use (Mezquita, Stewart, & Ruipérez 2010), physical fitness (Rhodes & Smith 2006), cholesterol (Sutin et al. 2010), and overall health (Goodwin & Friedman 2006). As political scientists take a heightened interest in the prospect of genetically driven political attitudes (Alford, Funk, & Hibbing 2005), they begin to exploit personality as a natural (supposedly heritable), exogenously determined differentiation between voters. Gerber, Huber, Doherty, Dowling and other coauthors (2010, 2011, 2012, 2013) are the most prolific in their attempts to connect individual personality with political attitudes and behavior. Their results suggest that the Big Five personality traits predict partisanship, racial attitudes, and right-wing authoritarianism. Furthermore, Gerber, Huber, Doherty and Dowling claim that the Big Five traits are particularly advantageous for political science research, because

> [They] are not obviously associated with political attitudes and behaviors. Instead, they are broad dispositions that are theorized to shape responses to the full range of stimuli people encounter in the world. Thus, just as socioeconomic status is associated with a broad range of forms of political and social engagement, political research on Big Five traits may provide a way to situate political judgments and behaviors within the context of a broader theoretical account of how an individual engages with their environments. (Gerber, Huber, Doherty, & Dowling 2011)

Just as nature drives political attitudes through genetic predispositions, so does nature drive personality differences, which "may be linked to virtually all aspects of political behaviour" (Mondak & Halperin 2008). Such a relationship between politics and personality can occur either directly – where attitudes and behavior are an explicit function of personality – or indirectly – where situations or contexts have a conditioning effect on the relationship between personality and political behaviors (Gerber et al. 2010; Mondak & Halperin 2008).

Survey mechanisms like TIPI are used to identify individual characteristics through the "language of personality" or lexical analysis (Gerber, Huber, Doherty, & Dowling 2011). "[M]ost of the socially-relevant and salient personality characteristics have become encoded in the natural language" (Gerber, Huber, Doherty, Dowling, & Ha 2010; John & Srivastava 1999). That is, in answering survey questions aimed at measuring personality, a respondent judges herself using language that relates to one of the Big Five dimensions. Scholars claim that an individual's self-evaluative responses to personality survey questions are stable, and the resultant Big Five dimensions "can be successfully conceptualized as need-like constructs that motivate people to respond in a certain way to environmental circumstances" (Denissen & Penke 2008, 1286). Alternatively, more recent efforts at measuring personality have used an open-vocabulary method, whereby text analysis software extracts and evaluates linguistic features from existing text – like an individual's Facebook activity – rather than having respondents identify themselves with particular language (Schwartz et al. 2013). No matter how personality is measured, when an individual observes a change in the political environment, she is likely to react in a specific way that is largely dependent on her existing personality.

Still, current evidence regarding the stability of personality traits is largely contradictory.

On one hand, core personality traits were long thought to be heritable, and therefore rooted in biological causes; once fully matured, an individual would be relatively invariant in her Big Five traits (McCrae & Costa Jr. 1996). Other research suggests that fully matured individuals can vary in their Big Five personality dimensions, though the changes in adults are smaller in magnitude compared to younger individuals. Scholars attribute the variation to continued biological maturation and not social or environmental factors (McCrae et al. 2000).

If, on the other hand, we were to observe a continuous loop of causality between personality and life events, then we should be less confident in the explanatory power of personality. In other words, if personality is endogenous, it would respond to environmental changes while personality could also lead to an individual to change her environment. Therefore, events may influence personality directly, or events may cause changes in an individual's social environment, which, in turn, influences personality (Srivastava et al. 2003). People often endure life changes ranging from car accidents and job promotions to marriages/divorces and relocations/retirements. An individual's personality might compel that person to seek out experiences that a different person would avoid. As such, personality-driven choices are a part of a causal look, serving to reinforce personality traits and dispositions (Srivastava et al. 2003) in the same way as political scientists have come to expect an individual to access information from elites and the media that is consistent with her existing ideology (Zaller 1992).

Furthermore, previous research suggests that the electorate's behavior concerning support for incumbents is influenced by, first, disasters/catastrophes and, second, public officials' subsequent reactions (e.g. Gasper & Reeves 2011). Pairing this finding with our understanding of personality suggests that a catastrophic event for example, might influence an individual's personality, which in turn would condition her reactions to an event and the decisions made by office-holders.

Given these findings, it is likely not safe to assume that personality is stable across time. In contrast, it would not be surprising to find that an individual's self-evaluative responses to personality survey questions vary based on stimuli, which likely include social factors as contributing to an individual's surroundings. In other words, an individual's variation in personality would be cause for concern, especially if the variation occurs in response to political phenomena. If personality is regularly in a state of updating and adaptation, then it is not hard to see that the expected relationship could be reversed; political events (e.g. elections), political behavior (e.g. voting or

approval of public officials), and other factors (e.g. catastrophes or [inter]national crises) could all predict change in an individual's personalities. Once again, consider the notion that personality, as measured by TIPI, is influenced by the changes in the political context or the individual's own political beliefs. As such, the traditional conception of personality as an explanatory variable with respect to these political variables may be misunderstood. Therefore, it is necessary to further investigate the status of personality as a variable that predates political behavior.

4 Measuring personality: the Big-Five factor structure and the Ten Item Personality Inventory

Academic measurement of personality is dominated by the Big-Five factor structure. The Big-Five factor model is a strategy that maps personality along five dimensions (or factors): extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience. Each of these dimensions groups several elements that define an individual's personality and that distinguishes it from others. For example, the emotional stability dimension measures if an individual tends to be tense, anxious, rigid or concerned, in contrast to other adjectives such as relaxed, calm, and tolerant to stress (McCrae & John 1992).

The five dimensions are built based on a number of "traits" associated with each factor. In other words, the score obtained by a certain individual in each dimension summarizes the presence (or absence) of bipolar elements implied by a broader category. There are other measurements that accomplish similar results, but with much less parsimonious structures that in practice lead to low levels of efficiency and high costs.[1]

The number of personality elements per factor determines the level of complexity and richness of the measurement. The Ten Item Personality Inventory (TIPI) is a popular method to construct the Big-Five structure since it uses a sufficient number of elements without sacrific-

[1]For example, one of the most important measures, the Cattell system (Cattell & Mead 2008), structures personality around 16 dimensions and 8 sub-dimensions. Indexes like NEP-PIR, NEO-FFI or BFI (Benet-Martínez & John 1998; John & Srivastava 1999; McCrae & Costa Jr. 1996) aim to measure scores across only five factors but based on a pool of items that ranges from 44 to 240. Needless to say, this represents some practical limitations regarding the data collection process.

ing a crucial part of its reliability and validity.[2] It involves only ten traits, with two items per dimension/factor that account for a score of a given individual within each of the five facets of personality. One of the most comprehensive previous examinations of the inventory demonstrates its criterion validity and validity to other scales measuring the five factor model with a significant higher number of traits (Gosling et al. 2003).[3]

The TIPI is based on two items per factor, with each item in turn consisting of two adjectives (e.g., "extraverted, enthusiastic" is an item of the extraversion factor), and another pair of adjectives with contrasting valence or direction (e.g., "reserved, quiet" as the contrasting element of the extraversion factor).[4] Once provided with the pair of descriptions, respondents identify how well both describe their personality on a scale of 1 ("Disagree Strongly") to 7 ("Agree Strongly").

In a given survey, each panelist receives a "score" for each dimension that captures the respondent's personality in that particular factor. High scores in a given subscale mean that the respondent's personality is highly defined by the attributes of that block. Each factor's score is calculated by summing the positive trait with the "inverse" or reversed score of its contrasting negative trait and then dividing the sum by two. The score for the negative trait can be obtained by subtracting the number scored in a particular reversed question from 8. For example, given the seven-point scale, a panelist identifying as 6 for "extraverted, enthusiastic" and a 3 for "reserved, quiet" would receive a score of 5.5 for extraversion, calculated as:

$$\text{Extraversion} = \frac{(\text{Enthusiasm} = 6) + (\text{Reserved} = 3)}{2} = \frac{6 + (8 - 3)}{2} = 5.5 \tag{1}$$

Considering that the values of this index range from 1 to 7, we could identify this individual as having a relatively extraverted personality.

Since the method used to build this scale is based on the assumption of orthogonality be-

[2]TIPI has been found – in large part – to be a valid abbreviation to the longer instruments. Gerber et al. (2010), who state that TIPI "was designed to [...] achieve high test-retest reliability", showed that TIPI correlated with the 44-item BFI between 0.65 and 0.87, and with the 240-item NEO PI-R between 0.56 to 0.68. Other research examining the briefness of TIPI by Hofmans, Kuppens, & Allik (2008) suggests that TIPI is a valid alternative to other big five instruments, even when translated into another language.

[3]Based on two waves that were approximately six weeks apart, the authors also find evidence for short-term retest reliability. Employing data from six waves that span almost two years, our analysis below provides a far more rigorous test of that claim.

[4]Figure A1 in the Appendix that is available from the authors upon request presents the actual battery of questions that TAPS uses to measure personality.

tween factors, which assumes no correlation among them, each dimension is generally analyzed separately. It also relies on the assumption that the dimensions are not mutually exclusive and that the addition or aggregation of factor scores does not lead to an easily interpretable overall score. In other words, a high aggregated score does not imply a specific type of personality that would contrast one depicted by a lower aggregated score.

5 The American Panel Survey Personality Data

Data for the following analysis are provided by the The American Panel Survey (TAPS). TAPS is a nationally-representative panel survey that conducts an online poll of up to 2,000 adult respondents monthly, starting in December of 2011 by Knowledge Networks (now GfK Knowledge Networks) for the Weidenbaum Center at Washington University.[5] At the beginning of each month, members of the panel receive notification to complete the new survey. Each wave remains open for approximately one month and takes between 15 and 25 minutes to complete. TAPS encompasses a wide variety of economic, sociological, and political questions asked on a large scale. Such breadth of data provides researchers with a unique opportunity to investigate trends and changes at the individual level. For example, if an individual remains active in the panel for two years, TAPS collects over 1,000 variables at 24 different points in time for one individual. The survey instruments repeat many questions over multiple waves. Such design invites investigation of individual-level change over both the short- and long-term.

One such set of variables pertains to TIPI. Panelists answered these questions over six waves (February, June, October 2012; May, September, November 2013). As Table 1 displays, 2,014 of those participating in the panel completed at least two waves of the TIPI battery.[6] Like

[5]More technical information about the survey is available at http://taps.wustl.edu. The sampling frame used to select the 2,000 respondents is the U.S. Postal Service's computerized delivery sequence file (CDSF), which covers around 97% of the physical addresses in all fifty states including P.O. boxes and rural route addresses. This frame is appended with information regarding householders' names, demographic characteristics of the inhabitants, and landline telephone numbers obtained from other sources such as the U.S. Census files and commercial data bases (e.g. White pages). The respondents are recruited based on a random stratified sample, where Hispanics and young adults between 18 and 24 years of age are slightly oversampled in order to account for their tendency to under-respond to surveys. Through the support of the Weidenbaum Center, those individuals without internet access are provided with a computer and internet access.

[6]Given the total N of 2,789 for the TAPS dataset, this means that we have less than 30% missingness in our outcome variable.

all panel surveys, attrition occurs, but over one-third (720) of those in our dataset completed each wave in which the personality index was measured. More than half of those in the set responded to at least 5 waves (1,233). We focus on those respondents who completed at least two waves of the TIPI so as to be able to evaluate the stability of their responses. Table 1 offers an overview of the descriptive statistics for the variables included in our analyses.

[TABLE 1 ABOUT HERE]

Previous research finds associations of various factors, both political and non-political, with the development of the individual's personality. Since our interest lies in which traits are associated with stability of common personality measurements, it serves our study well to explore the connections between factors cited in other studies with TIPI variability. Given that political labels tend to correlate with scores on the Big Five, we employ measurements for both *symbolic ideology* and *partisan identification*. We measure the former by allowing panelists to place themselves on a seven point scale ranging from "very liberal" (-3) to "very conservative" (+3). Partisan identification is measured categorically, as Democrat, Republican, and Independent/Other party (this category serves as the baseline in all models). Table 1 demonstrates that our sample's political identifications are very similar to national averages. Those identifying as Democrats outnumber Republicans, while the average panelist identifies to the conservative side of the symbolic ideological spectrum.

In addition to political affiliations and orientations, variation in the five items is often associated with other political behaviors. Consequently, we include a variable that measures panelists' *political interest* on a 4-point scale("very interested", "somewhat interested", "not very interested", and "not at all interested"). Since higher values indicate lower levels of interest, Table 1 shows that our panelists identify as being very interested in public affairs. Furthermore, evidence exists that the big five traits affect *political knowledge* (Gerber, Huber, Doherty, Dowling, & Ha 2010). To search for such a connection we provide our panelists with a 10-question battery on various topics in American politics. The overall measurement is thus the summed number of questions answered correctly. In this sample, the mean of correct responses is between 6 and 7. We include a measure for *religiosity* (church attendance as measured on a five point scale) since research demonstrates that the frequency of religious observance varies cross-nationally with multiple dimensions of the

Big Five (Saroglou 2010). A value of "2.35", as Table 1's corresponding mean displays, represents a level of attendance between biweekly and bimonthly. We also include other controls upon the recommendations of Gerber and his coauthors (2010): *education*, a dichotomous measure indicating that the panelist is *employed*, *income*, dichotomous indicators for *Black* and *Hispanic*, presence of *children*, *sex*, level of *happiness*, and *age*.

Our dynamic variables of interest fall into categories of either social or political. For the former, subjects were asked essential biographical information upon entering the panel. At various points in the panel's duration, these questions were re-asked. The traits that are most important include whether the panelist indicated she had experienced a *job loss* (asked in November 2011 and again in November 2012) or changed her marital status (*married* or *divorced*; asked in November 2011 and again in January 2014). Unsurprisingly, the number of panelists experiencing these events is not overwhelming; however, a conspicuous proportion report employment change. Table 1 shows that over ten percent of those with a job in November 2011 are no longer working a year later. Marital status is much more static. Few report entering marriage or experiencing divorce.

In addition to charting personal changes, TAPS regularly surveys its sample's views on political affairs. Each month at least one-half of the panel provide their level of approval on key political actors, such as Congress (*Congressional Approval*) and the President (*Presidential Approval*), on a five-point scale from strongly approve to strongly disapprove. Variables indicating an individual's level of change in perception are measured by taking the variance of the responses over the seven waves that occurred most closely to the instances of gathering TIPI data.[7] Table 1 shows that while changes in approval of elites do occur within the panel, the average variance of such opinions is quite small. The mean value of such a measurement is near zero. Additionally, on average, presidential approval is somewhat more stable than approval of Congress as a whole. Finally, our analysis operationalizes changes in political knowledge by taking the difference between scores of the 10 item knowledge battery asked in January 2012 and May 2013. The maximum amount of change in both directions (i.e. from correct to incorrect and vice versa) is seven questions. While improvement does occur over the 16 month interval, the average panelist, who answered both batteries displays nearly zero change. Finally, panelists responding to more

[7]Since each panelist answers this question (on average) every other month, individual wave perceptions are pooled over two months to maximize the number of observations.

waves will have greater opportunity to provide more variant TIPI factor scores. To control for such a confounding effect, we include a variable in our model that indicates how many *waves* of TIPI each panelist completed.

6 Measuring Variability of TIPI scores

For the present study we calculate TIPI scores for each of the five subscales that the Big-Five structure considers. Typically, scores on the Big-Five dimensions remain disaggregated, particularly when researchers are interested in the different effects each dimension may have on a political variable (e.g. Rammstedt & John 2007). However, we also sum an individual's five subscale scores together to create an overall additive scale.[8] This facilitates the analysis below, while also ensuring the preservation of the maximum available information in TAPS. It is important to highlight that we are interested in the variation and volatility of personality and not in the nature or direction of these changes. The aggregation of the subscales allows us to observe potential changes through all the dimensions and reach more precise and accurate estimates of personality volatility.

The key outcome variable for this analysis measures the stability of the Big-Five Model as tracked by TIPI. In order to capture this phenomenon, we rely upon changes in an individual panelist's responses over a certain period of time. Such a decision allows us to account for the different number of waves completed by panelists and also register multiple changes across waves.

To measure the stability of TIPI across all waves in which a given panelist participates, we calculate the variance of each subscale, as well as the one of the aggregated scale using the following formula:

$$s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \bar{x_i})^2, \tag{2}$$

where x_{ij} is the score in either one of the five subscales or the aggregated scale obtained by individual i in a given wave j, $\bar{x_i}$ the mean of scores of the respective scale in all waves answered by

[8] Accordingly, the aggregated TIPI score would be calculated as: $Aggregate_{ij} = Openness_{ij} + Conscientiousness_{ij} + Extraversion_{ij} + Agreeableness_{ij} + EmotionalStability_{ij}$, where the aggregated measure for respondent i in wave j equals the sum of the respondent's five subscale scores in that wave.

panelist i, and n the number of waves completed by panelist i.

To provide a better idea of the new measure we introduce, Table 3 presents summary statistics for the aggregated TIPI scale, and its average across all six waves, as well as information on the variance measure introduced above. We have aggregated TIPI scores for an average of 1,578 respondents per wave. Given the construction of that variable, the theoretical minimum of the aggregated scale would be 5 (if a respondent were to score 1 on all five subscales). While a score of 5 is only observed in one wave, the theoretical maximum of 35 (a maximum score of 7 on all five subscales) is observed in all six waves. The mean score of the aggregated scale across all respondents in all six waves is 25.80.

[TABLE 3 ABOUT HERE]

7 Is TIPI stable?

One of the main questions that drives the present study concerns the assumed stability of personality traits as measured by the TIPI scale. Although other authors have previously looked for an answer to this question, we aim to test their findings and improve our understanding of the TIPI scale's reliability by using a larger panel that will allow a deeper analysis of individual personality change over a much longer time frame.

When testing stability we face a problem inherent to survey analysis: measurement error. This "refers to the inaccuracy with which the underlying attitude is reflected by the survey instrument" (Achen 1975, pp. 1221). Achen called the attention to the idea that respondent's instability can be attributed to certain factors such as context, the vagueness of questions, the amount and clearness of the answers available, etc. Consequently, these elements are part of the volatility of attitudes that will be observed when analyzing panel survey responses through time and should be considered when studying the change and stability of any indicator.

Based on these considerations, we conduct a series of analyses that compare the correlations between scales found by other studies such as the the one presented by Gerber, Huber, Doherty and Dowling (2013). We also present results from a model proposed by Ansolabehere, Rodden, & Snyder (2008) that accounts for measurement error in order to calculate "true correla-

tions" between the subscales.

7.1 Comparing datasets and methods

There are several ways to test the stability of indicators in general and accordingly also of personality measurement more specifically. In their article, Gerber, Huber, Doherty and Dowling (GHDH) study the stability of personality scores obtained via TIPI across two waves of the Cooperative Congressional Election Study 2010 (CCES): a pre-election wave and a post-election wave. To assess stability, they compute polychoric correlations between standardized scores based on the means and standard deviations of the pre-election wave.

For our study, we employ a slightly different approach and first calculate the traditional TIPI scores that range from 1 to 7 for each of the subscales and the additive scale without further transformation. Then, we compute a series of correlations between the subscales and aggregated scale scores across the six waves to analyze their level of association over time. To calculate the volatility of values for a given subscale, we correlate its scores in a certain wave with those observed in each of the remaining waves.[9] Based on this method we find preliminary evidence that might contradict the claimed stability of the TIPI scale. Table 3 shows that the average variability of the aggregated TIPI scale is 4.76, which differs from the zero-mean that would indicate a completely stable indicator.

Moreover, Table 2 shows summary statistics for all possible across-wave correlations of the different subscales. For example, we can compute correlations among openness scores between waves one to six. Upon examining those correlation scores, the smallest one (indicating the smallest correlation between openness scores among any two waves) is 0.554 for the correlation between openness scores in waves 3 and 4, while the highest one is 0.644 (between waves 1 and 4). On average, openness scores are correlated at 0.586 across our six waves.

[TABLE 2 ABOUT HERE]

In order to make comparisons between the findings presented by GHDH and our results, we subsequently used both methods discussed above and applied them to the two datasets under

[9]Since we do not assume a linear relationship between the subscales at different points in time, we compute Spearman's rank correlation coefficients rather than Pearson correlation coefficients. Nonetheless, we do assume that the relationship between periods is monotonic.

analysis: the CCES and TAPS. The most striking finding from our analysis using TAPS is the absence of strong correlations among the personality indicators that should exist in order to claim stability. Figure 1 displays the distribution and correlation coefficients of the different subscales through the six waves under analysis. Correlation coefficients range from 0.56 to 0.74 across all subscales. If we only focus on the aggregated scale we find slightly higher levels, but even those never exceed a correlation coefficient of 0.73.

[Figure 1 ABOUT HERE]

These coefficients are slightly higher when we apply the method GHDH suggest and test for polychoric correlations instead of Spearman's rank correlations. However, although the coefficients are higher than those calculated by Spearman correlations, the absolute differences do never exceed 0.03, which indicates that there are no significant differences between the methods employed by GHDH and us. However, when applying our method to the GHDH data, the differences are more distinct. Figure 2 shows the Spearman correlation coefficients between the scores registered in the pre-election and post-election waves of CCES 2010. Here, the degree of association ranges from 0.68 to 0.82 in contrast to the 0.56 to 0.74 found previously in the TAPS data.

[Figure 2 ABOUT HERE]

In short, when conducting our analysis based on the 6-wave TAPS data, the results we find imply a considerably weaker stability of personality as captured by the TIPI scale than when using the CCES data GHDH utilize. The coefficients for each subscale across all TAPS waves are statistically different and lower than those computed between the two waves of CCES regardless of the method used. This finding threatens the claim of high correlations and stability between the subscales across waves, considering that we even observe a minimum average correlation of 0.59 for one of the subscales in TAPS. Table 4 offers a summary of the correlation coefficients found in each dataset that were computed by both the GHDH method and the one proposed in the present article.

[TABLE 4 ABOUT HERE]

Apart from drawing attention to the influence of modeling techniques and the effect of different correlation calculations, these findings most importantly question whether personality as captured by the TIPI scale can be understood as the stable concept it is commonly assumed to be. Moreover, the findings suggest that TIPI instability is not easily observed over short time frames, but becomes more obvious across longer periods of time. After all, the CCES data covers two waves that respondents answered with a separation of only 26 days on average. In contrast, the multiple waves on TAPS span a period of almost two years and therefore offer an invaluable source to test and retest TIPI reliability. This highlights the important role that data characteristics (such as the time frame in this case) can play in analysis that rely on personality as crucial variables. Moreover, another interesting finding relates to the differences between correlations among the different subscales. Our analysis indicates that certain dimensions could be more stable than others: while Extraversion registers an average correlation of 0.70, Openness does not even reach the 0.60. This might warrant more questions about the factors and reasons behind the stability of each individual dimension.

7.2 Measurement error

Another element that should be taken into account when assessing stability of survey instruments such as TIPI is measurement error. Measurement error could have an influence on the findings if the observed volatility (if any) were related to factors independent from the true attitudes of a respondent (context, respondents' skills, clarity of the survey instrument, etc.).

We account for potential measurement error throughout our analysis in a number of ways. First, the TAPS format itself limits some common sources of measurement error. For example, since TAPS is a self-reported online survey, issues such as interviewer bias or coding mistakes are eliminated. Moreover, by averaging multiple survey items and building the additive TIPI scale, we are achieving more accurate overall measures by "neutralizing" potential deviations from true attitudes.[10] In this same line, we also considered other alternatives that suggest averaging correlations through multiple waves as well as scale building. Most importantly, we also implement the main method suggested by Ansolabehere et al. (2008) to correct correlations in light of measurement

[10]This is one of the techniques suggested in Ansolabehere, Rodden, & Snyder (2008) to minimize measurement error.

error: estimation of parameters of the measurement model.

The standard measurement model applied to our analysis can be defined as:

$$W_{ikm} = X_{im} + e_{ik} \tag{3}$$

where W_{ikm} represents the observed answers to the k traits of each subscale m in wave i, X_{im} the *true* score for a given subscale we intend to measure and e_{ikm} a random error term. This implies that simple correlations between W_{ikm} will yield biased results due to the error term.

Since we are interested in the correlation between the "true scores" for each subscale X_m through i waves, we use the formula for each of them:

$$\rho_{X_i,X_j} = \frac{K-1}{K/\rho_{\bar{W}_i,\bar{W}_j} - \overline{1/\rho_{W_i,W_j}}} \qquad \forall\, i \neq j \tag{4}$$

and,

$$\bar{W}_i = \frac{1}{K}\sum_{k=1}^{K} W_{ik} \tag{5}$$

Table 5 presents both the regular correlation coefficients and the corrected coefficients computed with formula (4). As we can see, once we account for measurement error, the magnitude of the correlations significantly increases and reaches levels that suggest a high level of stability of the different subscales. In general, there is an increase between 0.10 and 0.22 in the degree of association for any given subscale across the different waves. This contrasts with previous findings presented above and gives support to the claims made by Converse (1964) and Achen (1975) about the underestimation of stability due to measurement errors.

[TABLE 5 ABOUT HERE]

These results should be interpreted with caution. The error model assumes that the K traits measure a single issue (X_m) and, accordingly, that there exists a high level of correlation between them (within a specific wave). However, it is known that TIPI lacks internal consistency within each subscale and that the correlation between the traits is low (Gosling et al. 2003). As a conse-

quence, the correction technique will tend to over-report corrected correlations, as the coefficients will be increased by more than they should based purely on the measurement error.

Overall, our findings suggest moderate stability of TIPI after accounting for measurement error and comparing different sources of data and methods. We argue that there are three specific reasons to be concerned about the general levels of correlation. First, the variables analyzed measure the same object, are worded identically and asked in very similar contexts, all of which should reduce different possible sources of measurement error. Second, we should consider that the phenomenon intended to measure by the TIPI battery (personality) is assumed to be an *extremely* stable trait, especially across short time periods, as claimed by the literature. Third, although the lapses between one survey and another are long enough to allow for changes in respondents' contexts, they can still be defined as relatively short term and will therefore usually not be related to drastic life changes (e.g. childhood to adolescence), that could be argued to significantly influence personality. Consequently, we would expect higher correlation coefficients than we observe from variables that are truly stable. For example, the correlation coefficients for variables gender and age, which are both asked upon entry into the TAPS panel and then again in December 2013, are between .92 and .98. Such a finding suggests that TIPI does not reach the same level of stability over time.

We can also compare these correlations to other variables that are generally considered stable in the political science literature. Although there is an ongoing debate about the volatility of party identification, there exists plenty of evidence in the literature that suggests a high level of stability of this characteristic. Green, Palmquist, & Schickler (2002, 69) offer evidence to support this claim. After analyzing different panel surveys over either a year or periods of two years, these authors find that correlations of party identification over time in a given survey range from 0.965 to 0.989 [11]. These numbers dwarf those found and shown in Figure 1 and suggest some degree of volatility in personality that has been overlooked until now.

[11] While party identification is generally measured on a 3-point-scale, each TIPI item has 7 possible values. Accordingly, one might suspect that this bigger range of possible answers makes TIPI harder to answer and could lead to more variation. To control for this possibility and make the results more comparable, we conducted an additional set of correlation analyses for which we recoded TIPI responses into 3 categories (original values of 1 and 2 were aggregated into a first category; 3, 4 and 5 into a second one, and 6 and 7 into a third category). Once we recalculated the new TIPI scales, we analyzed the degree of association between waves. The results show even lower correlations than those reported with the original coding.

This finding now leads us to a new question: what are the variables related to personality variation? We explore some potential answers in the following section.

8 Associations Between Volatility and Covariates

After showing that respondents' personality as measured by the TIPI shows variation across time, we will now try to determine the characteristics of the respondents that are associated with individual-level personality volatility. We do so by running a set of OLS regressions on the variability measure of each TIPI subscale and the additive scale that aim to capture the effects of those variables that are most commonly found to influence personality traits. More specifically, we will examine a general model that combines a set of demographic characteristics and political variables in order to analyze how they affect (in)stability in respondents' personality.[12] The number of waves of the panel in which a given respondent participated is included in all models to account for potential effects of survey participation itself, such as panel conditioning, as well as to control for differences in variability in the outcome variable associated with the total number of observations per respondent. Recall that the TIPI variability measure reflects the variance the respective TIPI score for a given respondent across all waves they have answered.

[TABLE 6 HERE]

Table 6 shows the estimated coefficients and standard errors for six pooled OLS regressions based on each of the TIPI subscales and the aggregate scale.[13] Education is the only variable that has a significant effect on variation across all five subscales and the aggregated TIPI scale. For all six models, the higher a respondent's education, the lower their personality variability across the six waves of our study. With education being coded on a 12-point scale, the coefficient of -0.259

[12]We also ran all analyses based on separate models for demographics and political variables. The result confirm the substantive findings reported below and are available from the authors upon request.

[13]Before conducting the main analysis, we tested for potential multicollinearity of the variables by analyzing all possible pair-wise correlations between them. Only 7 pairs of variables are correlated at $r \geq 0.40$: having kids and age ($r = -0.40$), income and education ($r = 0.43$), political knowledge and education ($r = 0.41$), Democratic party ID and ideology ($r = -0.43$), Republican party ID and ideology ($r = 0.46$), Democratic and Republican party ID ($r = -0.45$), and political knowledge and political interest ($r = -0.43$). We impute missing data so the regression results are based upon generating 10 imputed datasets using the *mice* package in *R*. Moreover, we also ran our analysis on the non-imputed dataset. The results confirm the substantive findings reported below and can be obtained from the authors upon request.

for Model (6) indicates that the variance of the overall TIPI scale will be 0.78 higher for somebody with a high school diploma as compared to a respondent with a bachelor's degree (and even 1.56 higher than for somebody with a doctorate). To better interpret those numbers: the median variance on the aggregated score across all respondents is 2.67. An increase of 1.56 to a variance of 4.23 would mean that a respondent no longer exhibits a median personality variation, but is now among the top third respondents with the highest variance.

The variables for variation in Presidential as well as Congress approval are significant in five of the six models estimated. In contrast to education, their coefficient estimates are positive, indicating that respondents with higher variation in their approval ratings also tend to exhibit higher variation in their personality measurements. These results suggest that exogenous factors that may affect political behaviors could also affect the personality measures of survey respondents. For example, catastrophic events may alter one's political outlook and their responses to survey items such as TIPI. The relationship the present multivariate analysis discovers, however, suggests that an individual's responses to such events may vary based on how the exogenous shock first affects her political perspectives.

Moreover, the estimates for political knowledge, political interest and change in political knowledge are found to have a significant effect on personality variation in three out of six models. Higher political knowledge and interest as well as an increase in the change of political knowledge between the waves in January 2012 and May 2013 are associated with an increase in personality stability as captured by TIPI. [14] Other results seem to follow conventional wisdom. For example, newly-weds exhibit a higher variation on the emotional stability subscale.[15]

In summary, after showing that individual personality as captured by the TIPI tends to vary significantly when measured repeatedly over time, we also find that it does so in systematic ways. Our analysis shows that both socio-demographic and political variables consistently and

[14] Since it could be argued that these variables are associated with respondents' levels of sophistication and therefore personality volatility might occur mainly among "bad survey takers", we also regressed party identification variation on the same sociodemographic and political variables included in the models above. Results show that from the covariates related to respondent's sophistication, only change in political knowledge turns out to be significant at the 95% level. Consequently, general volatility in responses cannot be completely attributed to the quality of the survey taker.

[15] While the adjusted R^2 values of these models are quite small (between .04 and .06), this does not diminish the substance of the results. Our goal is to test the relationships between the included political and socio-demographic variables and the outcome measurement. For this reason, we are more interested in the effects of explanatory variables than the overall fit statistics.

significantly affect variation in respondent's TIPI scores in ways that are in line with our theoretical expectations. This not only questions whether personality should be understood as a highly stable and time-invariant variable, but it also raises serious doubts about the nature of personality as predating both socio-demographic and – more importantly – political variables. In fact, since we show a number of political variables to actually systematically drive personality (in)stability, this might have far-reaching implications regarding the use of personality measures as explanatory variables in studies of political behavior.

9 Implications of Variability

To provide an example of the implications of variability in personality traits as measured by the TIPI, we illustrate the consequences of our findings by applying them to the study of Gerber, Huber, Doherty, Dowling and Ha (GHDDH; 2010). In their analysis, the authors show personality traits to affect both economic and social policy attitudes (in very distinct ways), as well as respondents' self-reported ideology. While the study advances our knowledge about the different effects personality has across different policy dimensions and in different contextual environments (most notably between blacks and whites in the United States), it is similar to previous research in that it neglects the possibility of personality variance within individuals across time.

Therefore, we review GHDDH's results in light of the findings presented above. More specifically, on page 120 in Table 3, the authors present six regression models analyzing the aggregate effect of personality on ideological self-placement and economic and social policy attitudes. To highlight the importance of our findings regarding an individual's personality variation, we fit Models 1, 3, and 5 presented in that table to our data and calculate predicted values for self-reported ideology as well as economic and social policy attitudes for all our respondents for each wave. As a result, each respondent has one predicted value for each of the three measures in each wave that they have completed the TIPI battery. Moreover, since the personality traits are the only time-variant covariates in those models, any differences we can observe in the ideology estimates, are purely due to personality variation across time.[16]

[16]The models in GHDDH make use of state-specific fixed effects, which the authors do not report and which we therefore also cannot include in our predictions. However, for the purpose of this application, the addition of fixed

Figure 3 illustrates the effect of personality variation on the estimate of self-reported ideology as based on the results reported in GHDDH, where ideology ranges from -2 (Very Conservative) to +2 (Very Liberal). Along the x-axis, respondents are ordered according to their mean predicted ideology across all waves in which they completed the TIPI battery. Values on the y-axis represent the absolute difference between a respondent's highest and lowest predicted ideology. For example, if one respondent has estimated ideology values of -0.25, 0, and 1 then their x-value would be 0.25, and their y-value 1.25. The solid line at the bottom of the figure indicates a difference of 0 between a respondent's highest and lowest predicted ideology score, and a quick examination reveals that almost none of our respondents lie on that line. More specifically, only 8 respondents out of a total of 1,730 for which we could predict ideology scores, have the same predicted scores for all of their waves. In contrast, the absolute majority of respondents shows some significant variation. A fair amount of respondents exhibit differences in their predicted ideology that are larger than 1, which again emphasizes the highly unstable nature of personality traits as measured by TIPI, given that ideology is only operationalized on a 5-point scale in this case.

[FIGURES 3 AND 4 ABOUT HERE]

To provide an example of what we would expect to find if personality traits were stable, Figure 4 illustrates the stability of self-reported age in the TAPS data. Respondents were asked to indicate their year of birth twice, once upon entering the panel and then again in December 2013. Similar to Figure 3, along the x-axis, respondents are ordered according to the mean of their self-reported year of birth in those two waves. Values on the y-axis represent the absolute difference between respondent's answers in the two waves. The solid line at the bottom of the figure indicates a difference of 0 between a respondent's two self-reported years of birth. As expected for a stable variable that can be assumed to suffer from little measurement error, the absolute majority of respondents lies on that line, indicating that they reported the exact same year of birth in both waves. This contrasts strongly with Figure 3, where almost no respondents were found to have the same predicted ideology scores for all of their waves.

In addition to Figure 3, Table 7 provides some additional summary statistics. Focusing on the row for self-reported ideology, we can see that the average prediction for self-reported ideology

effects would not make any meaningful difference.

across all respondents in all waves is -0.164, which is reasonably close to the reported mean in GHDDH (-0.155). $\bar{x}_{Min}$ and $\bar{x}_{Max}$ assume the lowest/highest value for each respondent across the different waves and report the respective mean across respondents. In other words, the average of the lowest predicted ideology value across all respondents is -0.422, whereas the average of the highest value is 0.079, which means that solely based on their variation in personality, an average respondent will vary in their predicted ideology between being conservative-leaning and being slightly liberal-leaning.

[TABLE 7 ABOUT HERE]

Figure 5 presents the same analysis for the effect of personality variation on the estimates of economic and social policy attitudes. Again, mean estimated attitudes are plotted along the x-axis and differences between the minimum and maximum attitudes along the y-axis. Just as with the predictions for ideology above, we can see that almost none of our respondents have the same predicted economic or social policy attitudes for all waves in which the completed the TIPI battery. Instead, there is again a significant amount of variation that is only based on variation in personality traits as captured by the TIPI. The corresponding results in Table 7 confirm those findings.

[FIGURE 5 ABOUT HERE]

To conclude, the application of our findings to GHDDH as an exemplary study that skillfully analyzes the link between personality traits and political orientations underlines the high importance of the consideration of personality variation over time. Given that the variation in personality as captured by the TIPI that we find in our analysis has the ability to significantly influence and change subsequent analysis that rely on personality traits (such as GHDDH), a review of the oftentimes assumed link between personality and political values, orientations and behavior is necessary.

10 Conclusion

The Big Five personality battery is a frequently used tool to quantify personality traits, which in turn are commonly assumed to predate and explain political behavior ranging from attitudes to

legislative voting. Just as many scholars before us have been concerned with the stability and fluidity of political attitudes (e.g. Green et al. 2004; Highton 2012), we examine how an individual's personality traits vary across time. Challenging the conventional wisdom regarding the stability of personality traits over time linking personality and political behavior, we present analyses that warrant a careful reconsideration of those assumptions. First, we find individual personality as captured by the TIPI to vary significantly over time. Second, we show this variability to be consistently associated with political and social variables, which raises serious doubts about the nature of personality as a factor that predates both socio-demographic and – more importantly – political variables.

Our study is unique in that it employs data by The American Panel Survey (TAPS), which allow us to analyze the variability as well various political associations with personality traits not only for one or multiple cross-sections, but for a representative panel of 2,014 respondents that covers six waves between February 2012 and November 2013. Based on this panel data, we show that both socio-demographic and political variables systematically and significantly affect variation in respondent's TIPI scores. More specifically, education, political knowledge and an over-time increase in political knowledge are associated with an increase in personality stability, whereas being employed and showing greater variation in Presidential as well as Congress approval are associated with significantly less stability in personality as captured by the TIPI.

In applying our results to a previous study by Gerber, Huber, Doherty, Dowling, & Ha (2010), we illustrate the far-reaching implications of our findings. The volatility in personality traits that our study uncovers has the ability to significantly influence and change the results of analyses, which rely on personality traits as explanatory variables. Consequently, we deem a careful review of the frequently assumed link between personality and political values, orientations and behavior necessary.

Certainly, these findings have numerous substantive implications, many of which are unexamined by the current study. Survey respondents' employment status, political knowledge, and public approval of national office holders are all functions of environmental phenomena. World events ranging from economic calamities to natural disasters, from armed conflicts to political scandals may be associated with an one's job loss, an increase political information, or feelings of cynicism about government. Additionally, several natural waxing and waning processes – be it

economic conditions or even seasonal weather patterns – may influence an individual's personality, which in turn affects their likelihood of voting (especially in primaries and other time-varying elections), willingness to engage in political groups, and the like.

In offering this alternative perspective where personality variations are linked – as psychology literature has suggested – to significant events, we seek to contribute to the esteemed scholarship that has long found personality to be highly stable. Going beyond the substantive implications of our research, it may be the case that scholars' measures of personality are not appropriate tools to explain political behavior. In turn, the results of this study should inspire future research, first, to realize that a single measure of an individual's personality may be an inaccurate snapshot, second, to account for the instability of personality measures like the TIPI over time, and third, to reconsider the assumption that personality predates political behavior.

References

Achen, C. H. (1975). Mass political attitudes and the survey response. *The American Political Science Review*, 1218–1231.

Alford, J. R., Funk, C. L., & Hibbing, J. R. (2005). Are political orientations genetically transmitted? *American Political Science Review*, *99*, 153–167.

Ansolabehere, S., Rodden, J., & Snyder, J. M. (2008). The strength of issues: Using multiple measures to gauge preference stability, ideological constraint, and issue voting. *American Political Science Review*, *102*(02), 215–232.

Benet-Martínez, V., & John, O. P. (1998). *Los Cinco Grandes* across cultures and ethnic groups: Multitrait-multimethod analyses of the big five in spanish and english. *Journal of Personality and Social Psychology*, *75*(3), 729.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter.* New York: Wiley.

Caprara, G. V., Barbaranelli, C., & Zimbardo, P. G. (2002). When parsimony subdues distinctiveness: Simplified public perceptions of politicians' personality. *Political Psychology*, *23*(1), pp. 77-95.

Carney, D. R., Jost, J. T., Gosling, S. D., & Potter, J. (2008). The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, *29*(6), 807–840.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*.

Cattell, H. E., & Mead, A. D. (2008). The sixteen personality factor questionaire (16pf). In G. M. G. Boyle & D. Saklofske (Eds.), *The sage handbook of personality theory and assessment: Volume 2 – personality measurement and testing.* London: SAGE Publications Ltd.

Clarke, H. D., & McCutcheon, A. L. (2009). The dynamics of party identification reconsidered. *Public Opinion Quarterly*, *73*(4), 704–728.

Converse, P. (1964). The nature of belief systems in mass publics. *Ideology and Discontent*, 206–61.

Denissen, J. J., & Penke, L. (2008). Motivational individual reaction norms underlying the five-factor model of personality: First steps towards a theory-based conceptual framework. *Journal of Research in Personality*, *42*(5), 1285–1302.

Dietrich, B. J., Lasley, S., Mondak, J. J., Remmel, M. L., & Turner, J. (2012). Personality and legislative politics: The big five trait dimensions among u.s. state legislators. *Political Psychology*, *33*(2), pp. 195-210.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*(1), 417–440.

Feldman, S. (1988). Structure and consistency in public opinion: The role of core beliefs and values. *American Journal of Political Science*, 416–440.

Gasper, J. T., & Reeves, A. (2011). Make it rain? retrospection and the attentive electorate in the context of natural disasters. *American Journal of Political Science*, *55*(2), 340–355.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2011). The big five personality traits in the political arena. *Annual Review of Political Science*, *14*(1), 265-287. doi: 10.1146/annurev-polisci-051010-111659

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2012). Personality and the strength and direction of partisan identification. *Political Behavior*, *34*(4), 653–688.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2013). Assessing the stability of psychological and political survey measures. *American Politics Research*, *41*(1), 54–75.

Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., & Ha, S. E. (2010). Personality and political attitudes: Relationships across issue domains and political contexts. *American Political Science Review*, *104*(01), 111–133.

Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., Raso, C., & Ha, S. E. (2011). Personality traits and participation in political processes. *The Journal of Politics*, *73*(3), pp. 692-706.

Goodwin, R. D., & Friedman, H. S. (2006). Health status and the five-factor personality traits in a nationally representative sample. *Journal of Health Psychology*, *11*(5), 643–654.

Goren, P. (2005). Party identification and core political values. *American Journal of Political Science*, *49*(4), 881–896.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality*, *37*(6), 504–528.

Green, D. P., Palmquist, B., & Schickler, E. (2002). *Partisan hearts and minds.* New Haven, CT: Yale University Press.

Green, D. P., Palmquist, B., & Schickler, E. (2004). *Partisan hearts and minds: Political parties and the social identities of voters*. Yale University Press.

Ha, S. E., Kim, S., & Jo, S. H. (2013). Personality traits and political participation: Evidence from south korea. *Political Psychology*, *34*(4), pp. 511-532.

Highton, B. (2012). Updating political evaluations: Policy attitudes, partisanship, and presidential assessments. *Political Behavior*, *34*(1), 57–78.

Hofmans, J., Kuppens, P., & Allik, J. (2008). Is short in length short in content? an examination of the domain representation of the ten item personality inventory scales in dutch language. *Personality and Individual Differences*, *45*(8), 750–755.

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research.* New York: The Guilford Press.

Lupia, A. (1994). Shortcuts versus encyclopedias: information and voting behavior in california insurance reform elections. *American Political Science Review*, *88*(01), 63–76.

MacKuen, M. B., Erikson, R. S., & Stimson, J. A. (1992). Peasants or bankers? the American electorate and the u.s. economy. *American Political Science Review*, *86*(3), pp. 597-611.

McCrae, R. R., & Costa Jr., P. T. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives.* New York: The Guilford Press.

McCrae, R. R., Costa Jr., P. T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., ... Smith, P. B. (2000). Nature over nurture: temperament, personality, and life span development. *Journal of Personality and Social Psychology*, *78*(1), 173.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175–215.

Mezquita, L., Stewart, S. H., & Ruipérez, M. Á. (2010). Big-five personality domains predict internal drinking motives in young adults. *Personality and Individual Differences*, *49*(3), 240–245.

Mondak, J. J., & Halperin, K. D. (2008). A framework for the study of personality and political behaviour. *British Journal of Political Science*, *38*(2), pp. 335-362.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. *Journal of research in Personality*, *41*(1), 203–212.

Rhodes, R., & Smith, N. (2006). Personality correlates of physical activity: a review and meta-analysis. *British Journal of Sports Medicine*, *40*(12), 958–965.

Saroglou, V. (2010). Religiousness as a cultural adaptation of basic traits: A five-factor model perspective. *Personality and Social Psychology Review*, *14*(1), 108–125.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, *8*(9), e73791.

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: set like plaster or persistent change? *Journal of Personality and Social Psychology*, *84*(5), 1041.

Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior*, *30*(3), pp. 341-366.

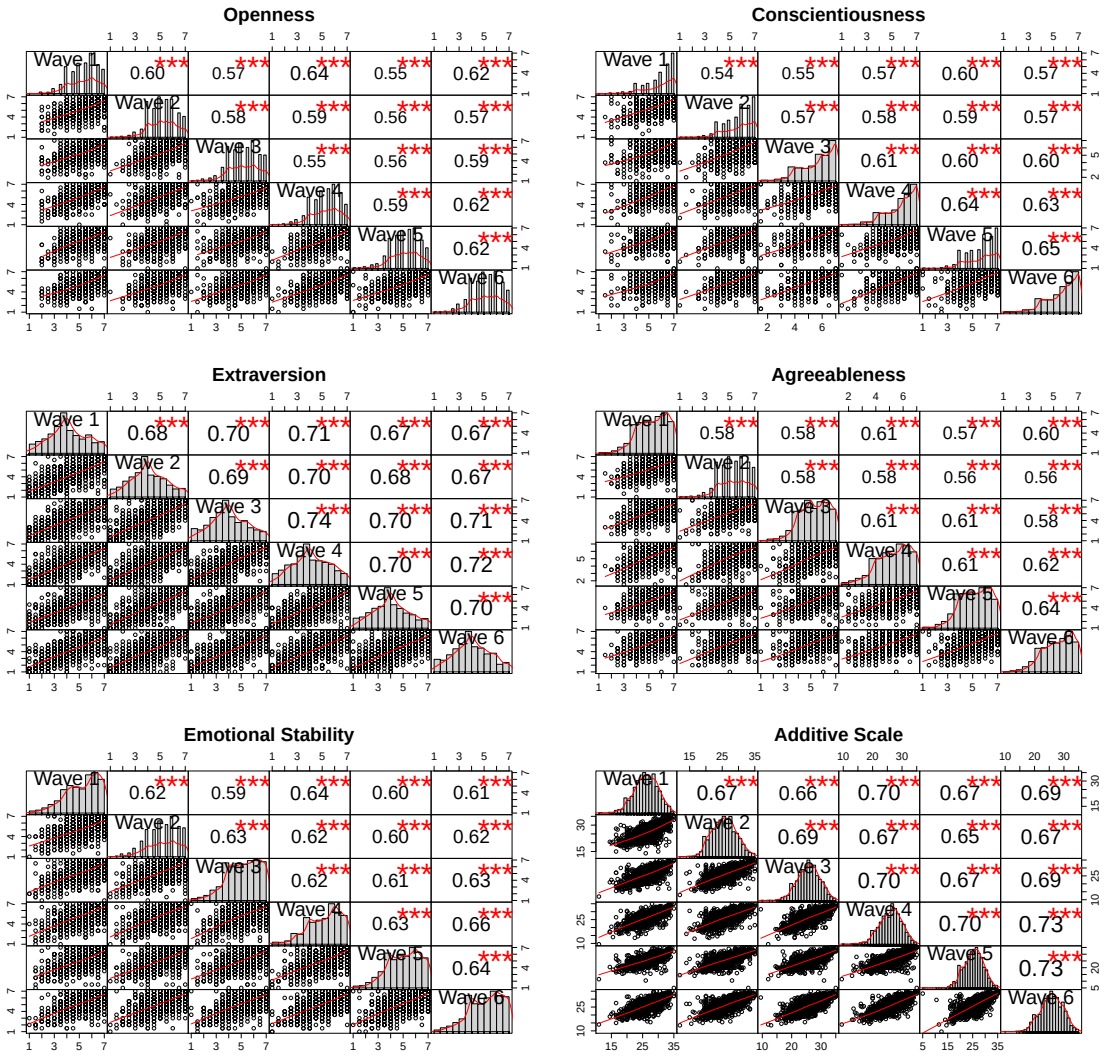
Sutin, A. R., Terracciano, A., Deiana, B., Uda, M., Schlessinger, D., Lakatta, E. G., & Costa Jr, P. T. (2010). Cholesterol, triglycerides, and the five-factor model of personality. *Biological psychology*, *84*(2), 186–191.

Wiggins, J. S. (1996). *The five-factor model of personality: Theoretical perspectives*. Guilford Press.

Zaller, J. (1992). *The nature and origins of mass opinion*. Cambridge university press.

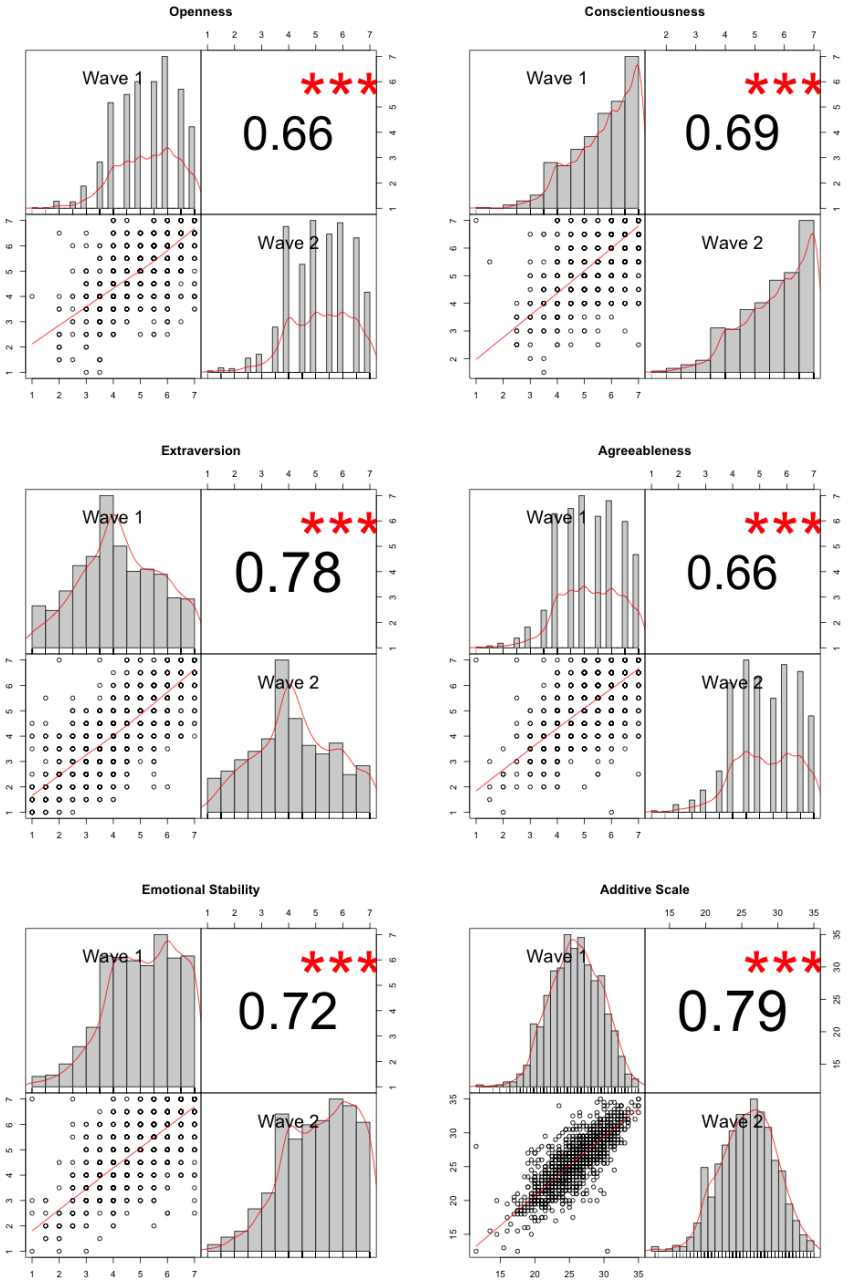
11 Tables and Graphs

Figure 1: Personality Correlation Across Waves (TAPS data)



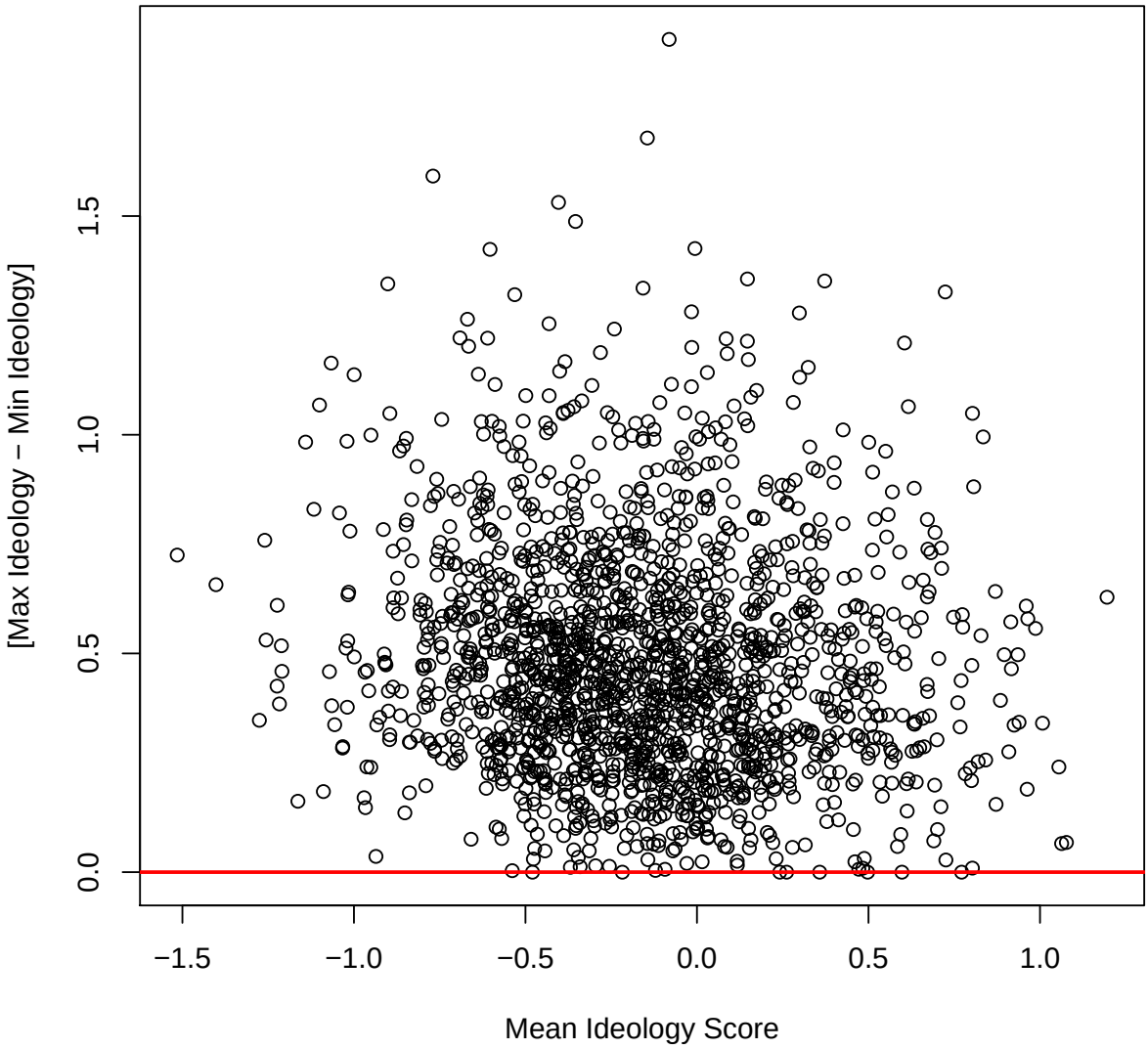
Note: The figure graphs correlations and variances of the different TIPI subscales across the six waves of TAPS data. The diagonal of each graph shows a set of histograms that provide information about the distribution of scores for a given TIPI dimension in a given wave. The lower diagonal of each graph plots the distribution of scores in one wave against the distribution in another wave. The upper diagonal presents Spearman correlation coefficients for the respective relationships and indicates their significance level. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Figure 2: Personality Correlation Across Waves (CCES data)



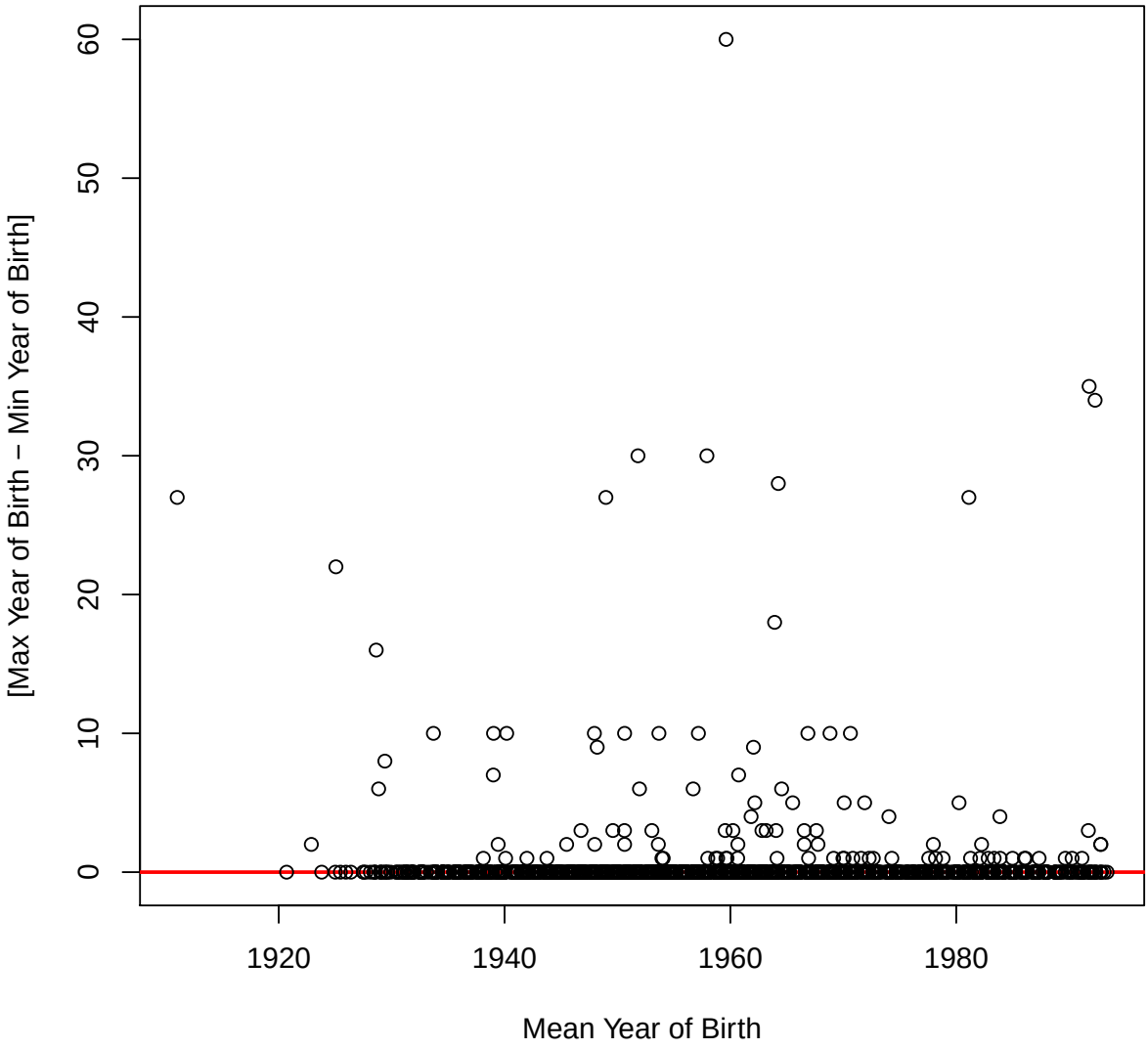
Note: The figure graphs correlations and variances of the different TIPI subscales across the two waves of CCES data. The diagonal of each graph shows a set of histograms that provide information about the distribution of scores for a given TIPI dimension in a given wave. The lower diagonal of each graph plots the distribution of scores in one wave against the distribution in another wave. The upper diagonal presents polychoric correlation coefficients for the respective relationships based on standardized scores.

Figure 3: Effect of Variation in Personality on Ideology Estimates



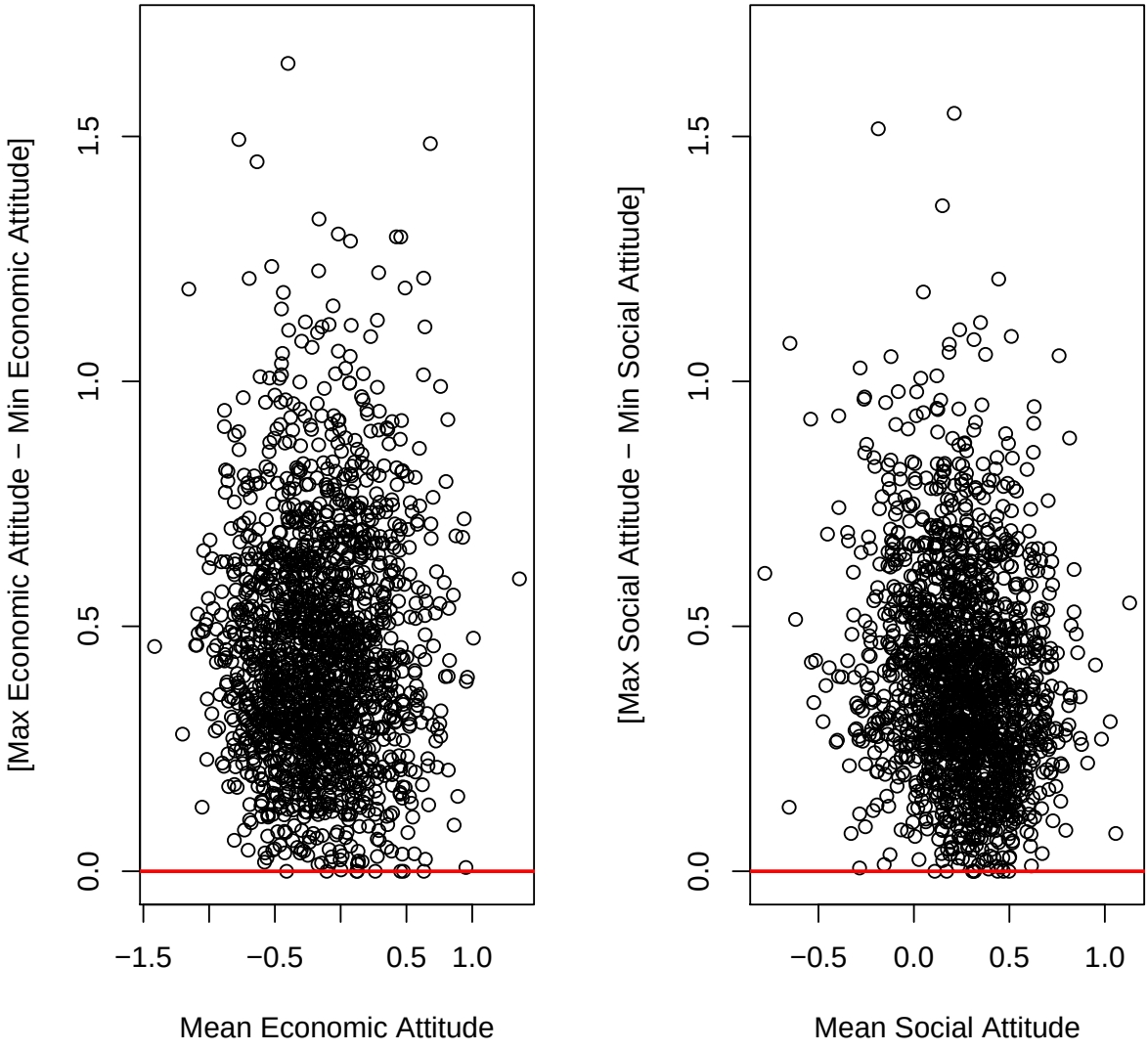
Note: The figure illustrates the effect of variation in personality on ideology estimates. Ideology is estimated based on the results and coefficients reported in Gerber, Huber, Doherty, Dowling and Ha (2010: 120, Table 3, Model 1). Personality variation is based on TAPS data. The figure plots respondents as ordered according to their mean predicted ideology across all waves in which they completed the TIPI battery along the x-axis. Values on the y-axis represent the absolute difference between a respondent's highest and lowest predicted ideology. The solid line at the bottom of the figure indicates a difference of 0 between a respondent's highest and lowest predicted ideology score and therefore corresponds to the y-value that would be expected for a perfectly stable variable.

Figure 4: Stability of Self-Reported Year of Birth in the TAPS Data



Note: The figure illustrates the "stability" of self-reported age in the TAPS data. Respondents were asked to indicate their year of birth twice, once upon entering the panel and then again in December 2013. Along the x-axis, respondents are ordered according to the mean of their self-reported year of birth in those two waves (values are jittered). Values on the y-axis represent the absolute difference between respondent's answers in the two waves. The solid line at the bottom of the figure indicates a difference of 0 between a respondent's two self-reported years of birth and therefore corresponds to the y-value that would be expected for a perfectly stable variable such as year of birth, when it is measured without any measurement error.

Figure 5: Effect of Variation in Personality on Policy Attitude Estimates



Note: The figure illustrates the effect of variation in personality on economic and social policy attitudes estimates. Policy attitudes are estimated based on the results and coefficients reported in Gerber, Huber, Doherty, Dowling and Ha (2010: 120, Table 3, Models 3 and 5). Personality variation is based on TAPS data. The figure plots respondents as ordered according to their mean predicted policy attitudes across all waves in which they completed the TIPI battery along the x-axis. Values on the y-axis represent the absolute difference between a respondent's highest and lowest predicted policy attitudes. The solid line at the bottom of the figure indicates a difference of 0 between a respondent's highest and lowest predicted policy attitudes and therefore corresponds to the y-value that would be expected for a perfectly stable variable.

Table 1: Descriptive statistics: outcome variable and covariates (without imputation)

	min	max	mean	N
TIPI variance	0.00	157.75	4.76	2014
Waves completed	2.00	6.00	4.60	2014
Age	18.00	100.00	50.84	1757
Gender [Female=1]	0.00	1.00	0.52	2011
Education [in years]	2.00	15.00	11.06	2000
Income [16 categories]	0.00	15.00	5.59	1901
Race [Black=1]	0.00	1.00	0.09	1975
Race [Hispanic=1]	0.00	1.00	0.12	1975
Have kids	0.00	1.00	0.31	1806
Religiosity [6 point scale]	0.00	5.00	2.35	1805
Employment status [Employed=1]	0.00	1.00	0.41	1968
Happiness [5 point scale]	0.00	4.00	0.94	1807
Job loss	0.00	1.00	0.11	2014
Got married	0.00	1.00	0.02	2014
Got divorced	0.00	1.00	0.00	2014
Symbolic Ideology [High=Conservative, 7 point scale]	-3.00	3.00	0.10	2005
Political interest [4 point scale]	0.00	3.00	0.84	2001
Party ID [Democrat=1]	0.00	1.00	0.37	1514
Party ID [Republican=1]	0.00	1.00	0.26	1514
Political knowledge [10 question battery]	0.00	10.00	6.65	1347
Party ID variation	0.00	4.50	0.28	1711
Presidential approval variation	0.00	5.33	0.44	2003
Congress approval variation	0.00	5.33	0.64	2004
Political knowledge change	-7.00	7.00	0.07	1168

Note: Based on respondents that completed at least 2 waves where TIPI questions were included. The number of individuals that met this criterion is 2,014. 202 respondents completed 2 waves, 339 completed 3 waves, 240 completed 4 waves, 513 completed 5 waves, and 720 completed all 6 waves. The dataset covers six waves of The American Panel Survey from February 2012 to November 2013.

Table 2: Summary statistics for across-wave correlations of the different subscales

	min	max	mean
Openness	0.554	0.644	0.586
Conscientiousness	0.541	0.648	0.591
Extraversion	0.667	0.744	0.697
Agreeableness	0.560	0.638	0.593
Emotional Stability	0.590	0.659	0.621
Aggregated TIPI Scale	0.647	0.728	0.685

Note: Correlations computed between the scores for each personality dimension across 6 waves from February 2012 to November 2013 of The American Panel Survey. This is a summary of the information presented in Figure 2.

Table 3: Summary statistics for scores on the aggregated TIPI-scale

	min	max	mean	N
Wave 1	11.00	35.00	26.17	1419
Wave 2	12.00	35.00	25.69	1573
Wave 3	9.50	35.00	25.70	1583
Wave 4	9.50	35.00	26.04	1686
Wave 5	5.00	35.00	25.44	1601
Wave 6	9.50	35.00	25.77	1607
Average	9.42	35.00	25.80	1578
Variance	0.00	157.80	4.76	2014

Note: The aggregated TIPI scale is calculated by adding the scores for the 5 personality dimensions obtained by each individual per wave. The study only includes respondents that completed at least 2 of the 6 waves under analysis.

Table 4: Comparison of correlations

Method	Data	Open.	Cons.	Extra.	Agree.	Emot. Stab.	Add. Scale
Gerber et al.	CCES[1]	0.68	0.70	0.70	0.73	0.82	NA[3]
	TAPS[2]	0.60	0.62	0.71	0.61	0.64	NA[3]
Boston et al.	CCES	0.66	0.69	0.78	0.66	0.72	0.79
	TAPS[2]	0.59	0.59	0.70	0.59	0.62	0.69

Notes: Coefficients reported come from correlations based on Polychoric correlations of standardized scores (Gerber et al.) and Spearman's rank correlations (Boston et al.), when applied to both CCES and TAPS data.
[1] Numbers as reported in Gerber et al. 2013. We replicated their method and obtained very similar coefficients.
[2] Average of correlations across 6 waves.
[3] The number of categories limits the calculation of polychoric correlations.

Table 5: Comparison of coefficients with and without measurement error correction

Openness

	Correlations with measurement error correction				
Wave 1	0.76	0.69	0.84	0.66	0.80
0.67	**Wave 2**	0.68	0.71	0.64	0.72
0.57	0.58	**Wave 3**	0.67	0.66	0.73
0.64	0.59	0.55	**Wave 4**	0.72	0.80
0.55	0.56	0.56	0.59	**Wave 5**	0.72
0.62	0.57	0.59	0.59	0.62	**Wave 6**
Correlation without measurement error correction					

Conscientiousness

	Correlations with measurement error correction				
Wave 1	0.72	0.90	0.95	0.74	0.79
0.54	**Wave 2**	0.79	0.78	0.71	0.82
0.55	0.57	**Wave 3**	0.81	0.79	0.80
0.57	0.58	0.61	**Wave 4**	0.76	0.87
0.60	0.59	0.60	0.64	**Wave 5**	0.81
0.57	0.57	0.60	0.63	0.62	**Wave 6**
Correlation without measurement error correction					

Extraversion

	Correlations with measurement error correction				
Wave 1	0.81	0.83	0.84	0.78	0.84
0.68	**Wave 2**	0.84	0.84	0.82	0.81
0.70	0.69	**Wave 3**	0.89	0.86	0.87
0.71	0.70	0.74	**Wave 4**	0.85	0.89
0.67	0.68	0.70	0.70	**Wave 5**	0.85
0.67	0.67	0.71	0.72	0.70	**Wave 6**
Correlation without measurement error correction					

Agreeableness

	Correlations with measurement error correction				
Wave 1	0.68	0.73	0.74	0.64	0.71
0.58	**Wave 2**	0.67	0.73	0.62	0.71
0.58	0.58	**Wave 3**	0.71	0.67	0.70
0.61	0.58	0.61	**Wave 4**	0.68	0.71
0.57	0.56	0.61	0.61	**Wave 5**	0.71
0.60	0.56	0.58	0.62	0.64	**Wave 6**
Correlation without measurement error correction					

Emotional Stability

	Correlations with measurement error correction				
Wave 1	0.80	0.80	0.82	0.75	0.82
0.62	**Wave 2**	0.79	0.78	0.74	0.79
0.59	0.63	**Wave 3**	0.82	0.75	0.82
0.64	0.62	0.62	**Wave 4**	0.79	0.86
0.60	0.60	0.61	0.63	**Wave 5**	0.81
0.61	0.62	0.63	0.66	0.64	**Wave 6**
Correlation without measurement error correction					

Table 6: Effect of political and sociodemographic variables on TIPI subscales variation

	Dependent variable: Variation in					
	Openness	Conscientiousness	Extraversion	Agreeableness	Emotional stability	Aggregated scale
	(1)	(2)	(3)	(4)	(5)	(6)
Waves completed	-0.036^{***}	-0.036^{**}	-0.018	0.006	-0.010	-0.077
	(0.013)	(0.014)	(0.013)	(0.012)	(0.013)	(0.128)
Age	0.003	0.004^{**}	0.002	0.001	0.002	0.017
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.013)
Gender [Female=1]	-0.023	-0.015	-0.018	-0.047	0.038	-0.220
	(0.038)	(0.039)	(0.036)	(0.033)	(0.038)	(0.364)
Education	-0.021^{*}	-0.024^{*}	-0.031^{***}	-0.018^{*}	-0.026^{**}	-0.259^{**}
	(0.012)	(0.012)	(0.011)	(0.010)	(0.013)	(0.110)
Income	-0.008	-0.009	-0.001	-0.001	0.004	0.039
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.056)
Race [Black=1]	0.087	0.233^{***}	0.151^{**}	0.074	0.076	0.851
	(0.069)	(0.071)	(0.064)	(0.059)	(0.069)	(0.670)
Race [Hispanic=1]	-0.048	-0.044	0.033	-0.041	-0.022	-1.087^{*}
	(0.062)	(0.063)	(0.057)	(0.053)	(0.060)	(0.578)
Have kids	0.003	0.019	0.111^{**}	-0.018	0.013	0.232
	(0.046)	(0.049)	(0.044)	(0.039)	(0.049)	(0.408)
Church attendance	0.004	-0.001	0.006	0.004	-0.009	0.024
	(0.012)	(0.013)	(0.012)	(0.010)	(0.012)	(0.123)
Employment status [Employed=1]	0.016	0.012	0.040	0.055	0.103^{**}	1.010^{**}
	(0.044)	(0.043)	(0.040)	(0.036)	(0.042)	(0.395)
Happiness	0.007	-0.026	-0.039	-0.001	0.002	-0.302
	(0.027)	(0.026)	(0.025)	(0.023)	(0.025)	(0.243)
Job loss	0.042	0.007	0.075	0.042	0.099^{*}	0.282
	(0.060)	(0.061)	(0.056)	(0.051)	(0.059)	(0.560)
Got married	0.050	0.163	-0.009	0.206^{*}	0.441^{***}	2.083
	(0.144)	(0.148)	(0.136)	(0.125)	(0.143)	(1.356)
Got divorced	-0.394	-0.345	-0.157	0.089	-0.261	-3.196
	(0.361)	(0.370)	(0.342)	(0.312)	(0.359)	(3.410)
Ideology	0.008	-0.014	-0.003	0.004	0.011	-0.030
	(0.014)	(0.014)	(0.014)	(0.012)	(0.015)	(0.131)
Political interest	0.015	-0.058^{**}	-0.065^{***}	-0.072^{***}	-0.017	-0.307
	(0.025)	(0.027)	(0.025)	(0.023)	(0.028)	(0.235)
Party ID [Democrat=1]	-0.017	0.025	-0.035	0.049	0.060	-0.159
	(0.052)	(0.057)	(0.052)	(0.044)	(0.060)	(0.516)
Party ID [Republican=1]	-0.055	0.029	-0.043	0.040	0.073	-0.719
	(0.067)	(0.060)	(0.055)	(0.051)	(0.065)	(0.650)
Political knowledge	-0.017	-0.056^{***}	0.009	-0.044^{***}	-0.034	-0.473^{***}
	(0.013)	(0.016)	(0.015)	(0.013)	(0.021)	(0.116)
Party ID variation	-0.019	-0.041	0.033	0.025	-0.029	-0.278
	(0.034)	(0.035)	(0.031)	(0.027)	(0.031)	(0.310)
Presidential approval variation	0.041	0.071^{***}	0.083^{***}	0.046^{**}	0.076^{***}	0.879^{***}
	(0.026)	(0.027)	(0.025)	(0.023)	(0.026)	(0.245)
Congress approval variation	0.101^{***}	0.016	0.089^{***}	0.039^{*}	0.056^{**}	0.535^{**}
	(0.024)	(0.025)	(0.023)	(0.021)	(0.024)	(0.231)
Political knowledge change	-0.036^{*}	-0.053^{***}	0.010	0.009	-0.033	-0.413^{**}
	(0.019)	(0.018)	(0.019)	(0.015)	(0.022)	(0.168)
Constant	0.893^{***}	1.232^{***}	0.840^{***}	0.950^{***}	0.908^{***}	9.632^{***}
	(0.175)	(0.183)	(0.170)	(0.153)	(0.169)	(1.583)
Observations	2,014	2,014	2,014	2,014	2,014	2,014
AIC (mean)	4823.31	4912.68	4602.35	4232.05	4788.76	13868
Adjusted R^2	0.046	0.068	0.044	0.049	0.050	0.059

Note: Pooled results from 10 imputed datasets *p<0.1; **p<0.05; ***p<0.01

Table 7: Application of Findings to Gerber et al. 2010

	$\bar{x}_{Min}$	$\bar{x}_{Mean}$	$\bar{x}_{Max}$
Predicted Self-reported Ideology	-0.4215	-0.1642	0.0791
Predicted Economic Policy Attitudes	-0.3914	-0.1654	0.0595
Predicted Social Policy Attitudes	0.0615	0.2618	0.4518

Note: Cell entries are based on the models presented in Gerber et al (2010: 120), Table 3, Models 1, 3 and 5 and data from six waves of The American Panel Survey. $\bar{x}_{Mean}$ reports the average mean value across all respondents in all waves. $\bar{x}_{Min}$ and $\bar{x}_{Max}$ take the lowest/highest value for each respondent across the different waves and report the respective mean across respondents. Self-reported Ideology is measured on a 5-point-scale (-2 = Very Conservative, 2 = Very Liberal, 0 = Not Sure). Economic and Social Policy Attitudes are scaled to have a mean of 0 and a standard deviation of 1. For more information on models and measurement, refer to Gerber et al (2010).