Econographics*

Jonathan Chapman
NYUAD
jchapman@nyu.edu
www.jnchapman.com

Mark Dean
Columbia
mark.dean@columbia.edu
www.columbia.edu/~md3405/

Pietro Ortoleva
Princeton
pietro.ortoleva@princeton.edu
pietroortoleva.com/

Erik Snowberg
Caltech, UBC, CESifo, NBER
snowberg@caltech.edu
hss.caltech.edu/~snowberg/

Colin Camerer
Caltech
camerer@hss.caltech.edu
hss.caltech.edu/~camerer/

January 24, 2019

Abstract

We study the pattern of correlations across a large number of behavioral regularities, with the goal of creating an empirical basis for more comprehensive theories of decision-making. We elicit 21 behaviors using an incentivized survey on a representative sample ($n = 1{,}000$) of the U.S. population. Our data show a clear and relatively simple structure underlying the correlations between these measures. Using principal components analysis, we reduce the 21 variables to six components corresponding to clear clusters of high correlations. We examine the relationship between these components, cognitive ability, and demographics. Common extant theories explain some of the patterns in our data, but each theory we examine is also inconsistent with some patterns.

JEL Classifications: C90, D64, D81, D90, D91

Keywords: Econographics, Reciprocity, Altruism, Trust, Costly Third-Party Punishment, Inequality Aversion, Risk Aversion, Common-Ratio Effect, Endowment Effect, WTA, WTP, Ambiguity Aversion, Compound Lottery Aversion, Discounting, Overconfidence, Cognitive Ability, Demographics

*We thank Douglas Bernheim, Ben Enke, Benedetto De Martino, Stefano Della Vigna, Xavier Gabaix, Daniel Gottlieb, Eric Johnson, David Laibson, Graham Loomes, Ulrike Malmandier, Matthew Rabin, Jörg Spenkuch, Victor Stango, Dimitri Taubinsky, Peter Wakker, Michael Woodford, Jonathan Zinman, and the participants of seminars and conferences for their useful comments and suggestions. Daniel Chawla provided excellent research assistance. Camerer, Ortoleva, and Snowberg gratefully acknowledge the financial support of NSF Grant SMA-1329195.

1 Introduction

Decades of research in economics and psychology has identified a large number of behavioral regularities—specific patterns of behavior present in the choices of a large fraction of decision-makers—that run counter to the standard model of economic decision-making. This has lead to an enormous amount of research aimed at understanding each of these behaviors. However, significantly less work has gone into linking these regularities with each other, either theoretically or empirically. Instead, most regularities have been studied in isolation, with specific models developed for each one. This has lead to concerns about model proliferation in behavioral economics.[1]

In this paper, we study the pattern of correlations across a large number of behavioral regularities, with the goal of creating an empirical basis for more comprehensive theories of decision-making. We focus on standard measures of preferences and beliefs in the (behavioral) economics literature, as described in Section 2. We elicit 21 behaviors from a representative sample of the U.S. population ($n = 1{,}000$). The *econographics*—a neologism describing measures of behaviors that have an effect on economic decision-making—cover broad areas of social preferences (8 measures), attitudes towards risk and uncertainty (9 measures), overconfidence (3 measures), and time preferences (1 measure).[2] Whenever possible, these elicitations are incentivized: the compensation participants receive depends on their choices. We also include two measures of cognitive abilities and several demographic variables. Moreover, we took steps to limit attenuation of correlations due to measurement error by eliciting many of our measures twice and using the Obviously Related Instrumental Variables (ORIV) technique developed in Gillen et al. (2018).

Our approach allows us to document the underlying structure of these measures. This

[1] See, for example, Fudenberg (2006); Levine (2012); Kőszegi (2014); Bernheim (2016); and Chang and Ghisellini (2018). There are many important defenses of the wide range of models, including that they may be the most accurate representation of behavior, see, for example, Conlisk (1996); Kahneman (2003); DellaVigna (2009); Chetty (2015); and Thaler (2016).

[2] Each measure we consider here is an econographic, however, we do not measure all possible econographics. We focus on behaviors that are core areas of interest to behavioral economists in the domains of risk, social preferences, time, and overconfidence. See Section 2 for more detail.

structure, broadly speaking, could take one of three forms. First, there may be no discernible structure, suggesting the continued usefulness of examining and theorizing about behaviors in relative isolation. Second, there may be a structure that matches well with some existing theoretical approach, providing support for extensive use of that theory. Third, it could be that a structure exists, but it does not match any extant theory. Our results are somewhere between the second and third possibility: extant theories match some of the correlations we observe, but no theory explains the overall patterns well—even within a specific domain.

Some parts of the structure we reveal are the subject of significant, often contradictory, theoretical speculation. In the domain of social preferences there are several possible organizing principles: altruistic preferences, social welfare preferences (Charness and Rabin, 2002), inequality aversion (Fehr and Schmidt, 1999), and reciprocity (Rabin, 1993).[3] As an example, consider three of our measures: giving in the dictator game, and sender and receiver behavior in the trust game. If inequality aversion is the correct organizing principle, it should explain all three behaviors. That is, we should see a cluster of correlations encompassing all three measures. Alternatively, behavior in the dictator game could be driven by altruistic preferences, sender behavior in the trust game by inequality aversion, and receiver behavior in the trust game by reciprocity, leading to a much different pattern of correlations. Similarly, in the domain of risk and uncertainty there are many possible forces that could shape behavior: curvature of the utility function, probability weighting, loss aversion, and ambiguity aversion. Each could potentially drive behavior in several of our tasks.

Examining all links between the econographics—including many that are not the subject of extant theories—has a number of benefits for our goal of creating an empirical basis for more comprehensive theories of decision-making. This is especially true when compared with the standard approach of theorizing about a particular connection between two behaviors and then testing that theory. First, as correlations are not transitive, identifying underlying

[3]We use altruism to refer to a specific behavior—the amount given in the dictator game. We use "altruistic preferences" to mean more general models of altruism that may have implications for multiple behaviors.

structures requires observing all links (or lack thereof) between behaviors.[4] Second, once that structure is identified, its components can be examined to see if they line up with existing constructs—such as demographics or cognitive ability. Third, by measuring all of the behaviors simultaneously in a representative sample, we ensure that the patterns we identify are not due to shifting participant populations between studies.[5]

Describing, in a compact way, the 378 correlations between 21 econographic variables, as well as cognitive and demographic variables, is a daunting task. However, we are aided by the fact that many of the econographics fall into clusters of variables—which feature high intra-cluster correlations and low inter-cluster correlations. To summarize these clusters, we make use of principal components analysis (PCA), a statistical technique that produces components—linear combinations of variables—that explain as much variation in the underlying behaviors as possible, as discussed in Section 3. These components highlight latent dimensions underlying the econographic variables.[6]

Overall, our main finding is that there is a clear and relatively simple structure underlying our data. We can summarize 21 econographics with six components, as shown in Section 4, and summarized in Table 1. Two of these components underlie social preferences and beliefs about others, two underlie risk and uncertainty. One component has explanatory power in both the domains of social and risk preferences. A final component underlies overconfidence measures. Time preferences are spread across a number of components, but load most heavily on the Punishment component of social preferences. Each econographic, except for

[4]For example, suppose a theory predicts a relationship between behaviors A and B, and this has been tested, finding a correlation 0.5. Another theory predicts a relationship between behaviors B and C; this is tested in a different study, also finding a correlation of 0.5. Even if there is no theoretically predicted relationship between A and C, the correlation between A and C—which could be anywhere between 0 and 1—is important for understanding the true structure of these three behaviors. If the correlation is 1, then A and C are redundant, and so at most one of the two prior theories can be correct. If the correlation is 0, then $B = A + C$, and the two theories should be complimentary. If the correlation is 0.5, this suggests that an underlying latent factor determines all three behaviors, and the focus should be on developing a theory consistent with this fact.

[5]Moreover, a correlation that might seem large in isolation (that is, statistically significant), may be quite small in the context of other correlations.

[6]We also used more sophisticated machine learning techniques, such as clustering, but these did not produce additional insight. See Appendix B.

Table 1: Twenty-one econographics can be summarized with six components.

	Social Components				*Risk Components*		
Components	*Generosity*	*Punishment* (*Impulsivity*)	*Inequality Aversion/WTP*		*WTA*	*Uncertainty*	*Overconfidence*
Econographics	Reciprocity: High	Anti-social Punishment	Dislike Having More	WTP	WTA	Ambiguity Aversion	Overestimation
	Reciprocity: Low	Pro-social Punishment	Dislike Having Less	Risk Aversion: CR Certain	Risk Aversion: Gains	Compound Lottery Aversion	Overplacement
	Altruism	(Patience)		Risk Aversion: CR Lottery	Risk Aversion: Losses		Overprecision
	Trust		Components combine in joint analysis		Risk Aversion: Gain/Loss		

Notes: Simplified summary of patterns shown in Table 7. The names of econographics correspond with those in the literature, but are sometimes the same as the general principles discussed in the third paragraph of the introduction, see Footnote 3. For definitions of specific econographics, see Section 2. As noted in the text, time preferences load weakly on the Punishment component, slightly changing its meaning: this is indicated by the parentheses around that measure, and the alternative name of the component.

time preferences, loads heavily on only one of the six components. Existing models predict some of the relationships we observe, but, to our knowledge, none match the overall pattern we observe; suggesting the need for new theories taking these patterns into account.

In building towards our main finding, we are aided by a fact that was not ex ante obvious: risk and social preferences are largely—but not completely—independent. Thus, we first study social and risk preferences separately, before combining them together with measures of time preferences and overconfidence.

Our results show there is significant scope for representing our eight measures of social preferences in a more parsimonious way. The behaviors we measure break down into three clusters, as summarized in Table 1. In particular, altruism, trust, and two different types of reciprocity form a cluster. Pro- and anti-social punishment are a second cluster. Finally, two different types of inequality aversion—dislike of having more than another person and a dislike of having less—form a third cluster. In total, this suggests that differences in behavior within the broad domains we identify—generosity, punishment, and inequality aversion—are less substantial than differences in behavior across those domains.

Risk preferences show a structure that is less parsimonious than standard theory. Our nine measures of risk preferences form three clusters, as summarized in Table 1. Two of these contain risk attitudes. Willingness to pay (WTP) for a lottery ticket is part of one cluster along with risk aversion, as measured by lottery equivalents for a certain amount and a lottery. Willingness to accept (WTA) is part of another cluster together with risk aversion for gains, losses, and gains and losses, as measured by certainty equivalents. These two different clusters are largely unrelated to each other.[7] The third cluster consists of aversion to compound lotteries and ambiguity aversion. These behaviors are highly correlated, consistent with previous studies (Halevy, 2007; Dean and Ortoleva, 2015; Gillen et al., 2018). Our richer

[7] Note that this finding is distinct from the different "domains" of risk attitudes often discussed in psychology—see Weber and Johnson (2008) for a review—as we examine only one domain. Our findings are also distinct from economic studies that document poor correlations between different elicitations of risk attitudes—see Friedman et al. (2014) for a review—as we document a particular pattern of both high and low correlations between measures of risk attitudes.

data allow us to document that they are largely unrelated to other aspects of risk preferences.

Analyzing risk and social preferences, along with overconfidence and time preferences, results in six components, as summarized in Table 1. Two social components and two risk components are largely unchanged. The three overconfidence measures form a new component. Patience loads most heavily (and negatively) on the punishment component of social preferences. However, one social component—inequality aversion—and one risk component—related to WTP—combine with each other. Similar links have been suggested in prior research based on the idea that from behind the "veil of ignorance" more inequality creates more risk (Carlsson et al., 2005). However, our results indicate this is only true for one of the three components of behaviors related to risk and uncertainty.

In order to facilitate the use of our findings in future theorizing, we examine, in Section 5, the correlations between the components we identify and various other observables, specifically, cognitive abilities and demographics. There are a number of potentially interesting correlations. Four of the six components are correlated with cognitive abilities. The strongest relationship is with the component containing measures of punishment and the heaviest weighting of time preferences. In particular, higher cognitive ability is associated with a lower propensity to punish and greater patience. Most correlations indicate a typical positive relationship between cognitive ability and normatively rational behavior. However, there is one exception: high-cognitive ability participants are more generous. Strong relationships between components and demographics are less common. The strongest links are with education and income, although it is important to note that these variables do not seem to be simply proxying for cognitive ability, but have their own individual effect.

Section 6 draws out the implications for existing theories, some of which were alluded to above. We show that while there is some overlap between our data and common theories of social and risk preferences, no theory explains the overall patterns well, even within a specific domain. For social preferences, the three outcome-based models we consider (altruistic preferences, social welfare, and inequality aversion) predict that many of the measures that make up the Generosity component should be positively correlated. However, these theories also predict relationships between these measures and those that make up the Inequality Aversion component differ from those we observe. For risk preferences, models delineating risk and uncertainty do not match our data. Moreover, although reference points are clearly important in explaining the split between risk aversion measures associated with WTA and those with WTP, common theories of reference dependence make additional predictions that are not consistent with our data.

Our study is uniquely suited to our aim of understanding the empirical basis for more parsimonious behavioral models across the risk, social, time, and confidence domains. There is a significant literature, summarized in Section 7, that examines the correlation between two or three behavioral measures and/or cognitive ability. These more limited sets of relationships are not able to recover the nuanced structure we document here. There are a small number of studies which do measure more behaviors (Burks et al., 2009; Dean and Ortoleva, 2015; Stango et al., 2017a; Dohmen et al., 2018). As described in the same section, these studies differ from ours in terms of behaviors examined, incentivization, representativeness, and how they deal with the attenuating effects of measurement error.

We conclude, in Section 8, by providing some further interpretation of our results, and their limits. One limit is worth previewing here: we have chosen one way to organize the description of the relationships in our data. Many others are possible. In order to keep the analysis simple, we emphasize the difference between high and low correlations between econographics, and pay less attention to middling correlations. While these middling correlations may be of great interest to some, we focus on the bigger picture: our data suggest an underlying organization for a long list of behaviors that are often treated independently.

2 Design and Econographics

Four design decisions followed from our goal of providing an empirical basis for the underlying structure of theories of behavioral decision-making. The most important design decision was the selection of behaviors to elicit. The other three decisions were: to focus on behavior rather than parameters of models, to incentivize as many measures as possible, and to use a representative sample. We discuss these decisions, and then turn to a description of the econographics we chose to measure.[8]

We follow a two-part process to collect a slate of measures capturing the core behaviors of interest to behavioral economists. First, we chose to focus on three domains that have been at the center of the study of preferences and beliefs in economics: risk, social, and time preferences.[9] We also included measures of overconfidence, as these have recently seen rapid increases in scholarly interest. Our focus on preferences and beliefs implies that we left out behaviors generally considered biases or mistakes, such as base-rate neglect or the law of small numbers.[10] Second, within these domains we focused on the behaviors that are the core areas of interest in behavioral economics. Within risk preferences, for example, this meant a focus on behaviors that violate the standard model—each of which has generated voluminous theoretical and experimental literatures. To assess these behaviors, we make use of existing approaches that allow for a continuous measure of the extent to which a behavior is exhibited. In the interest of concision, we primarily refer to the research that inspired each measure, rather than attempting to review the entire literature around each behavior.[11]

[8] Appendix A gives implementation details omitted here. The specific question wordings, screenshots, and other details of experimental design can be found online at hss.caltech.edu/∼snowberg/wep.html. Implementation and analytic techniques are discussed in Section 3.

[9] We were unable to measure present bias, see Footnote 18.

[10] Our focus on preferences and beliefs—rather than mistakes and biases—and on continuous measures of behaviors, are the major points of differentiation between our work and Stango et al. (2017a). See Section 7 for more detail.

[11] Ultimately, these criteria, and the way we used them to derive a set of measures, involved some judgement. As there are very few attempts to measure econographic profiles (see Section 7 for a review), we believe our list of measures is likely to be useful to a wide swath of economists. However, the omission of particular measures may make our data less useful to specific researchers. By analogy to demographics, a researcher who studies testosterone differences in the general population might get little out of a demographic profile that does not include the length of individuals' first and third fingers (Pearson and Schipper, 2012),

As our focus is on understanding the underlying structure of behavior, we study behavioral measures. For example, when measuring risk aversion we elicit certainty equivalents for lotteries, and use (a linear transform of) them in our analyses, rather than trying to identify a parameter of some utility function (for example a constant relative risk aversion, or CRRA, utility function). This allows our results to be used to inspire theories that connect behaviors without committing to specific functional forms, which are almost surely mis-specified.[12]

Two final design decisions—that our study is incentivized and representative—are relatively straightforward. Much of the literature we build on comes from laboratory studies in economics, which are almost always incentivized. While there are good reasons to use non-incentivized measures, these are largely spurred by concerns about feasibility. These concerns were not substantial in our case, and we were motivated to move past them. In order to make our empirical basis representative of a broad range of people, rather than just specific subgroups, we used a representative sample. These two design decisions drove a number of implementation details discussed in the next section.

A challenge of using a representative sample is that some participants will be poorly educated. Elicitations must be designed with this in mind. As such, many of our behavioral measures are based on the same elicitation technique: indifferences elicited using a multiple price list (MPL) method. This was chosen as it allows for a more efficient estimation of indifference points than asking individual binary choices—which, for an experiment of our scope, would be infeasible—yet it is seen as easier for participants to understand than incentivized pricing tasks (Cason and Plott, 2014). A training period, with examples and supervised trial sessions, preceded the actual survey.[13] The techniques we used to (statistically) deal with

but most researchers would be grateful for a fuller demographic profile in a dataset, even if it excluded that particular measure.

[12]Additionally, recovering a parameter of a behavioral model through non-linear transformations of quantities measured with error is problematic. For example, some participants state relatively high, or low, certainty equivalents for lotteries that result in huge (positive and negative) CRRA coefficients. These values dominate correlations, causing them to be more informative about measurement error than behavior. Upon publication, our data will be made publicly available, and researchers interested in particular parametric formulations can use them to test those theories, if they so desire.

[13]We also included dominated options at the endpoints of the MPL scales wherever possible. The undominated options in these rows were pre-selected, following Andreoni and Sprenger (2012a). The software also

measurement error, detailed in Section 3.2, further helped in recovering valid estimates from these low-education populations.

2.1 Social Preferences

There are many examples where people's actions take into account the preferences and beliefs of others, even in non-strategic settings. The motivating factors behind these acts are often given the broad term of social preferences.

Altruism is defined as giving to strangers while expecting nothing in return. In experimental economics, it is usually measured using the dictator game (Forsythe et al., 1994; Falk et al., 2013), in which one participant decides unilaterally how to split money between themselves and another person. Following this literature, we measure **Altruism** as the amount given to another person in the dictator game.

Trust and reciprocity are intertwined. To understand why, imagine that a stranger asks for money for a sure-thing investment. In order to provide money in such a venture you must *trust* that he will give some of the proceeds back to you. The act of giving money back is *reciprocation*, which may depend on how much money you gave to the stranger. We measure these concepts through a standard trust game: one participant (the sender) decides how much of an endowment to send to a second (the receiver). This amount is doubled by the experimenter, and then the receiver decides how much to send back—which is also doubled (Berg et al., 1995). We measure **Trust** as the amount sent by the participant when they are in the role of the sender. The original sender will also take the role of receiver in a different interaction: **Reciprocity: Low** corresponds to the amount sent back when receiving the lowest amount, and **Reciprocity: High** corresponds to the amount sent back when receiving the maximum amount.[14]

imposed a single crossing point.

[14] These two types of reciprocity are sometimes referred to as positive and negative reciprocity. However, negative reciprocity is usually defined as the response to the lowest possible action. In this case, the lowest action is sending no money, to which the receiver cannot respond. Thus, to avoid confusion with the standard usage we use low and high rather than negative and positive. Note also that each participant's partner in

People are willing to punish others for what they perceive as bad behavior—even when that behavior does not directly affect them—and punishment is costly. To measure this, we allow participants to observe a trust game in which the sender gives all the money they have, and the receiver returns nothing. We give each participant a stock of points to punish the receiver—that is, to pay a cost to reduce the points the receiver gets: the amount used is **Pro-social Punishment**. Prior studies document that a significant minority of people—the percent possibly depending on culture—also punish the sender (Herrmann et al., 2008). Thus, we also give a separate stock of points that can be used to punish the person who sent all their money. The amount used to punish the sender is **Anti-social Punishment**.

Many people seem uncomfortable with having a different amount (greater or less) than others, a phenomenon known as *inequality aversion* (Fehr and Schmidt, 1999; Charness and Rabin, 2002; Kerschbamer, 2015). **Dislike Having Less** is how much a person is willing to forgo in order to ensure that they will not have less than another person. **Dislike Having More** is how much a person is willing to forgo in order to ensure that they will not have more than another person.

2.2 Measures of Risk Attitudes

To measure attitudes towards risk and uncertainty, we elicit the valuation of various prospects. All lotteries involve only two possible payoffs, and most assign 50% probability to each.

Following the standard approach, we identify the behavioral manifestation of risk aversion as valuing a lottery at less than its expected value. Extensive research shows that the patterns of valuation depend on whether a lottery contains positive payoffs, negative payoffs, or both a positive and a negative payoff.[15] Thus, we include three measures of risk aversion: **Risk Aversion: Gains** elicits a participant's certainty equivalent for a lottery containing non-

these interactions differs between when they are the sender and the receiver. Moreover, when the participant plays the role of receiver, we use the strategy method: that is, we elicit the response for every possible amount sent. For more details of implementation see Appendix A.

[15]The most familiar theoretical model of these patterns of valuation is Prospect Theory (Kahneman and Tversky, 1979).

negative payoffs, **Risk Aversion: Losses** elicits a participant's certainty equivalent for a lottery with non-positive payoffs, and **Risk Aversion: Gain/Loss** elicits a participant's certainty equivalent for a lottery with one positive and one negative payoff (Cohen et al., 1987; Holt and Laury, 2002). The difference between the expected value of the lottery and a participant's value are used in the analysis, so larger numbers indicate more risk aversion.

The endowment effect is the phenomenon that, on average, people value a good more highly if they possess, or are *endowed* with, it. In our implementation, **WTP** is the amount a participant is *willing to pay* for a lottery ticket, and **WTA** is the amount the participant is *willing to accept* for the same ticket when she or he is endowed with it. The difference between WTA and WTP is the **Endowment Effect** (Plott and Zeiler, 2005). We discuss this as a risk preference because the object being bought and sold is a lottery ticket, but more importantly, because of the patterns revealed in the analysis of Section 4.3.

Risk attitudes often change when one of the available options offers certainty, as demonstrated through the common ratio effect (Allais, 1953). Under Expected Utility, when the winning probabilities of two lotteries are scaled down by a *common* factor, a person's ranking over those lotteries should not change. However, this is often not the case. To capture this effect, we ask the participant to make two choices, one that measures risk aversion with a certain alternative, and another in which both options are risky. In **Risk Aversion: CR Certain** we elicit the amount b such that the participant is indifferent between a certain amount a and a lottery paying b with probability α (and zero otherwise): that is, a lottery equivalent of a sure amount. In **Risk Aversion: CR Lottery**, we elicit the amount c such that the participant is indifferent between a lottery paying a with probability $1/x$ (and zero otherwise), and c with probability α/x (and zero otherwise). Under Expected Utility, $b = c$. The **Common Ratio** measure is then $b - c$ (Dean and Ortoleva, 2015).

Ambiguity aversion is a preference (or beliefs that lead to a preference) for prospects with known probabilities over those with unknown probabilities. To measure it, we use an *ambiguous urn* filled with balls of two different colors. One color gives the participant a positive payoff, and the other gives them zero. Participants do not know the proportion of the different color balls in the urn, but are allowed to choose which color gives a positive payoff. They are then asked for their certainty equivalent for a draw from this urn. If participants have a prior over the composition of the urn, they must believe a draw from the urn has a winning probability of at least 50%, yet many participants prefer a 50/50 lottery with known odds. The difference between the certainty equivalent for a draw from the *risky urn*—with a known composition of 50% of each color—and the ambigous urn is **Ambiguity Aversion**.[16] Similarly, a draw from a risky urn is usually more highly valued than one in which the number of balls is unknown, but drawn from a uniform distribution—that is, a compound lottery. The difference between the certainty equivalents for a draw from the risky urn and one from a compound urn is **Compound Lottery Aversion** (Halevy, 2007).

2.3 Overconfidence

Overconfidence can be divided into three types. **Overestimation** refers to a person's estimate of her performance on a task (versus her actual performance). **Overplacement** refers to her perceived performance relative to other participants (versus her real relative performance). In order to measure these phenomena, we ask participants to complete two tasks, and then ask them to estimate their performance on one, and their performance relative to others taking the survey on the other. The difference between these subjective estimations and actual performance (in absolute or percentile terms) give us overestimation and overplacement, respectively (Moore and Healy, 2008).

Overprecision refers to a belief that one's information is more precise than it actually is. We ask participants to estimate a quantity (such as the year the telephone was invented), and then tell us how close they think they were to the correct answer. To difference out

[16] The risk measure here is **Risk Aversion: Urn**, which matches the description above. Empirically, this measure is highly correlated with Risk Aversion: Gains. Note that both Compound Lottery Aversion and Ambiguity Aversion difference out the same quantity. If this quantity is measured with error, this can create spurious correlation between Compound Lottery Aversion and Ambiguity Aversion. This issue, and our solution, is discussed in Section 3.2.

overprecision from justified precision, we regress how close the participant thought they were on a fourth order polynomial of their accuracy (Ortoleva and Snowberg, 2015b,a).[17]

2.4 Patience (Time Preferences)

A payoff sometime in the future is generally seen as less valuable than a payoff of the same size today. The value today of a fixed future payoff is **Patience** (Andersen et al., 2008).[18]

3 Implementation and Analysis

This section describes our representative, incentivized survey, and the statistical techniques used to eliminate the attenuating effects of measurement error.

3.1 Survey Implementation

Administering an incentivized survey to a representative population presents challenges not normally dealt with in lab environments. To surmount these challenges, we partnered with YouGov, a worldwide leader in online surveys serving the public, businesses, and governments.[19] Our study was given to a representative sample of 1,000 U.S. adults between March 30 and April 14, 2016. We consulted extensively with YouGov on our study design to utilize its expertise in survey design and implementation.

Constructing a representative sample is difficult given variation in response rates. In order to do so, most modern surveys weight on demographics. YouGov supplements this with

[17] Note that the fact that both overprecision and overplacement refer to the same factual questions will create correlated measurement error between them. This can create spurious correlation between the two measures. This issue, and our solution, are discussed in Section 3.2.

[18] In an earlier survey we attempted to measure present bias, but found little or no evidence of it, similar to some recent studies of financial payments (see, for example, Augenblick et al., 2015). This may have occurred due to the fact that points were, in general, not instantly convertible into consumption. Thus, we did not attempt to measure it here. A general concern about such experiments is whether the participants trust the experimenter to follow through with payment (Andreoni and Sprenger, 2012a). Using a survey company with established relationships with its panelists seems to have largely mitigated this, as discount rates were quite low (δ was quite high).

[19] Companies that use in-person or phone surveys, such as Gallup, were unable to administer incentives.

its own panel of participants. It continually recruits new people, especially from difficult-to-reach and low-socio-economic-status groups. To generate a representative sample, it randomly draws people from various Census Bureau products, and matches them on observables to members of its panel. Oversampling and differential response rates lead to the over- and under-representation of certain populations. Thus, YouGov provides sample weights to recover estimates that would be obtained from a fully representative sample. According to Pew Research, YouGov's sampling and weighting procedure yields more representative samples than traditional probability sampling methods, including Pew's own probability sample (Pew Research Center, 2016, YouGov is Sample I). We use these weights throughout the paper.[20]

Incentivized questions pose additional challenges: stakes, and whether the experimenter is seen as credible in making future payments or running randomizations as specified. Two randomly selected questions were chosen for payment.[21] To enhance the credibility of our study, we took advantage of YouGov's relationship with its panel, and restricted the sample to those who had already been paid (in cash or prizes) for their participation in surveys.

All outcomes to incentivized questions were expressed in points. This is an internal YouGov currency used to pay participants. It can be converted to U.S. dollars, or prizes, using the approximate rate of \$0.001 per point.[22] The average payment to participants was

[20]As economists rarely run their own surveys in representative populations, it is worth explaining how the survey research literature uses the term "representative." With few exceptions—censuses, and samples in rural areas of developing countries based on a census—representative samples are representative on observables, not on unobservables. While random samples have the potential to be representative on both observable and non-observables, low response rates render these samples less representative on both observables and expressed preferences, as the Pew study documents. Commonly used representative surveys in economics, such as the Current Population Survey, use weighting to account for non-response. The CPS also uses imputation to adjust for item non-response, which is not present in our survey (see www.census.gov/programs-surveys/cps/technical-documentation/methodology/imputation-of-unreported-data-items.html).

[21]We chose to pay for two randomly selected questions to increase the stakes while making fewer participants upset about their payoffs. Paying for two questions instead of one may theoretically induce some wealth effects, but these are known to be negligible, especially in an experiment such as ours (Charness et al., 2016). Paying for randomly selected questions is incentive compatible under Expected Utility, but not necessarily under more general risk preferences, where it is known that no such mechanism may exist (Karni and Safra, 1987; Azrieli et al., Forthcoming). A growing literature suggests this theoretical concern may not be empirically important (Beattie and Loomes, 1997; Cubitt et al., 1998; Hey and Lee, 2005; Kurata et al., 2009), but there are some exceptions (Freeman et al., 2015).

[22]The conversion from points to awards can only be done at specific point values, which leads to a slightly

around 9,000 points (or \$9). The survey took participants between 45 minutes and an hour. This compensation level is quite high for an internet survey, and represents a rate of pay approximately three times the average for surveys.

3.2 Measurement Error

Measurement error causes a downward bias in correlations. This would make it more difficult to pick out the clusters of inter-related measures. Moreover, this downward bias would also make it difficult to know which behaviors are actually *not* related. To circumvent this issue, we use the ORIV technique of Gillen et al. (2018), which uses the instrumental variables approach to errors-in-variables to produce efficient, consistent estimates. At the heart of this method is duplicate observations of variables—multiple elicitations that are similar, but not exactly the same—that are likely to have orthogonal measurement error. This technique takes duplicate measures and estimates a stacked regression where each X^i is used as both an independent variable and an instrument. This is done twice, once for each Y^i, effectively averaging across all four specifications in the stacked regression model.[23]

Our setting requires us to deal with an additional issue: constructed variables will often have correlated measurement error due to the nature of their construction. For example, the Compound Lottery Aversion and Ambiguity Aversion measures are both constructed by taking some behavior (the certainty equivalent for a compound or ambiguous urn), and subtracting off the same quantity measured with error (certainty equivalent of a simple risky urn). This leads to correlation in the measurement error of Ambiguity and Compound Lottery Aversion. To avoid spurious correlations, we make use of the fact that we have two observations for each measure, and modify the ORIV procedure so that the measurement error in the instrument is uncorrelated with the measurement error in the left-hand-side

convex payoff schedule. This is of little concern here as these cash-out amounts are further apart than the maximum payoff from this survey.

[23]When we do not have an available duplicate for one of the measures, we use an approximation of ORIV detailed in Footnote 31 of Gillen et al. (2018).

variable.[24] We use this formulation for all sets of variables constructed from two elicitations: Compound Lottery and Ambiguity Aversion, and Overplacement and Overprecision.

3.3 Multiple Hypothesis Testing

This paper displays a large number of correlations and standard errors. There are no theoretical predictions for most of the correlations we examine. Thus, we omit any description of statistical testing or significance from our tables. However, if one were interested in null-hypothesis statistical testing (NHST), the appropriate critical value for significance at the 5% level is between 1.96 for a single hypothesis test, and 3.82 using a Bonferroni (Dunn, 1958, 1961) correction for all 378 correlations underlying Sections 4 or 5.

3.4 Principal Component Analysis

We examine correlation matrices directly, and use principal component analysis (PCA) to summarize them. The aim of PCA is to extract the m *components* most useful for explaining $n > m$ variables. Components are linear combinations of the variables. The first component is constructed to capture the highest possible fraction of variance in the data (subject to the constraint that the linear weights sum to one), the second to capture the highest fraction of the remaining variance, conditional on being orthogonal to the first component, and so on.[25]

Once components are identified, the key question is, "How many are necessary to provide a good description of the underlying data?" Heuristically, we want to retain components only when the marginal explanatory power is high. In order to determine the number of

[24]Formally, X^i and Y^i, $i \in \{1, 2\}$ are two elicitations of X^* and Y^* measured with error. For a given i, X^i and Y^i are constructed by subtracting the same quantity (measured with error) from two different quantities. Thus, the measurement error in X^i and Y^i are correlated for a given i, but the measurement error in X^i and Y^j, $i \neq j$ are not. The stacked regression in ORIV is then modified to become:

$$\begin{pmatrix} Y^a \\ Y^b \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \beta \begin{pmatrix} X^a \\ X^b \end{pmatrix} + \eta, \quad \text{with instruments } \; W = \begin{pmatrix} X^b & 0_N \\ 0_N & X^a \end{pmatrix}.$$

[25]See Abdi and Williams (2010) for an introduction. We also tried more sophisticated machine learning techniques, such as clustering, but these produced no additional insight. See Appendix B.

components to retain we use an approach which captures this intuition: parallel analysis. Parallel analysis creates many random datasets with the same numbers of observations and variables as the original data. The average eigenvalues of the resulting correlation matrices are then computed. Components are kept as long as their associated eigenvalues are greater in the actual data than the average in the randomly generated data.[26]

The retained components help to understand the relationship between the original variables in the dataset. The correlation between a component and a variable is called the variable's *loading* on that component. Variables that load heavily on the same component are highly related. In order to facilitate interpretation, retained components can be rotated relative to the data. Following standard practice, we rotate the resulting components using the Varimax rotation (Furr, 2017, p. 92). This rotates the basis identified from the retained components to maximize the variance of the squared loadings, easing interpretation of the components. However, as we will see, the patterns in the components largely line up with apparent patterns in the correlation matrices, making interpretation straight-forward.

In Section 5, we analyze whether the components we identify correlate with cognitive and demographic measures. To reduce measurement error in the components, we multiply the weights by the average of the two copies of each variable. This leads to components that are not orthogonal, although, in practice, they are quite close.[27]

4 Relationships between Econographics

The next two sections attempt to explain, as succinctly as possible, the relationships between 21 econographics variables. We begin by examining each econographic separately. Next, we study the relationships between econographics through a visual inspection of correlation ma-

[26] The value of additional components is graphically represented by a "scree" plot. This shows the eigenvalues of the correlation matrix of the underlying data, and compares this with the average value of the eigenvalues produced by parallel analysis. The scree plots for our analysis can be found in Appendix C.

[27] We could also construct two measures of each component, one from each set of duplicate variables. These could then be used as instruments for each other. However, as the two resulting measures of each component are correlated between 95%–99%, this results in only slightly greater coefficients, with standard errors that are roughly twice as large due to the presence of a first stage. As such, we prefer averaging.

trices, followed by principal components analyses to verify the observed patterns. This leads to our central finding that the 21 econographics are well summarized by six principal components. In the next section, we examine the relationship between these principal components and cognitive abilities and demographics. As noted in the introduction, there are many ways one might summarize these 378 relationships. Our approach is driven by the desire to create an empirical basis for the underlying structure of more comprehensive theories of behavioral decision-making, and by the hope that this is a relatively straight-forward way to do so.

We describe the relationships between econographics in three steps. First, we examine social preferences, then risk preferences, and then combine social and risk preferences together with overconfidence measures and patience. This is done for simplicity, and because each of the first two types of preferences may be of independent interest. Moreover, as our results are largely driven by clusters of correlations, these clusters will not disappear when additional measures are added to the analysis (although they may be augmented).

4.1 Summary Statistics

Most of the behaviors we measure appear to be quite similar to those found in the literature. For all but one of the behaviors we elicit, a strict majority of participants exhibit behavior consistent with the literature—for example, 71% of participants appear to be ambiguity averse or ambiguity neutral.

The summary statistics in Table 2 show that behavior in our data is consistent with standard findings in the laboratory. In addition to summary statistics, we also show the percent of participants whose responses are in the "expected" direction. Surveying the information in this column, the majority of participants are risk averse or risk neutral over gains, risk loving over losses, inequality averse, exhibit an endowment effect and the common ratio effect, are ambiguity averse, have a negative reaction to compound lotteries, and are over-confident. (A majority of participants are not all of these things simultaneously.) Note that although patience in this table is represented as a discount rate, in the correlational

Table 2: Summary Statistics of Econographics Measures

Variable	Description / Unit	Mean Value	Standard Deviation	% Expected Direction	Corr. btw. Duplicates
Reciprocity: Low	% of Possible Points Returned	0.42	0.23		0.80
Reciprocity: High	% of Possible Points Returned	0.39	0.22		0.96
Trust	% of Possible Points Sent	0.45	0.26		
Altruism	% of Possible Points Sent	0.41	0.27		
Anti-social Punishment	% of Possible Points Used	0.21	0.35		
Pro-social Punishment	% of Possible Points Used	0.50	0.42		
Dislike Having More	% of Income Forgone for Equal Split	0.08	0.45	78%	0.74
Dislike Having Less	% of Income Forgone for Equal Split	0.02	0.37	58%	0.69
Risk Aversion: Gains	(EV − CE)/EV	−0.04	0.49	50%	0.64
Risk Aversion: Losses	(EV − CE)/EV	−0.29	0.52	73%	0.69
Risk Aversion: Gain/Loss	(EV − CE)/EV	−0.04	0.55		0.71
WTA	% of Expected Value	0.91	0.41		0.70
WTP	% of Expected Value	0.65	0.36		0.75
Endowment Effect	% of Expected Value	0.26	0.57	74%	0.75
Risk Aversion: CR Certain	1 − EV of LE as % of Sure Amount	−0.32	0.44	71%	0.76
Risk Aversion: CR Lottery	1 − EV of LE as % of EV of Lottery	−0.35	0.44	73%	0.70
Common Ratio	(EV − CE)/EV	−0.03	0.52	62%	0.62
Ambiguity Aversion	Ambiguous CE − Risky CE	0.07	0.55	71%	0.64
Compound Aversion	Compound CE − Risky CE	0.03	0.56	68%	0.57
Overestimation	Perceived − Real # Correct (of 3)	0.69	1.18	87%	0.37
Overplacement	Perceived − Real Percentile	6.4	39	60%	0.30
Overprecision	Standardized Subjective Precision	−0.02	1.00		0.42
Patience	Monthly Discount Rate	0.87	0.23		0.78

Notes: EV = Expected Value, CE = Certainty Equivalent, LE = Lottery Equivalent. % Expected Direction refers to the percent of the participants (weighted) that give an answer in the direction expected given the current literature: risk averse or risk neutral for most risk questions; risk loving or risk neutral for risk aversion over losses; equality seeking for distributional preferences; endowment effect greater than 0; overweighting small probabilities for common ratio; and ambiguity / compound averse or neutral. When there are two measures of a quantity, those measures are normalized and stacked, so the sample statistics are drawn from 2,000 observations from 1,000 people.

analysis we code the variable as discussed in Section 2.4. Either coding gives the same (directional) interpretation to correlations. However, the coding in Section 2.4 is linear in a participant's answer, allowing for the measurement error correction discussed in Section 3.2.

Our data exhibit fairly standard levels of noise. As discussed in Gillen et al. (2018) and Snowberg and Yariv (2018) the correlation of duplicate measures—subtracted from one—gives the level of noise in a particular elicitation. Gillen et al. (2018) report correlations of around 0.65 between duplicate measures, using data from Caltech undergraduates. In most cases the correlations we observe—in the final column of Table 2—are somewhat higher. This implies that our data is *less* noisy than similar data obtained from Caltech undergraduates. The exceptions are the overconfidence measures, which are noisier than the rest.[28] When there is no correlation listed we only have one elicitation of that behavior.

4.2 Links Between Social Preferences

There is ample opportunity to create a more parsimonious representation of social preferences: altruism, trust, anti- and pro-social punishment, and distributional preferences. These measures fall into three clusters shown in correlation Table 3: one formed by the two measures of reciprocity, and altruism and trust; a second formed by the willingness to punish pro- and anti-social behavior; and a third formed by our two measures of inequality aversion. These clusters are characterized by high within-cluster correlations and low correlations between measures in different clusters. To make these clusters visually apparent, we present the correlation matrix in the form of a "heat map" where the shade of red indicates the magnitude of the correlation.[29]

The first correlation in the first cluster (0.86 between the two reciprocity measures), is a useful example for interpretation. There are distinctions between those who are more and less reciprocal when a partner is more or less generous, resulting in a less than perfect cor-

[28] Gillen et al. (2018) do not report data from overconfidence measures.

[29] The shading is driven by a concave function of magnitude so that there is more differentiation between magnitudes of 0 and 0.25 than there is between 0.25 and 1. We also show both the upper and lower part of this symmetric matrix.

Table 3: ORIV Correlations of Social Measures

	Reciprocity: Low	Reciprocity: High	Altruism	Trust	Anti-social Punishment	Pro-social Punishment	Dislike Having More	Dislike Having Less
Reciprocity: Low		0.86 (.027)	0.34 (.063)	0.49 (.052)	−0.03 (.062)	0.06 (.058)	0.16 (.053)	−0.07 (.054)
Reciprocity: High	0.86 (.027)		0.34 (.053)	0.49 (.041)	−0.03 (.053)	0.08 (.050)	0.22 (.048)	−0.05 (.054)
Altruism	0.34 (.063)	0.34 (.053)		0.60 (.040)	0.05 (.049)	0.03 (.047)	0.25 (.057)	−0.04 (.061)
Trust	0.49 (.052)	0.49 (.041)	0.60 (.040)		−0.05 (.041)	0.07 (.045)	0.22 (.054)	−0.20 (.059)
Anti-social Punishment	−0.03 (.062)	−0.03 (.053)	0.05 (.049)	−0.05 (.041)		0.43 (.034)	−0.20 (.061)	0.22 (.059)
Pro-social Punishment	0.06 (.058)	0.08 (.050)	0.03 (.048)	0.07 (.045)	0.43 (.033)		−0.04 (.061)	0.04 (.061)
Dislike Having More	0.16 (.053)	0.22 (.047)	0.25 (.057)	0.22 (.054)	−0.20 (.061)	−0.04 (.061)		0.26 (.059)
Dislike Having Less	−0.07 (.054)	−0.05 (.054)	−0.04 (.061)	−0.20 (.059)	0.22 (.059)	0.04 (.061)	0.26 (.059)	

Notes: Bootstrapped standard errors from 10,000 simulations in parentheses. Colors in heatmap change with each 0.05 of magnitude of correlation.

relation. Yet, the predominant behavioral distinction, reflected in the very high correlation, is how reciprocal someone is in all conditions. This overarching behavior is (empirically) related to both trust and altruism, although more closely to the former. As noted in the introduction, although some readers may have anticipated some of these correlations, the fact that the literature distinguishes between, say, different forms of reciprocity, indicates that these anticipations are not universally shared.

The second cluster contains the two punishment measures.[30] Like other clusters, they are highly correlated with each other, but with few other measures. However, these measures, unlike any others in our data, are characterized by both an extensive and intensive margin. The extensive margin is whether or not someone punishes, and the intensive how much punishment a person metes out. This leads to obvious questions about how these two margins contribute to the overall relationship. The correlations on both margins are roughly equal. About 2/3 of participants engage in pro-social punishment, whereas only 1/3 engage in anti-social punishment. However, almost everyone who engages in anti-social punishment also engages in pro-social punishment. Of those that engage in both types of punishment, there is a strong positive correlation of about 0.4 in the amount they choose to spend on punishment of both parties. Both the extensive and intensive margin are poorly related to other measures of social preferences.

The third cluster contains the distributional preferences measures. It features the weakest intra-cluster correlation. Moreover, both measures in this cluster—Dislike Having More and Dislike Having Less—are moderately correlated with other econographics. These moderate correlations extend to some measures in the risk domain, leading to this component combining with risk preferences in the analysis of all 21 econographics in Section 4.4.

Before turning to the principal components analysis (PCA), it is worth noting that there is very little in the literature that indicates which specific correlations we should, and should

[30]As a reminder, both refer to costly punishments meted out on participants in a sender-receiver game where the sender sent his or her entire endowment, and the receiver returned none. Pro-social punishment refers to the amount used to punish the receiver, and anti-social punishment refers to the amount used to punish the sender.

Table 4: Principal Components Analysis of Social Preferences

	Generosity	*Punishment*	*Inequality Aversion*	Unexplained
Reciprocity: Low	**0.52**	0.04	−0.04	0.27
Reciprocity: High	**0.52**	0.03	0.00	0.27
Altruism	**0.40**	0.02	0.11	0.54
Trust	**0.49**	−0.02	−0.07	0.35
Anti-social Punishment	−0.03	**0.70**	0.05	0.24
Pro-social Punishment	0.07	**0.64**	−0.03	0.37
Dislike Having More	0.17	−0.24	**0.65**	0.28
Dislike Having Less	−0.13	0.17	**0.75**	0.22
Percent of Variation	34%	19%	16%	32%

Notes: First three principal components using the varimax rotation. Weights greater than or equal to 0.25 in bold.

not, find. This could lead to a lot of plausible storytelling. For example, one might believe that the dislike of having more (one form of inequality aversion) entirely motivates altruism, and a dislike of having less (another form) motivates punishment. Yet, these patterns are not present in our data. Alternatively, one might expect that the predominant feature of distributional preferences is the presence or absence of a preference for equality, which would generate the observed correlation between Dislike of Having More and Dislike of Having Less. Whatever one's priors (or lack thereof), these results should be informative. We discuss what we believe are the biggest takeaways for theory in Section 6.

The PCA of these correlations shows the same decomposition: results appear in Table 4. Three components are suggested for inclusion under parallel analysis—see Appendix Figure C.1 for the scree plot. Together, these three components explain 68% of the variation in the eight measures of social preferences we explore here.

These three components have fairly obvious interpretations. The first is generosity in behaviors that directly influence the well-being of another person. The second is a general affinity for punishment. The third and final component captures both types of inequality aversion. The components are thus named accordingly: *Generosity*, *Punishment*, and

Inequality Aversion.

Overall, these results show there is a clear structure to the connections between different social preferences. These measures tend to group in clusters with high within-cluster correlations and low across-cluster correlation. This structure likely leads to similar findings using other latent-variable discovery techniques.

4.3 Links Between Risk Attitudes

In this subsection we show that models of decision-making under risk and uncertainty may be too parsimonious. As shown by the clear clusters in Table 5, ambiguity aversion and compound lottery aversion group together, blurring the line between risk and uncertainty. Moreover, we find two clusters of risk attitudes related to WTA and WTP, respectively.

The easiest cluster to interpret contains Ambiguity and Compound Lottery Aversion.[31] This high correlation is consistent with extant empirical work (Halevy, 2007; Dean and Ortoleva, 2015; Gillen et al., 2018) and theoretical observation (Segal, 1990; Dean and Ortoleva, 2015). Note that both of these measures are essentially unrelated to measures of risk preferences, suggesting either a delineation between risk and uncertainty, with compound lotteries grouped with ambiguity aversion, or the lack of a clean line between risk and ambiguity.

The two remaining clusters all contain risk attitudes, and suggest two separate aspects of risk preferences. The first contains Risk Aversion: Gain, Loss, Gain/Loss, and WTA; and the second includes Risk Aversion: CR Certainty, CR Lottery, and WTP. Crucially, there are low correlations between the measures in different clusters (with the exception of the relationship between Risk Aversion: Gain/Loss and WTP—correlated 0.3, meaning that risk aversion using WTP is negatively correlated with risk aversion over losses). This suggests that there are two separate aspects of risk aversion that are largely unrelated to each other, or to preferences under uncertainty. Note that these two different aspects are both

[31] Note that these come from the certainty equivalent of a draw from an ambiguous (Ellsberg) or compound urn minus the certainty equivalent from a 50/50 risky urn. In order to prevent spurious correlation, we adapt the ORIV procedure as described in Section 3.2.

Table 5: ORIV Correlations of Risk Measures

	WTA	Risk Aversion: Gains	Risk Aversion: Losses	Risk Aversion: Gain/Loss	WTP	Risk Aversion: CR Certain	Risk Aversion: CR Lottery	Ambiguity Aversion	Compound Lottery Aversion	Endowment Effect	Common Ratio
WTA		−0.66 (.051)	−0.27 (.076)	−0.58 (.054)	−0.11 (.069)	−0.03 (.069)	0.12 (.071)	0.14 (.064)	0.09 (.067)	–	−0.14 (.070)
Risk Aversion: Gains	−0.66 (.051)		0.39 (.070)	0.60 (.049)	0.04 (.071)	0.09 (.068)	−0.13 (.070)	0.03 (.057)	0.01 (.061)	−0.49 (.055)	0.20 (.055)
Risk Aversion Losses	−0.27 (.077)	0.39 (.070)		0.65 (.056)	0.30 (.076)	−0.19 (.074)	−0.15 (.077)	0.06 (.065)	0.01 (.065)	−0.38 (.060)	−0.04 (.064)
Risk Aversion: Gain/Loss	−0.58 (.054)	0.60 (.049)	0.65 (.055)		0.19 (.069)	−0.14 (.070)	−0.21 (.072)	−0.02 (.058)	−0.02 (.072)	−0.53 (.047)	0.06 (.069)
WTP	−0.11 (.069)	0.04 (.070)	0.30 (.076)	0.19 (.069)		−0.45 (.041)	−0.28 (.061)	0.01 (.054)	0.00 (.059)	–	−0.17 (.060)
Risk Aversion: CR Certain	−0.03 (.070)	0.09 (.069)	−0.19 (.074)	−0.14 (.069)	−0.45 (.041)		0.41 (.060)	−0.11 (.063)	−0.05 (.068)	0.26 (.053)	–
Risk Aversion: CR Lottery	0.12 (.071)	−0.13 (.071)	−0.15 (.078)	−0.21 (.072)	−0.28 (.060)	0.41 (.060)		0.01 (.059)	0.06 (.069)	0.26 (.055)	–
Ambiguity Aversion	0.14 (.063)	0.03 (.057)	0.06 (.065)	−0.02 (.058)	0.01 (.055)	−0.11 (.063)	0.01 (.058)		0.47 (.080)	0.09 (.055)	−0.11 (.073)
Compound Aversion	0.09 (.068)	0.01 (.060)	0.01 (.064)	−0.02 (.072)	0.00 (.058)	−0.05 (.068)	0.06 (.068)	0.47 (.066)		0.06 (.057)	−0.10 (.083)

Notes: Bootstrapped standard errors from 10,000 simulations in parentheses. Colors in heatmap change with each 0.05 of magnitude of correlation.

within a fairly narrow domain: the valuation of risky lotteries. Thus, unlike psychology research, which has shown differences in risk preferences across outcome domains, these patterns suggest differences in preferences within the domain of money lotteries—the most studied domain in choice theory, and experimental and behavioral economics.

These two separate aspects align with WTA and WTP, respectively, giving this division substantive and theoretical importance. We find that certainty equivalents fall into one cluster while lottery equivalents into another. This is in line with prior research finding that different fixed elements in an MPL lead to different degrees of risk aversion. Here, as in the literature, this is consistent with the fixed element of the MPL acting as a reference point (Sprenger, 2015). Our results add, first, that these groups are largely uncorrelated with each other, and, second, naturally align with WTA and WTP. WTA—where the lottery is explicitly the reference point—naturally aligns with certainty equivalents: measures where the lottery is fixed. Similarly, WTP aligns with measures with a fixed monetary amount. At the same time, as we discuss in Section 6, and more in depth in Chapman et al. (2017), our data are difficult to reconcile with current theoretical explanations. The most visible signature of this difficulty in Table 5 is the lack of an obvious component of loss aversion positively related to WTA and negatively related to Risk Aversion: Gain/Loss.

Before proceeding, it is worth discussing the two measures in the correlational table that we do not include in the clusters or in the following principal components analysis: the endowment effect and the common ratio effect. Both measures are the difference of two other measures we already include: for example, the endowment effect is the difference of the WTA and WTP. Thus, including all three measures in a PCA makes little sense. As described above, our own recent research compiles extensive evidence that WTA and WTP should be treated as separate behavioral measures. Thus, we include the component measures (WTA and WTP) rather than the endowment effect (Chapman et al., 2017). We take the same approach with the common ratio measures: entering the constituent measures

Table 6: Principal Components Analysis of Risk Preferences

	Risk Aversion: WTA	*Risk Aversion: WTP*	*Uncertainty*	Unexplained
WTA	**−0.51**	−0.09	0.12	0.32
Risk Aversion: Gains	**0.55**	0.13	0.05	0.25
Risk Aversion: Losses	**0.39**	−0.21	0.09	0.45
Risk Aversion: Gain/Loss	**0.53**	−0.09	0.02	0.22
WTP	0.03	**−0.54**	−0.02	0.43
Risk Aversion: CR Certain	0.08	**0.62**	−0.06	0.29
Risk Aversion: CR Lottery	−0.05	**0.48**	0.11	0.53
Ambiguity Aversion	−0.01	−0.04	**0.69**	0.27
Compound Lottery Aversion	0.01	0.04	**0.69**	0.28
Percent of Variation	29%	21%	17%	34%

Notes: First three principal components using the varimax rotation. Weights greater than or equal to 0.25 in bold.

into the correlational and principal components analysis.[32]

Once again, the principal components analysis in Table 6 confirms the visual patterns in the correlation table. The first three components explain 66% of the variation in the nine measures of risk preferences considered here. The first and the second clearly capture different aspects of risk attitudes. Following the discussion above, we use the names *Risk Aversion: WTA* and *Risk Aversion: WTP* to describe them. The third encompasses more complex lotteries, which may induce uncertainty: hence we use the title *Uncertainty*.

4.4 Putting it all Together

Analyzing all 21 econographics together leads to six components. The structure of components from the risk and social domains are preserved, however, a social and risk component combine. The sixth component is comprised of the three overconfidence measures. Time preferences load on several components; most heavily on the Punishment component.

[32]Entering the common ratio measure directly in the principal component analysis produces qualitatively similar results to Table 6: see Table D.2.

Analyzing risk and social preferences, together with overconfidence and time preferences, is straight-forward because there are few important relationships between risk and social preferences. However, they are not completely unrelated. One of the social preference components (Inequality Aversion) combines with a risk-preference component (Risk Aversion:WTP). Otherwise, the structure in the previous two subsections is largely unaltered.

The first and second components in Table 7 are similar to the first components in the social and risk analyses, respectively. Thus, we retain the same names. Patience loads moderately on both components: more generous and less risk averse people are more patient.

The third component is of particular interest. It combines the measures included in the second risk component, Risk Aversion: WTP, with those of the third social component, Inequality Aversion. This suggests that some aspects of social and risk preferences are distinct, while others are more related. Moreover, this relationship follows a clear pattern building upon the previously identified components: Inequality Aversion combines with Risk Aversion: WTP. We note that both of these components deal with aversions to spreads—in possible payoffs or distributional assignments. Thus, we conjecture that a similar form of caution may lead participants to both dislike entering conditions of risk, as well as generically dislike unequal allocations. Note, however, that this pertains only to entering a situation of risk (WTP for a lottery), rather than leaving one (WTA). In line with its constituent parts, we call this component *Inequality Aversion/WTP*.

The fourth component loads heavily on all three overconfidence measures; accordingly, we call it *Overconfidence*. Given existing work on the conceptual distinctions between different types of overconfidence, the fact that they would load on a single component may not have been ex ante obvious (Moore and Healy, 2008).

The fifth is the social component Punishment, to which (a bit of) time preferences have been added. People who score highly on this component enjoy both pro- and anti-social punishment, and are impatient. Thus, we dub this new component, *Impulsivity*.

The sixth component is essentially the same as the third risk component. Thus, we

Table 7: Principal Components Analysis of All Measures

	Generosity	*Risk Aversion: WTA*	*Inequality Aversion/ WTP*	*Overconfidence*	*Impulsivity*	*Uncertainty*	Unexplained
Reciprocity: Low	**0.50**	0.00	0.00	0.02	0.05	−0.01	0.29
Reciprocity: High	**0.51**	0.01	−0.02	−0.07	0.06	0.04	0.27
Altruism	**0.39**	0.06	0.03	0.04	0.05	−0.03	0.54
Trust	**0.47**	0.01	−0.06	0.12	−0.04	−0.05	0.33
Anti-social Punishment	−0.01	0.00	0.03	0.05	**0.67**	−0.03	0.26
Pro-social Punishment	0.09	−0.05	−0.04	−0.08	**0.60**	0.03	0.42
Dislike Having More	0.22	0.00	**0.26**	−0.18	−0.11	0.17	0.57
Dislike Having Less	−0.06	−0.10	**0.40**	−0.02	0.22	0.07	0.48
WTA	−0.04	**−0.50**	−0.06	0.06	0.04	0.10	0.31
Risk Aversion: Gains	0.04	**0.55**	0.11	0.03	−0.05	0.06	0.23
Risk Aversion: Losses	−0.09	**0.36**	−0.16	0.05	0.16	0.09	0.42
Risk Aversion: Gain/Loss	−0.01	**0.50**	−0.11	−0.04	−0.01	0.03	0.24
WTP	−0.01	0.01	**−0.47**	0.14	0.01	−0.02	0.45
Risk Aversion: CR Certain	−0.02	0.10	**0.53**	0.07	−0.02	−0.11	0.38
Risk Aversion: CR Lottery	−0.03	−0.03	**0.42**	0.11	−0.03	0.04	0.56
Ambiguity Aversion	0.04	0.00	−0.03	−0.03	0.05	**0.70**	0.25
Compound Lottery Aversion	−0.07	0.01	0.02	0.08	−0.07	**0.65**	0.30
Overestimation	0.04	−0.03	0.01	**0.50**	0.10	0.01	0.54
Overplacement	0.03	0.04	0.16	**0.49**	−0.09	−0.05	0.56
Overprecision	−0.01	−0.02	−0.07	**0.62**	−0.02	0.06	0.33
Patience	0.19	−0.19	−0.05	−0.05	**−0.28**	0.05	0.65
Percent of Variation	14%	13%	11%	8%	8%	7%	40%

Notes: First six principal components using the varimax rotation. Weights greater than or equal to 0.25 in bold.

continue to call this component *Uncertainty*.

Overall the underlying structure can be summarized with four points: 1) Six easy-to-interpret and separate components explain 60% of the variance of our twenty-one variables; 2) The social and risk components do not change, but two of them—Risk Aversion: WTP and Inequality Aversion—collapse into one; 3) Time preferences are captured by several of the existing components, possibly signaling that they are related to many elements, although none particularly strongly; and 4) Overconfidence measures are largely separate, and captured by a single component.

5 Econographics and Other Measures

In this section we examine the correlation between our econographic components, cognitive abilities, and demographics. This has two potential benefits. First, it may give us clues as to the the psychological processes underlying these clusters. Second, it may also provide useful information about economic preferences. For example, understanding how the components relate to demographics might be useful in understanding how preferences might change as a population becomes older or better educated.

5.1 Econographics and Cognitive Measures

There are several important relationships between our six econographic components and the two measures of cognition on our survey: IQ and the three-question Cognitive Reflection Test (CRT; Frederick, 2005). These two cognitive ability measures measures have an ORIV correlation of > 0.8.[33]

Our IQ measure consists of two classes of questions drawn from the International Cog-

[33]The correlation uncorrected for measurement error is 0.46. Our survey also contained two instantiations of the "beauty contest" measure of strategic sophistication (Nagel, 1995). Unfortunately, as in Snowberg and Yariv (2018), we found that the distribution of responses is relatively uniform, with a spike in responses at 50 (out of 100), as shown in Figure D.1. It is difficult to discern who is strategically sophisticated from this pattern. Thus, we did not include these measures in our analysis.

Table 8: Correlations with Measures of Cognitive Ability

	IQ	CRT
Generosity	0.13 (.038)	0.09 (.038)
Risk Aversion: WTA	0.02 (.040)	0.04 (.039)
Inequality Aversion/WTP	−0.13 (.043)	−0.11 (.039)
Overconfidence	−0.23 (.050)	−0.12 (.049)
Impulsivity	−0.24 (.033)	−0.25 (.032)
Uncertainty	0.05 (.034)	0.02 (.037)

Notes: Bootstrapped standard errors from 10,000 simulations in parentheses. Colors in heatmap change with each 0.05 of magnitude of correlation. Blue hues indicate negative correlations, green hues positive correlations.

nitive Ability Resource, a public domain intelligence measure (ICAR; Condon and Revelle, 2014). We choose three matrix reasoning questions, similar to Raven's Progressive Matrices. In these questions, participants determine which of a set of possibilities correctly completed a graphic pattern. We also construct a second battery of three questions based on 3-D rotations: a drawing of a cube was shown, and participants had to identify which of a set of six other drawings of a cube were compatible. These questions were chosen to try to capture a variation in fluid intelligence in the general population.The specific questions were chosen to be of progressively greater difficulty.[34]

There are several important relationships between our six econographic components and IQ and the Cognitive Reflection Test (CRT), shown in Table 8. Once again, we present the correlations in the form of a heatmap where shade is indicative of magnitude. Unlike

[34]Condon and Revelle (2014) contains the percent of people that typically get each of their questions correct. We used this information to attempt to cut the population at the 20th, 40th, 60th, 80th, 90th, and 95th percentile. We were largely successful, with 21% answering no questions correctly, 24% answering one correctly, 24% answering two correctly, 15% answering three correctly, 7% answering four correctly, 5% answering five correctly, and 4% answering all six correctly.

prior sections that were primarily concerned with magnitudes, here, positive and negative correlations have very different interpretations. As such, we distinguish between signs of the correlation using different colors: blue for negative, and green for positive. Consistent with the high correlation between the IQ and CRT measures, the general pattern of correlations is largely the same: higher cognitive ability is positively correlated with Generosity, and negatively correlated with Inequality Aversion/WTP, Overconfidence, and Impulsivity.[35]

These results show that, in general, the underlying behavioral components that we have identified are not simple proxies for intelligence. Moreover, there is a more nuanced relationship between our measures of intelligence and the extent to which a participant is "behavioral"—that is, behaves differently than a risk-neutral expected utility maximizer. While it is the case that lower measured intelligence is related to higher impulsivity and inequality aversion, it is also related to lower generosity. Moreover, risk and uncertainty attitudes are largely unrelated to intelligence, with much of the positive relationship between Inequality Aversion/WTP, Overconfidence, and intelligence being driven by distributional preferences.[36]

5.2 Econographics and Demographics

There are interesting relationships in our data between our components and standard economic variables—such as income and education—or demographics—such as age and gender. Table 9 shows these correlations in Panel A, and the same correlations controlling for IQ in Panel B.[37]

Greater education and income are associated with higher Generosity and lower Impulsivity, as was the case with higher cognitive ability. This may be unsurprising, as education and income are often associated with higher cognitive ability. However, as Panel B of Table

[35]Note that IQ is used to calculate one of the three main components of the Overconfidence measure, and also has some weight in the other components, which might result in some spurious correlation. This is unlikely to be a matter of great concern as the pattern of correlations with the CRT is largely the same.

[36]Appendix Table D.1 shows the relationships between individual econographics and cognitive measures.

[37]Appendix Table D.1 shows the relationships between individual econographics and demographics.

Table 9: Correlations with Demographics

	Income	Education	Age	Male	Attend Church
Panel A: Correlations					
Generosity	0.10 (.046)	0.06 (.049)	0.09 (.047)	0.04 (.046)	0.15 (.046)
Risk Aversion: WTA	0.00 (.049)	0.06 (.046)	0.03 (.050)	−0.03 (.047)	0.02 (.054)
Inequality Aversion/WTP	0.05 (.044)	−0.03 (.046)	0.20 (.045)	−0.08 (.046)	−0.06 (.054)
Overconfidence	0.00 (.052)	0.01 (.050)	0.04 (.047)	0.10 (.048)	0.12 (.046)
Impulsivity	−0.19 (.045)	−0.22 (.039)	0.05 (.045)	0.04 (.047)	0.05 (.046)
Uncertainty	0.03 (.045)	0.01 (.044)	−0.01 (.043)	−0.04 (.045)	−0.04 (.041)
Panel B: Controlling for IQ					
Generosity	0.08 (.045)	0.04 (.048)	0.11 (.046)	0.04 (.049)	0.16 (.048)
Risk Aversion: WTA	0.00 (.047)	0.06 (.045)	0.04 (.049)	−0.03 (.047)	0.02 (.053)
Inequality Aversion/WTP	0.07 (.044)	−0.01 (.046)	0.18 (.044)	−0.07 (.046)	−0.06 (.049)
Overconfidence	0.04 (.052)	0.07 (.049)	0.01 (.044)	0.12 (.048)	0.12 (.049)
Impulsivity	−0.16 (.047)	−0.18 (.041)	0.01 (.045)	0.06 (.047)	0.04 (.044)
Uncertainty	0.02 (.045)	0.00 (.043)	0.00 (.043)	−0.05 (.046)	−0.04 (.040)

Notes: Bootstrapped standard errors from 10,000 simulations in parentheses. Colors in heatmap change with each 0.05 of magnitude of correlation. Blue hues indicate negative correlations, green hues positive correlations.

9 shows, these relations remain even after controlling for IQ. There is no association between income or education and Overconfidence or Inequality Aversion/WTP; however, once we control for cognitive abilities, we obtain that higher education is (slightly) correlated with higher Overconfidence. That is, even though more education increases participants' own

ability and knowledge, it seems to increase their perceptions of these skills even more. This suggests, if anything, that education is not a "cure" for non-standard preferences.

Older people are more generous and men are more overconfident, in accordance with the literature.[38] These results only become stronger once we control for cognitive abilities. Additionally, older people exhibit a higher Inequality Aversion/WTP score, while men are slightly lower. As with cognitive abilities, we find that neither demographic is related to the other risk component, Risk Aversion: WTA, adding nuance on the relation between these aspects and risk preferences. Lastly, frequency of church attendance is positively associated with Generosity. It is also positively associated with Overconfidence, although not statistically significantly at conventional levels.

6 Relation to Current Theories

As suggested in prior sections, commonly used theories can account for some, but not all, of the patterns in our data, even within the social or risk domain. In this section, we spell out the correlations that would be predicted by different theories, and the points of difference between theories and our empirical patterns.

6.1 Social Preferences

Theories of social preferences can capture some of the moments in our data, but no single theory can explain all the patterns we observe. Here we consider four common theories of social preferences: a simple model of altruistic preferences, social welfare preferences (Charness and Rabin, 2002), inequality aversion (Fehr and Schmidt, 1999), and reciprocity (Rabin, 1993).

A simple model of altruistic preferences—in which utility is increasing in one's own

[38]Although not significant after Bonferroni correction, it is significant without. As this fact has been previously shown, for example in Ortoleva and Snowberg (2015b), there is little risk of this being a spurious correlation.

monetary payoff and that of others—predicts a positive correlation between our measure of Altruism, Trust, Reciprocity: High, and Reciprocity: Low, in line with the Generosity component (see Tables 3, 4, and 7). This simple model also predicts that more altruistic people would have to be paid more to move to an uneven split in which the other person got less, but paid less to move to a split in which the other person got more. That is, it also predicts that measures in the Generosity component should be positively correlated with Dislike Having More and negatively related to Dislike Having Less. There is some evidence for such a pattern in our data, however, it does not reflect the most parsimonious reading, as expressed by PCA. Moreover, a model of altruistic prefereces would also predict a *negative* correlation between Dislike Having More and Dislike Having Less—the opposite of what we find in our data. The altruistic preferences model also predicts that people would accept a reduction in allocation to move away from an even split to one in which another person got more. In our data, the majority of non-indifferent participants require an *increase* in income (79% of non-indifferent participants in question 1 and 74% in question 2) to move away from the even split, the opposite of the theoretical prediction. Finally, a purely altruistic person would never engage in costly punishment.

Social welfare preferences modify the altruistic preferences model by assuming that the utility a decision-maker gets from the consumption of others grows faster when others are poorer than the decision-maker (Charness and Rabin, 2002). Thus, altruistic behavior is governed by two parameters, one that dictates behavior when a person has more than others, and one when they have less. Like the altruism model, this model also predicts a positive correlation between Altruism, Trust, Reciprocity: High, and Dislike Having More, as these are all situations in which the decision-maker will (choose to) have more than the other person. In contrast, Reciprocity: Low would be governed by a separate parameter, which would also determine Dislike Having Less and lead to a negative correlation between these two measures. This is not the primary pattern we see in the data. As with altruistic preferences, a person of this type would never engage in costly punishment.

Perhaps the most common characterization of social preferences is Fehr and Schmidt's model of inequality aversion. This adds two parameters to the standard model: one codes how much the decision-maker dislikes having more than other people, and the other how much they dislike having less. As with social welfare preferences, the first of these parameters governs behavior in Altruism, Trust, Reciprocity: High, and Dislike Having More, while Dislike Having Less is governed by the second parameter. The model of Bolton and Ockenfels (2000) would imply similar correlations. Again this does not match the most parsimonious reading of our data. Moreover, a model of inequality aversion would predict that no money should be returned in the Reciprocity: Low measure, which is not what we find. Inequality-averse participants might engage in Pro-social Punishment (as this would reduce inequality) but not Anti-social Punishment (as this would increase inequality). This is inconsistent with the strong positive correlation we find between these two measures.

Intention-based models, such as Rabin (1993) have little bite in our setting, as only one of our choice environments—the trust game—has a second player whose actions affect each others' payoffs. However, this can plausibly lead to a negative correlation between Reciprocity: High (in which participants have been treated fairly) and Reciprocity: Low (in which they have been treated unfairly). This is in contrast to what we find in our data.

6.2 Risk Preferences

Our results exhibit inconsistencies with standard models of preferences under risk and uncertainty. These models typically separate risk and uncertainty, with the distinction being whether probabilities of outcomes are known or unknown (Arrow, 1951; Gilboa and Marinacci, 2016). The curvature of the utility function, and possibly aspects of non-Expected Utility, should affect behavior in both domains, while beliefs and ambiguity aversion should only impact choice under uncertainty. In our data, the classic delineation between risk and uncertainty fails to hold: attitudes towards compound lotteries, where probabilities are known, are related to ambiguity aversion, where they are not. This empirical relationship has been shown before, with different possible explanations (Halevy, 2007; Dean and Ortoleva, 2015; Gillen et al., 2018). Segal (1990) proposes that ambiguous prospects are treated like compound lotteries, in which case probability weighting could explain both phenomena. This theory would further imply a link between Ambiguity/Compound Lottery Aversion and the Common Ratio effect. However, we do not observe this relationship in our data. An alternative is to argue that compound lotteries are complex objects that people perceive as ambiguous (Dean and Ortoleva, 2015). In either case, these results point to different structure, where compound lotteries are associated with prospects with unknown probabilities.

The split into two separate types of risk preferences for objective, non-compound, risk is also difficult to reconcile with models that treat risk aversion as a unitary phenomenon—for example, driven by the curvature of a utility function. As discussed at length in Chapman et al. (2017), the specific clusters we find can be interpreted using the framework of Sprenger (2015), in which question structure serves as an implicit frame and determines reference points. In particular, Sprenger shows that certainty and probability equivalents yield different levels of risk aversion. He explains this finding by appealing to the different reference points these frames induce. In line with this hypothesis, the two clusters in our data are related to the two explicit frames we administer in the WTA and WTP tasks: certainly equivalents are correlated with WTA, while probability equivalents are correlated with WTP.

A natural question is whether this pattern of correlations is in line with existing theories of reference dependence. Cumulative Prospect Theory (CPT) and the Kőszegi Rabin (KR) model predict specific relationships between WTA, WTP, and the Endowment Effect, and the three measures Risk-Aversion: Gains, Losses, and Gain/Loss. While deriving these specific predictions and testing them is beyond the scope of this paper, Chapman et al. (2017) shows that the predicted relationships fail to hold, and in some cases are in the opposite direction of predictions from reference-dependence models. Thus, while the reference point induced in these questions is of clear first-order importance, existing theories of reference dependence cannot account for the overall pattern of correlations observed in our data.

7 Literature

We review two distinct areas of research. First, we briefly discuss the small, but significant, literature that examines connections between two or three different behavioral regularities, or between economic behavior and measured intelligence. Second, we describe in more detail studies that have estimated the correlation structure between many different economic behaviors. As far as we are aware, there are two contemporaneous projects that examine multiple behaviors in representative samples, and a further two that use restricted samples.

No existing study is well suited to our purpose: to provide a basis for more parsimonious behavioral models by examining the relationships between a wide range of behaviors in the risk, social, time, and confidence domains. Bilateral (or trilateral) analyses cannot reveal clusters of behaviors that move together, and therefore may be governed by the same underlying process. Studies that do contain multiple measures either focus on restricted samples—which may produce unrepresentative patterns of connections—or on constellations of measures that contain important differences from those we use. Moreover, studies with several measures and representative samples do not use incentives, nor take steps to eliminate the attenuating effects of measurement error.

7.1 Studies with a Small Number of Behavioral Measures

There are several laboratory studies of the relationship between different aspects of risk preferences, and risk preferences and time preferences. Many of the findings from these studies extend to our representative population. However, our results create a more nuanced picture of the complicated relationships between different facets of risk attitudes. Pedroni et al. (2017) find that risk attitudes are poorly correlated—and that parameter estimates from a CPT model are unstable—across elicitation tasks based on objective probabilities, a finding consistent with our result that risk attitudes are multi-faceted. A few studies find weak or non-existent relationships between attitudes towards risk and ambiguity/uncertainty (see, for example, Cohen et al., 1987; Lauriola and Levin, 2001; Chakravarty and Roy, 2009; Cohen et al., 2011; Ahn et al., 2014).[39] Our findings are similar, and we also document the relationship between ambiguity and compound lottery aversion previously established in a student sample in Halevy (2007).[40] Several studies describe and test theoretical relationships between risk and time preferences—both of which are related to the curvature of utility functions.[41] Generally, in our data, there is some evidence the two are related, but our results isolate the specific part of risk preferences that are associated with time preferences.

There are also papers which study the relationship between risk and/or time preferences and measured intelligence (Burks et al. 2009; Dohmen et al. 2010; Benjamin et al. 2013—see Dohmen et al. 2018 for a recent review). In aggregate, the literature suggests that higher cognitive ability is associated with more risk aversion, but with some subtleties. Andersson et al. (2016) suggest that the driving factor in these results might be increased noise in choice, rather than differing preferences. Several other studies find the relationship between risk aversion and cognitive ability to be negative in the gain domain and positive in the loss domain. Our results suggest a further nuance, as we find differing relationships between cognitive ability and the different components of risk attitudes we identify.

There are fewer studies of the relationships between different aspects of social preferences,

[39]von Gaudecker et al. (2011) document significant heterogeneity in the distribution of preference parameters governing risky choice in a broad population. However, due to the statistical methodology they employ it would be difficult to examine correlations between these parameters.

[40]Abdellaoui et al. (2015) find a weaker relationship, although Gillen et al. (2018) show that findings of weak relationships in experiments may be due to measurement error. We utilize Gillen et al.'s techniques in order to avoid such issues.

[41]Andersen et al. (2008) find that taking into account risk attitudes reduces estimated discount rates, implying a relationship between risk attitudes and discount rates calculated under the assumption of risk neutrality (as we do in this paper). Andreoni and Sprenger (2012b) find evidence for a failure of *interchangeability*: inter-temporal preferences vary depending on whether the objects of choice are lotteries or certain payments, which they interpret as evidence for different utility functions for risky and risk-free prospects. Finally, Epper et al. (2011) find a significant relationship between decreasing impatience and probability weighting in an experiment that links the two behaviors. Unlike our experiment, they estimate probability weighting parametrically from choices over lotteries, rather than using questions explicitly designed to identify the existence of probability weighting.

or between social preferences and risk. Peysakhovich et al. (2014) report findings similar to ours: cooperative behavior is correlated across a number of settings including the trust and dictator games, but is unrelated to costly punishment. At least one study documents a correlation between risk and inequality aversion, based on the idea that from behind a "veil of ignorance" inequality creates risk (Carlsson et al., 2005). There is conflicting evidence on whether there is a connection between risk attitudes and sender behavior in a trust game (Eckel and Wilson, 2004; Schechter, 2007; Epstein et al., 2016). We find no such relationship.

7.2 Studies with Large Numbers of Behavioral Measures

Our work is differentiated from most prior studies that contain multiple measures in that it has a large, representative sample. However, two contemporaneous and complimentary projects have those same features. Stango et al. (2017a) measure a broad set of 17 behavioral factors in a representative sample of U.S. adults to study, in part, the relationships between them. Falk et al. (2017) survey 80,000 adults across 80 countries to document patterns of six behaviors.[42] These studies differ from ours in both purpose and implementation: they consider different measures, do not use incentives, and do not take steps to eliminate the attenuating effects of measurement error.

Stango et al. have three aims: to relate measured behavioral factors to financial outcomes, to combine behavioral factors into a single measure of how "behavioral" a person is, and to study the correlations between the measures.[43] They collect non-incentivized measures of a substantially different—and more heterogeneous—set of behaviors than we do. These include many biases and mistakes, for example: the gambler's fallacy, non-belief in the law of large numbers, and exponential growth bias—as well as measures of risk and time preference.[44]

[42] Becker et al. (2012) measure risk aversion, time preferences, trust, and altruism using incentivized questions in a representative sample. However, they do not report the correlations between these behaviors, as their goal is to study the links between these behaviors and personality.

[43] See also Stango et al. (2017b).

[44] This is related to a literature, largely in psychology, that studies the empirical relationship between different kinds of violations of rationality such as base rate neglect and syllogistic reasoning (see, for example, Stanovich, 2011). Our focus is narrower in that we are more concerned with violations of the standard model

Unlike our study, they do not include measures of social preferences. Their data is thus not optimized to find common components within and across risk and social preferences.

Stango et al. find 1 to 3 components that are difficult to interpret from an economic point of view underlying their 17 measures.[45] They conclude that, "Our methods and results suggest poor prospects for reducing the whole of (behavioral economics) to a manageable number of (latent) traits, as had been done for intelligence and personality." In contrast, we can summarize our data using a small number of economically meaningful components. The difference is likely due to both the different set of measures they use, representing measures using a behavioral/non-behavioral dichotomy (rather than a continuum), and the fact that they do not take steps to reduce the attenuating effects of measurement error.[46]

Falk et al. collect a subset of the measures we include here: patience, risk, positive reciprocity, punishment, altruism, and trust. Each is measured through a combination of qualitative self-reports and hypothetical-money questions. Again, because of their different aims, their data is unsuitable for our purpose. This is, in part, because their measures are less specific than ours: for example they do not quantify the various different aspects of risk preferences. Correlations between these preferences are generally lower than what we observe (all are below 0.33). This may be due to the attenuating effects of measurement error, or the fact that they examine a more diverse population (across 80 countries).

An immediate predecessor of this paper is Dean and Ortoleva (2015), which studies relationships between many of the same behaviors in a sample of 180 Brown undergraduates. These students are likely to be less heterogeneous on some dimensions, such as cognitive ability. Dean and Ortoleva focus more on risk and time, and less on social preferences.

of economic choice, such as ambiguity aversion and altruism. These behaviors may represent preferences that are different from those assumed by the classic economic model, rather than violations of rationality.

[45] As the authors state in their abstract, "Our quest for common factor parsimony largely fails: within-[person] correlations between behavioral factors tend to be low, and the common factor contributing to all 17 behavioral factors within-individual is weakly identified and does not help explain outcomes conditional on the other covariates."

[46] Measurement error biases estimated correlations towards zero, reducing the number of meaningful components recovered by standard dimension-reduction techniques, including the one we use—principal components analysis (PCA). See Sections 3.2 and 3.4.

Finally, Burks et al. (2009) make use of a large scale experiment carried out on a group of newly recruited truck drivers. The authors use parametric methods to measure risk aversion, short-term and long-term discounting (though a beta-delta model) and behavior in a sequential two-person prisoner's dilemma (similar to our trust game). They find a statistically significant (though quantitatively small) relationship between risk attitudes, patience, and sender behavior in the trust game: more patient people tended to be less risk averse and send more in the trust game.

Broadly speaking, to the extent that the studies use similar measures, all four tend to find similar relationships. However, these correlations represent a very limited subset of the 210 we examine here: 45 for Dean and Ortoleva, 21 for Stango et al., and 15 for Falk et al.. For example, Stango et al. find relationships between their various overconfidence measures, Falk et al. find a relationship between altruism and reciprocity, and Dean and Ortoleva find correlations between risk aversion over gains and losses, and between reactions to ambiguous and compound lotteries.[47]

8 Conclusion

We elicit 21 econographics from a representative sample of 1,000 U.S. adults in order to create an empirical basis for the underlying structure of more comprehensive theories of behavioral decision-making. We identify six easy-to-interpret econographic components that explain a large fraction of the variance in these 21 econographic variables: Generosity, Risk Aversion: WTA, Inequality Aversion/WTP, Overconfidence, Impulsivity, and Uncertainty. These components suggest that more parsimonious representations of social preferences are possible, but that canonical theories of risk preferences are perhaps too parsimonious. Moreover, they suggest limited, and nuanced, connections between risk and social preferences. By studying the relationship between the components we identify and cognitive measures and

[47]There are also some differences, but the reasons for these are inseparable from the reasons for broad disagreement, discussed above.

demographics, we document several stylized facts that may be useful for theorizing.

A strength of our study is the number of behaviors included in our analysis. However, the behaviors we included were limited by the current literature.[48] These, in turn, present limits for our analyses. A nuanced view of these limitations comes from thinking about what would happen if we had included more, or fewer, measures, which we do in the remainder of this paper. This exercise also speaks to the robustness of our results.

Including a duplicate elicitation of "Risk Aversion: Gains" has little qualitative effect on our conclusions, as shown in Table D.3. This duplicate measure loads heavily on the second component in Table 7, and this component now becomes the first (in terms of percent of variation explained).[49] Thus, the ranking of components may respond to the inclusion or exclusion of measures. As such, we have not attached meaning to the ordering in the text.

Removing Dislike Having More reduces the number of components in Table 4 to two, but has little qualitative effect on the overall analysis of Table 7, as shown in Table D.4.[50] Dislike Having Less still combines with the variables in the Risk Aversion: WTP component to form Inequality Aversion/WTP. Thus, it appears that minor perturbations are not particularly consequential, although these exercises have little to say about larger changes.

The robustness here stems from the fact that the most correlations between measures—displayed, for example, in Table 3 and 5—are either large or close to zero.[51] There are a few middling correlations—such as those between Dislike Having More and other social preference measures. These are the likely sources of fragility in PCA. Adding a variable with a number of middling correlations may cause parallel analysis to suggest the inclusion of an additional component, and this inclusion may lead to extensive changes in existing

[48]For example, we would have liked to have a measure of motivated reasoning. However, we were unable to find such a measure in the literature, and our design principles, discussed in Section 2, dictated that we not attempt to invent our own.

[49]This measure, Risk Aversion: Urn is described in Footnote 16. Note that this measure also loads on the uncertainty component, although the correlations that leads to this loading are spurious, as Risk Aversion: Urn is used in constructing both of the measures that comprise the uncertainty component.

[50]The number of components is determined, as before, using parallel analysis. In analyzing just the social preference measures, the Dislike Having Less loads on the punishment component.

[51]The preceding two paragraphs are not comprehensive tests of robustness. Rather, they illustrate patterns in the data to build intuition about how PCA responds to small changes in which econographics are included.

components. As such, the components are useful for making sense of the large correlation matrices in our analysis. However, the correlation matrices themselves are the most robust—and consequential—part of the analysis, as these are the fundamental patterns upon which latent variable models, such as PCA, are built.

Overall, this discussion suggests that our main conclusions—that there is an underlying structure to our measures that is informative for theorizing—are robust. However, we note that our findings relate to a fairly specific domain of economic behaviors: choices over money lotteries, time, the distribution of resources between two people, and beliefs about oneself and others. Adding more measures might create broader clusters that increase the average number of measures per component, or lead to a more diffuse set of underlying dimensions. Whatever the outcome of such explorations, it is worth noting that a failure to reduce all economically-relevant behaviors to just a handful of components is not a failure of behavioral economics. Chemistry has been incredibly successful with more than 100 elements. It should be our goal to accurately and adequately describe economic behavior with no more, and no fewer, components than are necessary. This study takes a step in that direction.

References

Abdellaoui, Mohammed, Peter Klibanoff, and Lætitia Placido, "Experiments on Compound Risk in Relation to Simple Risk and to Ambiguity," *Management Science*, 2015, *61* (6), 1306–1322.

Abdi, Hervé and Lynne J Williams, "Principal Component Analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, 2010, *2* (4), 433–459.

Ahn, David, Syngjoo Choi, Douglas Gale, and Shachar Kariv, "Estimating Ambiguity Aversion in a Portfolio Choice Experiment," *Quantitative Economics*, 2014, *5* (2), 195–223.

Allais, Maurice, "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 1953, *21* (4), 503–546.

Andersen, Steffan, Glenn W. Harrison, Morten I. Lau, and E.Ẽlisabet Rutström E. Elisabetom, "Eliciting Risk and Time Preferences," *Econometrica*, 2008, *76* (3), 583–618.

Andersson, Ola, Håkan J Holm, Jean-Robert Tyran, and Erik Wengström, "Risk aversion relates to cognitive ability: Preferences or Noise?," *Journal of the European Economic Association*, 2016, *14* (5), 1129–1154.

Andreoni, James and Charles Sprenger, "Estimating time Preferences from Convex Budgets," *The American Economic Review*, 2012, *102* (7), 3333–3356.

_ **and** _ , "Risk Preferences Are Not Time Preferences," *American Economic Review*, 2012, *102* (7), 3357–3376.

Arrow, Kenneth J., "Alternative Approaches to the Theory of Choice in Risk-taking Situations," *Econometrica*, 1951, *19* (4), 404–437.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger, "Working Over Time: Dynamic Inconsistency in Real Effort Tasks," *The Quarterly Journal of Economics*, 2015, *130* (3), 1067–1115.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy, "Incentives in experiments: A Theoretical Analysis," *Journal of Political Economy*, Forthcoming.

Beattie, Jane and Graham Loomes, "The Impact of Incentives upon Risky Choice Experiments," *Journal of Risk and Uncertainty*, March 1997, *14* (2), 155–168.

Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse, "The Relationship between Economic Preferences and Psychological Personality Measures," *Annual Review of Economics*, 2012, *4* (1), 453–478.

Bendor, Jonathan and Piotr Swistak, "The Evolution of Norms," *American Journal of Sociology*, 2001, *106* (6), 1493–1545.

Benjamin, Daniel J. Sebastian A. Brown, and Jesse M. Shapiro, "Who is Behavioral? Cognitive Ability and Anomalous Preferences," *Journal of the European Economic Association*, 2013, *11* (6), 1231–1255.

Berg, Joyce, John Dickhaut, and Kevin McCabe, "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 1995, *10* (1), 122–142.

Bernheim, B.D̃ouglas, "The Good, the Bad, and the Ugly: A Unified Approach to Behavioral Welfare Economics," *Journal of Benefit-Cost Analysis*, 2016, *7* (1), 12–68.

Bolton, Gary E. and Axel Ockenfels, "ERC: A Theory of Equity, Reciprocity, and Competition," *The American Economic Review*, 2000, *90* (1), 166–193.

Burks, Stephen V. Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini, "Cognitive Skills affect Economic Preferences, Strategic Behavior, and Job Attachment," *Proceedings of the National Academy of Sciences*, 2009, *106* (19), 7745–7750.

Carlsson, Fredrik, Dinky Daruvala, and Olof Johansson-Stenman, "Are People Inequality-Averse, or Just Risk-Averse?," *Economica*, 2005, *72* (287), 375–396.

Cason, Timothy N. and Charles R. Plott, "Misconceptions and Game Form Recognition: Challenges to Theories of Revealed Preference and Framing," *Journal of Political Economy*, 2014, *122* (6), 1235–1270.

Chakravarty, Sujoy and Jaideep Roy, "Recursive Expected Utility and the Separation of Attitudes towards Risk and Ambiguity: An Experimental Study," *Theory and Decision*, 2009, *66* (3), 199–228.

Chang, Beryl and Fabrizio Ghisellini, *Behavioural Economics: An Unfinished Business*, Forthcoming., 2018.

Chapman, Jonathan, Pietro Ortoleva, Erik Snowberg, Mark Dean, and Colin Camerer, "Willingness-To-Pay and Willingness-To-Accept are Probably Less Correlated than You Think," 2017. California Insitute of Technology, *mimeo.*

Charness, Gary and Matthew Rabin, "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics*, 2002, *117* (3), 817–869.

_ , Uri Gneezy, and Brianna Halladay, "Experimental Methods: Pay One or Pay All," *Journal of Economic Behavior & Organization*, 2016, *131*, 141–150.

Chetty, Raj, "Behavioral Economics and Public Policy: A Pragmatic Perspective," *American Economic Review*, 2015, *105* (5), 1–33.

Cohen, Michèle, Jean-Marc Tallon, and Jean-Christophe Vergnaud, "An Experimental Investigation of Imprecision Attitude and its Relation with Risk Attitude and Impatience," *Theory and Decision*, July 2011, *71* (1), 81–109.

— , **Jean-Yves Jaffray, and Tanios Said**, "Experimental Comparison of Individual Behavior under Risk and under Uncertainty for Gains and for Losses," *Organizational Behavior and Human Decision Processes*, 1987, *39* (1), 1–22.

Condon, David M. and William Revelle, "The International Cognitive Ability Resource: Development and Initial Validation of a Public-domain Measure," *Intelligence*, 2014, *43* (2), 52–64.

Conlisk, John, "Why Bounded Rationality?," *Journal of Economic Literature*, 1996, *34* (2), 669–700.

Cubitt, Robin, Chris Starmer, and Robert Sugden, "On the Validity of the Random Lottery Incentive System," *Experimental Economics*, September 1998, *1* (2), 115–131.

Dean, Mark and Pietro Ortoleva, "Is it All Connected? A Testing Ground for Unified Theories of Behavioral Economics Phenomena," 2015. Columbia University, *mimeo.*

DellaVigna, Stefano, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 2009, *47* (2), 315–372.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 2010, *100* (3), 1238–1260.

— , — , — , **and** — , "On the Relationship between Cognitive Ability and Risk Preference," *Journal of Economic Perspectives*, 2018, *32* (2), 115–34.

Dunn, Olive Jean, "Estimation of the Means of Dependent Variables," *The Annals of Mathematical Statistics*, 1958, *29* (4), 1095–1111.

— , "Multiple comparisons among means," *Journal of the American Statistical Association*, 1961, *56* (293), 52–64.

Eckel, Catherine C. and Rick K. Wilson, "Is Trust a Risky Decision?," *Journal of Economic Behavior & Organization*, 2004, *55* (4), 447–465.

Ellsberg, Daniel, "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, 1961, *75* (4), 643–669.

Epper, Thomas, Helga Fehr-Duda, and Adrian Bruhin, "Viewing the Future through a Warped Lens: Why Uncertainty Generates Hyperbolic Discounting," *Journal of Risk and Uncertainty*, December 2011, *43* (3), 169–203.

Epstein, Ziv, Alexander Peysakhovich, and David G. Rand, "The Good, the Bad, and the Unflinchingly Selfish: Cooperative Decision-making can be Predicted with High Accuracy when using only Three Behavioral Types," in "Proceedings of the 2016 ACM Conference on Economics and Computation" ACM 2016, pp. 547–559.

Ericson, Keith M Marzilli, John Myles White, David Laibson, and Jonathan D Cohen, "Money Earlier or Later? Simple Heuristics explain Intertemporal Choices better than Delay Discounting Does," *Psychological Science*, 2015, *26* (6), 826–833.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," 2017. University of Bonn, *mimeo.*

_ , _ , _ , David Huffman, and Uwe Sunde, "An Experimentally Validated Preference Survey Module," 2013. University of Bonn, *mimeo.*

Fehr, Ernst and Klaus M. Schmidt, "A Theory Of Fairness, Competition, And Cooperation," *The Quarterly Journal of Economics*, August 1999, *114* (3), 817–868.

_ and Urs Fischbacher, "Third-party Punishment and Social Norms," *Evolution and Human Behavior*, 2004, *25* (2), 63–87.

Forsythe, Robert, Joel L. Horowitz, N.Ẽugene Savin, and Martin Sefton, "Fairness in Simple Bargaining Experiments," *Games and Economic behavior*, 1994, *6* (3), 347–369.

Frederick, Shane, "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 2005, *19* (4), 25–42.

Freeman, David, Yoram Halevy, and Terri Kneeland, "Eliciting Risk Preferences using Choice Lists," 2015. University of British Columbia, *mimeo.*

Friedman, Daniel, R Mark Isaac, Duncan James, and Shyam Sunder, *Risky Curves: On the Empirical Failure of Expected Utility*, Routledge, 2014.

Fudenberg, Drew, "Advancing Beyond Advances in Behavioral Economics," *Journal of Economic Literature*, Aug 2006, *44* (3), 694–711.

Furr, R.M̃ichael, *Psychometrics: An Introduction*, Sage Publications, 2017.

Gilboa, Itzhak and Massimo Marinacci, "Ambiguity and the Bayesian Paradigm," in Horacio Arló-Costa, Vincent F. Hendricks, and Johan van Benthem, eds., *Advances in Economics and Econometrics: Theory and Applications*, Vol. 1 of *Springer Graduate Texts in Philosophy*, Springer, 2016, pp. 385–439.

Gillen, Ben, Erik Snowberg, and Leeat Yariv, "Experimenting with Measurement Error: Techniques and Applications from the Caltech Cohort Study," *Journal of Political Economy*, 2018, *Forthcoming.*

Halevy, Yoram, "Ellsberg Revisited: An Experimental Study," *Econometrica*, 2007, *75* (2), 503–536.

Herrmann, Benedikt, Christian Thöni, and Simon Gächter, "Antisocial Punishment across Societies," *Science*, 2008, *319* (5868), 1362–1367.

Hey, John and Jinkwon Lee, "Do Subjects Separate (or Are They Sophisticated)?," *Experimental Economics*, September 2005, *8* (3), 233–265.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Kahneman, Daniel, "Maps of Bounded Rationality: Psychology for Behavioral Economics," *American Economic Review*, 2003, *93* (5), 1449–1475.

_ **and Amos Tversky**, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, March 1979, *47* (2), 263–292.

_ **, Jack L. Knetsch, and Richard H. Thaler**, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 1990, *98* (6), 1325–1348.

Karni, Edi and Zvi Safra, "'Preference Reversal' and the Observability of Preferences by Experimental Methods," *Econometrica*, 1987, *55* (3), 675–685.

Kerschbamer, Rudolf, "The Geometry of Distributional Preferences and a Non-parametric Identification Approach: The Equality Equivalence Test," *European Economic Review*, May 2015, *76*, 85–103.

Kőszegi, Botond, "Behavioral Contract Theory," *Journal of Economic Literature*, 2014, *52* (4), 1075–1118.

Kurata, Hiroshi, Hiroshi Izawa, and Makoto Okamura, "Non-Expected Utility Maximizers Behave as if Expected Utility Maximizers: An Experimental Test," *Journal of Socio-Economics*, 2009, *38* (4), 622–629.

Lauriola, Marco and Irwin P. Levin, "Relating Individual Differences in Attitude toward Ambiguity to Risky Choices," *Journal of Behavioral Decision Making*, 2001, *14* (2), 107–122.

Levine, David K., *Is Behavioral Economics Doomed?: The Ordinary versus the Extraordinary*, Cambridge, UK: Open Book Publishers, 2012.

Moore, Don A. and Paul J. Healy, "The Trouble with Overconfidence," *Psychological Review*, 2008, *115* (2), 502–517.

Nagel, Rosemarie, "Unraveling in Guessing Games: An Experimental Study," *The American Economic Review*, 1995, *85* (5), 1313–1326.

Ortoleva, Pietro and Erik Snowberg, "Are Conservatives Overconfident?," *European Journal of Political Economy*, 2015, *40* (B), 333–344.

_ **and** _ , "Overconfidence in Political Behavior," *American Economic Review*, February 2015, *105* (2), 504–535.

Pearson, Matthew and Burkhard C. Schipper, "The Visible Hand: Finger Ratio (2D: 4D) and Competitive Bidding," *Experimental Economics*, 2012, *15* (3), 510–529.

Pedroni, Andreas, Renato Frey, Adrian Bruhin, Gilles Dutilh, Ralph Hertwig, and Jörg Rieskamp, "The Risk Elicitation Puzzle," *Nature Human Behaviour*, 2017, *1* (11), 803–809.

Pew Research Center, *Evaluating Online Nonprobability Surveys*, www.pewresearch.org, May 2016.

Peysakhovich, Alexander, Martin A. Nowak, and David G. Rand, "Humans Display a 'Cooperative Phenotype' that is Domain General and Temporally Stable," *Nature Communications*, 2014, *5*, 4939.

Plott, Charles R. and K. Zeiler, "The Willingness to Pay-Willingness to Accept Gap, The 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Valuations," *American Economic Review*, 2005, *95* (3), 530–545.

Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics," *The American Economic Review*, 1993, *83* (5), 1281–1302.

Schechter, Laura, "Traditional Trust Measurement and the Risk Confound: An Experiment in Rural Paraguay," *Journal of Economic Behavior & Organization*, 2007, *62* (2), 272–292.

Segal, Uzi, "Two-Stage Lotteries without the Reduction Axiom," *Econometrica*, 1990, *58* (2), 349–377.

Snowberg, Erik and Leeat Yariv, "Testing the Waters: Behavior across Subject Pools," 2018. California Institute of Technology, *mimeo.*

Sprenger, Charles, "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Journal of Political Economy*, 2015, *123* (6), 1456–1499.

Stango, Victor, Joanne Yoong, and Jonathan Zinman, "The Quest for Parsimony in Behavioral Economics: New Methods and Evidence on Three Fronts," 2017. National Bureau of Economic Research Working Paper #23,507.

_ , _ , and _ , "Quicksand or Bedrock for Behavioral Economics? Assessing Foundational Empirical Questions," 2017. National Bureau of Economic Research Working Paper #23,625.

Stanovich, Keith, *Rationality and the Reflective Mind*, Oxford, UK: Oxford University Press, 2011.

Thaler, Richard H., "Behavioral Economics: Past, Present, and Future," *American Economic Review*, 2016, *106* (7), 1577–1600.

von Gaudecker, Hans-Martin, Arthur Van Soest, and Erik Wengström, "Heterogeneity in Risky Choice Behavior in a Broad Population," *The American Economic Review*, 2011, *101* (2), 664–694.

Weber, Elke U. and Eric J. Johnson, "Decisions under Uncertainty: Psychological, Economic, and Neuroeconomic Explanations of Risk Preference," in Paul W. Glimcher, Ernst Fehr, Colin Camerer, and Russell Alan Poldrack, eds., *Neuroeconomics: Decision Making and the Brain*, New York, NY: Elsevier, 2008, pp. 127–144.

Online Appendix—Not Intended for Publication

A Survey Implementation

A.1 Measures

This section builds on Section 2 with a more direct description of how we instantiate the various measures discussed there. In the interest of brevity, measures are summarized as succinctly as possible. The specific question wordings, randomization of question ordering, screenshots, and so on are online at hss.caltech.edu/∼snowberg/wep.html. As mentioned in Section 3.2, in order to reduce issues related to measurement error, we elicit many quantities twice, and apply ORIV. We sometimes refer to these second elicitations as "duplicates."

A.1.1 Altruism, Trust, Reciprocity, and Punishment

Altruism: Participants are given 6,000 points to allocate, in 1,000 point increments, between themselves and another randomly chosen participant (Forsythe et al., 1994). The amount given to the other participant is our measure of altruism.

Trust: Participants play a standard sender-receiver trust game, in both the sender and receiver roles (Berg et al., 1995). As the sender, they are given another stock of 6,000 points to keep or to send to a randomly chosen participant—the receiver—in 1,000 point increments.[1] The amount sent is doubled, and then the receiver can send some portion of the total back, which is again doubled. We treat the amount sent as our measure of trust.

Reciprocity, High and Low: We elicit the responses of each participant, acting as a receiver, to each possible quantity of points he or she may receive in the trust game. These are in 2,000 point increments, from 2,000 to 12,000 points, reflecting the six possible amounts

[1] Participants are informed that this randomly chosen participant is different from the one in the dictator game.

that may be sent. As we do not observe the response of the receiver to the lowest possible action (sending 0 points), it is possible that receivers encode sending 1,000 points as a neutral, or even positive, action. As such, we call the response to receiving 2,000 points "Reciprocity: Low", with the amount sent back when receiving 4,000 points as the duplicate measure for ORIV. Symmetrically, we label the amount sent back when receiving 12,000 points as "Reciprocity: High," although it could safely be labeled positive reciprocity. The amount sent back when receiving 10,000 points is the duplicate measure for ORIV.

Pro- and Anti-social Punishment: Participants observe another sender-receiver trust game that they do not participate in, in which the sender sent 6,000 points and the receiver sent back 0 points, resulting in payoff of 12,000 for the receiver and 0 for the sender. Participants are given a stock of 4,000 points, which they can use to deplete the points of the sender and the receiver. Each point spent depletes the points of the targeted individual (the sender or the receiver) by six points. They can spend up to 2,000 points punishing each participant, and any unused points are kept. Participants are told (truthfully) that no one will have the opportunity to punish their own behavior, and that if their responses are randomly selected for implementation, no one else will have the opportunity to punish the sender or receiver.[2]

The amount used punishing the receiver is our measure of pro-social punishment, as this can be seen as a participant attempting to enforce the pro-social norm of trust. We label the amount used to punish the sender anti-social punishment (as extensively documented in Herrmann et al., 2008).[3]

A.1.2 Distributional Preferences

Our design for eliciting distributional preferences follows Kerschbamer (2015), which nests several popular models including Fehr and Schmidt (1999) and Charness and Rabin (2002).

[2]This measure was made incentive compatible by applying the chosen punishments to additional participants recruited to take part in a lab experiment.

[3]Note that in the context of public goods games, pro-social punishment is sometimes referred to in the literature as "third-party costly punishment." (Bendor and Swistak, 2001; Fehr and Fischbacher, 2004).

This technique results in two measures. In both of these, the participant decides on an allocation of points between themselves and another, randomly-matched, participant.

Dislike Having Less: In each line of the MPL, the participant decides between an even allocation (say, 4,000 points each), or an uneven allocation that favors another participant. In the uneven allocation, the other participant always receives the same amount, which is greater than both the even allocation and the allocation to the participant making the decision (say, 6,500 points). Each line of the MPL increments the amount to the participant making the choice (from, say, 1,600 points to 6,400 points). This defines how much the decision-making participant is willing to forgo in order to avoid having less than the other participant. Note that this quantity can be negative, in the sense that being willing to take, say, 3,000 points in order to ensure the other person will get 6,500 points means they are willing to forgo some points to ensure that the other person has more. Note, however, that a negative quantity here does not imply that the participant dislikes having more.

Dislike Having More: In this case, the participant is deciding between an even split, and another split that would give the other participant a lower total (say, 1,500 points) and the decision-making participant more (say, between 1,600 and 6,500 points). Once again, this quantity can be negative if the decision-making participant is willing to diminish their own points in order to ensure that they get more.

A.1.3 Risk and Loss Aversion

This and the following three sets of measures require participants to price lotteries in various ways. All lotteries involve only two possible payoffs, and most assign 50% probability to each. The specific point values of the lotteries used to measure each quantity can be found in Table A.1. Question wordings and screenshots can be found at hss.caltech.edu/∼snowberg/wep.html.

Risk Aversion: Gains: Participants are asked to give their certainty equivalent, using an MPL, for two lotteries containing either gains, or a gain and a zero amount.

Risk Aversion: Losses: Certainty equivalents are elicited for two lotteries that contain either only losses, or a loss and a zero amount.

Risk Aversion: Gain/Loss: Certainty equivalents are elicited for two lotteries that contain a loss and a gain of equal magnitude.

A.1.4 Common Ratio

The common ratio effect suggests non-linear preferences over, or misperceptions of, probabilities (Allais, 1953).

Risk Aversion: CR Certain: The participant gives the lottery equivalent for a fixed certain amount a. That is, for a fixed probability α, this question identifies b such that $a \sim \alpha b + (1-\alpha)0$ for that participant. (b is sometimes referred to as a *gain-equivalent.*)

Risk Aversion: CR Lottery: Fixing a ratio $x > 1$, we use the same a and α as above, and the participant gives lottery equivalent c such that $(1/x)a+(1-1/x)0 \sim (\alpha/x)c+(1-\alpha/x)0$.

Common Ratio: Under Expected Utility, $b = c$. We measure the common ratio effect as the difference between the two measures, $b - c$.

A.1.5 Uncertainty

Risk Aversion: Urn: This is similar to "Risk Aversion: Gains" however, the participant is told that his or her payment is tied to a random draw from a jar with 100 balls, divided 50/50 between two colors, rather that being told there is a 50% probability of a certain payment. One color is tied to a high payoff, and the other to a zero payoff. In this and the following two measures the participant is asked to choose which color they would like to result in the high payoff. Empirically, this measure is highly correlated with "Risk Aversion: Gains", and is primarily used to difference out the risk component from ambiguous and compound urns.[4]

Compound Lottery Aversion: The payoffs from this lottery are also tied to a draw from an urn. However, participants are told that the exact composition of the urn is not yet determined. Rather, a number between 0 and 100 will be drawn, and this will be the number of balls of a specific color in the jar, and the other color will make up the remaining balls. The certainty equivalent for this lottery, minus the certainty equivalent for Risk Aversion: Urn is the measure of compound lottery aversion (Halevy, 2007).

Ambiguity Aversion: Here, the payoffs are tied to a draw from an Ellsberg (1961) urn—that is, an urn with an unknown proportion of balls of each color. The certainty equivalent for this lottery, minus the certainty equivalent for Risk Aversion: Urn is the measure of ambiguity aversion.

Note that both Compound Lottery Aversion and Ambiguity Aversion difference out the same quantity. If this quantity is measured with error, this can create spurious correlation between Compound Lottery Aversion and Ambiguity Aversion. This issue, and our solution, are discussed in Section 3.2.

A.1.6 Endowment Effect

WTP: Participants are given a stock of points they can divide between keeping and buying a lottery ticket, using an MPL. The amount they are willing to spend on the ticket is used as their willingness to pay (WTP).

WTA: Participants are endowed with a lottery ticket, and their selling price is elicited using an MPL. The amount they are willing to sell the ticket for is their willingness to

[4] It is also used to illustrate an aspect of principal components analysis in Section 8.

Table A.1: Lotteries Used for Risk Measures

	Lottery 1	Lottery 2
Panel A: Fixed Lotteries		
Risk Aversion: Gains	$0.5 * 0 \oplus 0.5 * 5{,}000$	$0.5 * 1{,}000 \oplus 0.5 * 4{,}000$
Risk Aversion: Losses	$0.5 * -5{,}000 \oplus 0.5 * 0$	$0.5 * -4{,}000 \oplus 0.5 * -1{,}000$
Risk Aversion: Gain/Loss	$0.5 * -5{,}000 \oplus 0.5 * 5{,}000$	$0.5 * -4{,}000 \oplus 0.5 * 4{,}000$
Risk Aversion: Urn	$0.5 * 0 \oplus 0.5 * 10{,}000$	$0.5 * 0 \oplus 0.5 * 8{,}000$
Compound Lottery Aversion	$0.5 * 0 \oplus 0.5 * 10{,}000$	$0.5 * 0 \oplus 0.5 * 8{,}000$
Ambiguity Aversion	Ambiguous $\{0, 10{,}000\}$	Ambiguous $\{0, 8{,}000\}$
WTA	$0.5 * 1{,}000 \oplus 0.5 * 10{,}000$	$0.5 * 2{,}000 \oplus 0.5 * 8{,}000$
WTP	$0.5 * 1{,}000 \oplus 0.5 * 10{,}000$	$0.5 * 2{,}000 \oplus 0.5 * 8{,}000$
Panel B: Variable Lotteries		
Common Ratio: High Probability		
Fixed (a)	2,500	4,000
Variable	$0.2 * 0 \oplus 0.8 * b$ $b \in [2{,}200, 7{,}000]$	$0.25 * 0 \oplus 0.75 * b'$ $y \in [3{,}600, 10{,}000]$
Common Ratio: Low Probability		
Fixed (a)	$0.75 * 0 \oplus 0.25 * 2{,}500$	$0.8 * 0 \oplus 0.2 * 4{,}000$
Variable	$0.8 * 0 \oplus 0.2 * c$ $c \in [2{,}200, 7{,}000]$	$0.85 * 0 \oplus 0.15 * c'$ $c' \in [3{,}600, 10{,}000]$

Notes: Possible values for $b, c \in \{2200, 2500, 2800, 3100, 3400, 3700, 4000, 4300, 4600, 4900, 5200, 5500, 5800, 6100, 6400, 6700, 7000\}$, and $b', c' \in \{3600, 4000, 4400, 4800, 5200, 5600, 6000, 6400, 6800, 7200, 7600, 8000, 8400, 8800, 9200, 9600, 10000\}$.

accept (WTA).

Endowment Effect: The endowment effect is measured as the difference between WTA and WTP for the same lottery ticket (Kahneman et al., 1990; Plott and Zeiler, 2005).

The same two lottery tickets are used in both WTA and WTP. In order to avoid consistency effects, one of WTA or WTP is elicited near the beginning of the survey, and the other is elicited near the end. Whether WTA or WTP comes first for a given participant is randomly determined.

A.1.7 Overconfidence

Overconfidence can be divided into three types: overestimation, overplacement, and overprecision (Moore and Healy, 2008). Overestimation refers to a person's estimate of her performance on a task (versus her actual performance). Overplacement refers to her perceived performance relative to other participants (versus her real relative performance). Overprecision refers to a belief that one's information is more precise than it actually is.

Overestimation: After each block of three questions on the in-survey IQ test (described in Section 5.1) the participant was asked how many questions they thought they got correct (out of three). This, minus the actual number correct is used as our measure of overestimation.

Overprecision: Our measures of overprecision and overplacement follow Ortoleva and Snowberg (2015b,a). Participants are asked two factual questions: about the year the telephone was invented, and the current unemployment rate. After each factual question, the participant is asked for a qualitative assessment of the accuracy of their answer. This residual from a regression of this measure on a fourth-order polynomial of the participant's accuracy gives overprecision.

Overplacement: After giving their subjective perception of how accurate their answers were, the participant is asked to give their perception about their accuracy compared to a random group of 100 other survey takers. This answer is the perceived percentile of their accuracy. The actual percentile of their accuracy is subtracted from this number to give our measure of overplacement.

Note that as both overprecision and overplacement refer to the same factual questions, this will create correlated measurement error between them. This can create spurious correlation between the two measures. This issue, and our solution, is discussed in Section 3.2.

A.1.8 Patience (Time Preferences)

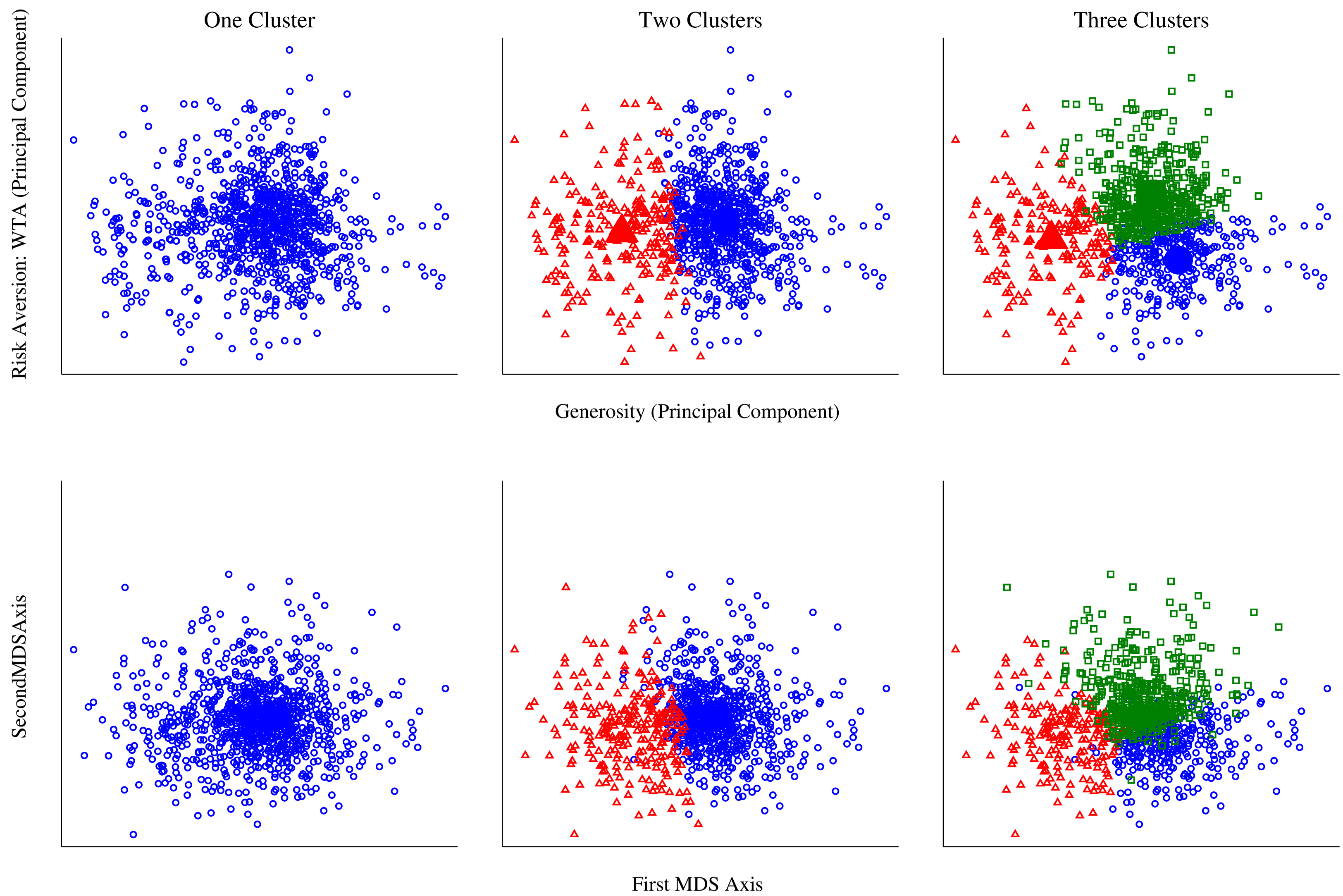
Patience is measured using the standard "Money Earlier or Later" (MEL) paradigm (for example, Marzilli Ericson et al., 2015). Participants were offered 4,000 points in 45 days, and, in the duplicate, 6,000 points in 90 days. In each row of an MPL, the participant could select the delayed payment, or an alternative amount immediately, where the alternative amount varied between rows.

B Machine Learning

As noted in various points in the text, a few readers have suggested applying machine learning techniques to our data, especially clustering algorithms. As we also noted in the text, these methods failed to produce additional insight. This brief appendix is intended to give readers an understanding of why these methods were not useful in summarizing our data. It is not a comprehensive summary of all of the machine learning techniques we tried, rather, it shows, through the use of clustering algorithms, why those techniques are not useful in this context. There are two causes: there are no strong clusters in our data, and reasons related to standard "curse of dimensionality" problems prevent the identification of more subtle clusters.

In order to provide some intuition, we graph our data, although visualizing a 21-dimensional space is challenging. We approach visualization in two ways. First, we use the first two principal components identified in the text as axes. By construction, these components contain the most information about the underlying data. We also use multi-dimensional scaling, a common visualization technique that tries to preserve the distance between points while projecting them onto a lower dimensional space. The axes produced by this process are not directly interpretable. The figures in this section should provide some insight into why we did not include any figures of raw data in the text: reducing 21-dimensional data to two dimensions loses a lot of information.

Figure B.1: The optimal number of clusters in our data is one.



Notes: Top three panels show scatter plots of data in the space defined by the first two principal components of all 21 measures. Color and shape indicates cluster. Large, filled-in shape indicates centroid of a cluster. Bottom three panels show scatter plots of the data in a space defined by multi-dimensional scaling (MDS).

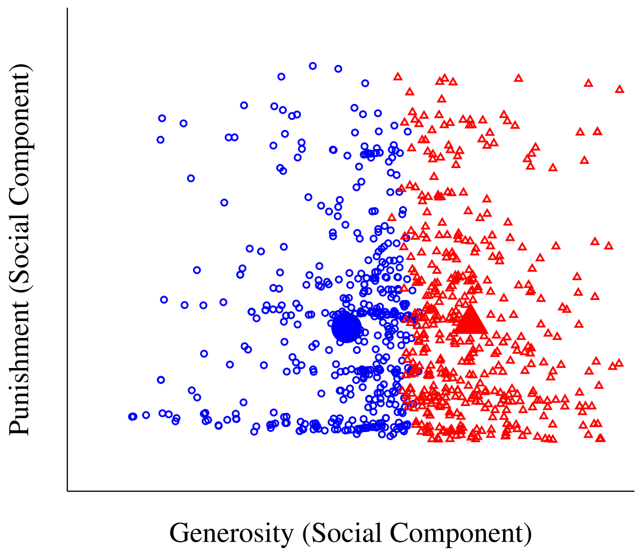
Using standard procedures, such as the so-called "elbow method", unsupervised clustering algorithms (whether using k-medians or k-means) produce an optimal number of clusters: in our case one cluster, illustrated in the first panel of Figure B.1. This figure plots the data along the first two principal components, averaging measures when we have more than one for a given attribute (such as risk aversion, etc.) Visual inspection should make it obvious why the optimal number of clusters is one: the data look to be distributed according to a single-mode joint distribution, such as a multi-variate normal. We note that the optimal clustering scheme takes into account information from all 21 dimensions, not just the two principal components that are displayed.

Forcing the clustering algorithm to include two or three clusters, as in the second and third panels of Figure B.1, seems to just impose arbitrary dividing lines on the single cluster in the first panel. Using a decision-tree approach to understand the classification shows that the most important determinant of the two clusters is whether or not someone is extremely risk loving on the Risk Aversion: Gain / Loss elicitation. Deeper decision trees provide just as little insight. Substantially similar results are observed by illustrating these three clustering schemes using MDS, as in the bottom three panels of Figure B.1.

Narrowing our focus to a particular domain, such as social preferences, does not help. Even with only our 8 social preference measures, the optimal number of clusters is still one, although the data is "lumpier," as displayed in Figure B.2. Forcing two clusters does not produce much insight. The clusters are largely defined by whether or not someone falls on one side or the other of the median level of Generosity (the first principal component of the social analysis). This dividing line goes through what appears to be the densest part of the distribution, rather than defining two largely distinct clusters.

While it is relatively clear that there are no strong clusters in our data, our data would not be optimal for finding more subtle clusters. To understand why, it is instructive to consider the dimensionality of the problem. While 21 dimensions may seem small compared to our 1,000 data points, in the realm of machine learning this is simply not the case. To

Figure B.2: Restricting focus to social preferences does not improve the performance of clustering algorithms.



Notes: Scatter plots of data in the space defined by the first two of the eight social preference measures. Color and shape indicate cluster. Large, filled-in shape indicates centroid of a cluster.

understand why, suppose we took our 21 continuous measures and divided them along the median value of that measure. This would produce $2^{21} \approx 2{,}000{,}000$ cells. Even with such a coarse classification, the 1,000 datapoints fill 979 cells. Restricting to the domain of social preferences, with $2^8 = 256$ cells, the data is still spread out across 133 of them.

It is worth noting that PCA is an often-used tool in machine learning. That is, displaying our main results does take advantage of a machine learning technique, just a particularly simple one that is understood by some economists. The visualizations here build on that analysis, although only in a limited way. However, trying to cluster participants on only their principal component scores (a six-dimensional space) also results in one optimal cluster. This should be unsurprising, as the first six principal components include most of the information in all 21 dimensions.

As noted at the beginning of this appendix, we also tried other types of machine learning analyses, such as outlier detection. Moreover, we also tried building clustering and decision-tree analyses on top of our principal components by, say, taking only one variable from each cluster identified by PCA. These different procedures do not produce additional insight.

C Scree Plots of PCAs

Figure C.1: Scree Plot for PCA of Social Measures

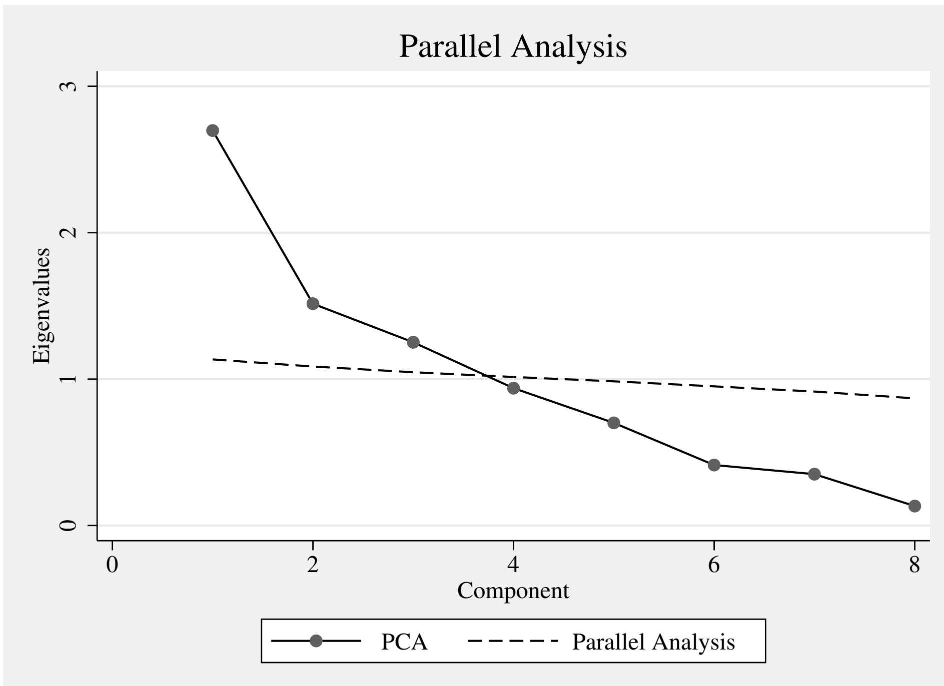


Figure C.2: Scree Plot for PCA of Risk Measures

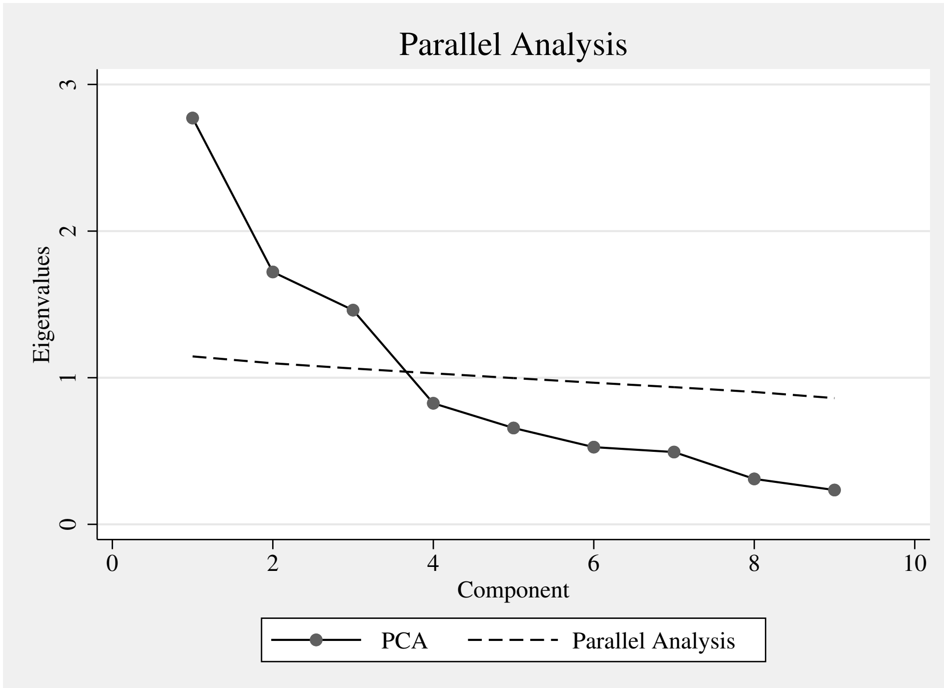
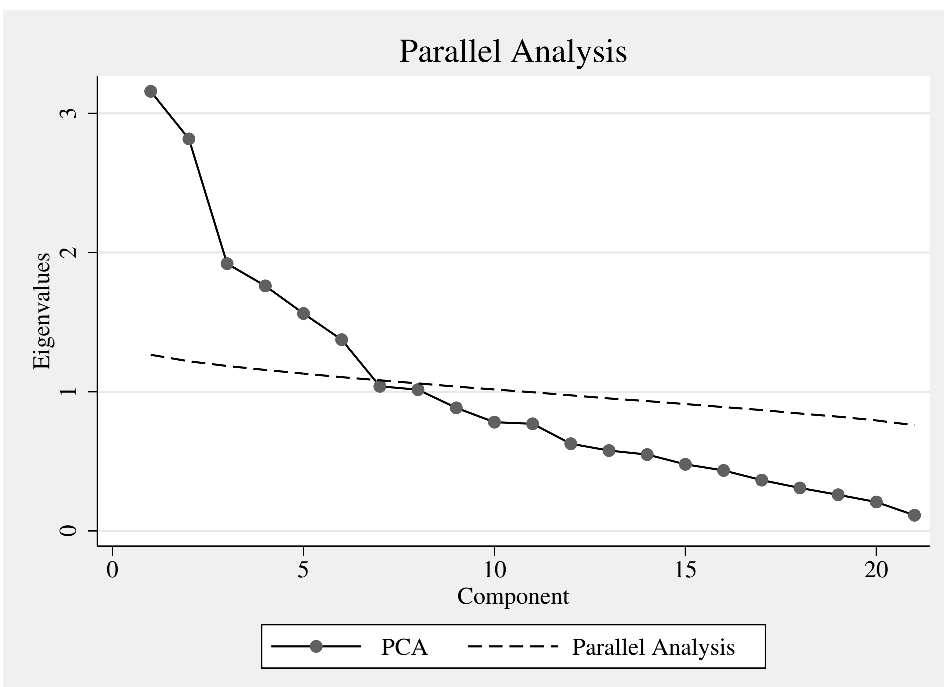


Figure C.3: Scree Plot for PCA of Risk, Social, Overconfidence and Time Preferences



D Other Analyses

Figure D.1: CDF of responses to "Beauty Contest" questions show no signs of strategic sophistication.

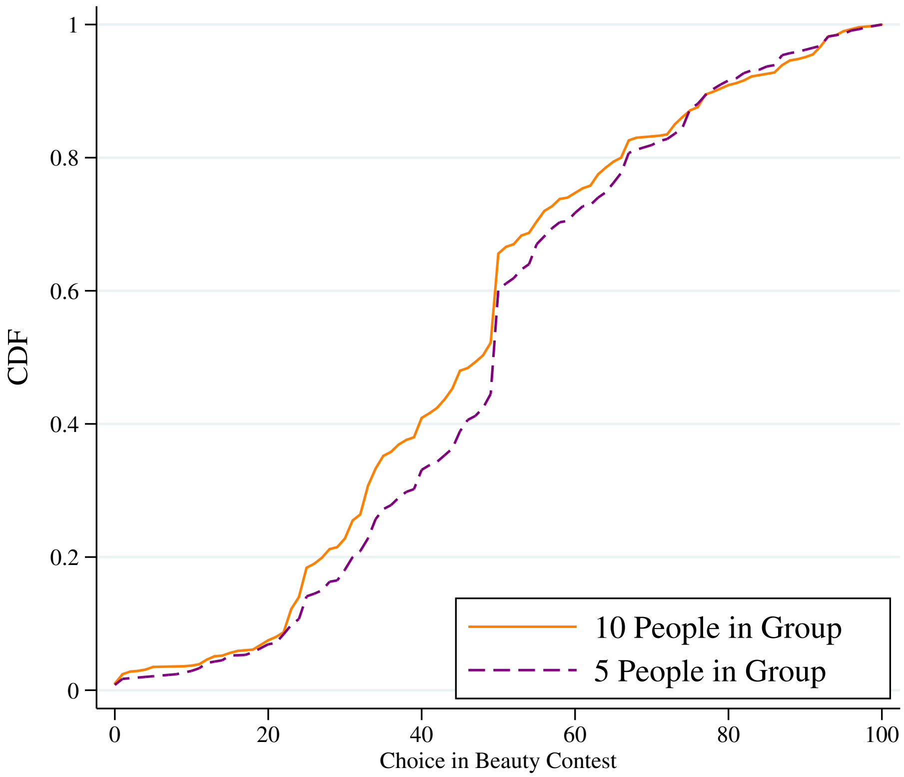


Table D.1: Correlations of Econographics with Cognitive Measures and Demographics

	IQ	CRT	Income	Education	Age	Male	Attend Church
Reciprocity: Low	0.06	0.05	0.05	0.00	0.06	0.04	0.19
	(.041)	(.041)	(.048)	(.052)	(.044)	(.048)	(.043)
Reciprocity: High	0.15	0.09	0.08	0.04	0.09	0.04	0.19
	(.038)	(.039)	(.044)	(.049)	(.044)	(.047)	(.042)
Altruism	−0.03	−0.05	0.01	0.02	0.06	0.03	0.04
	(.046)	(.047)	(.047)	(.046)	(.047)	(.046)	(.046)
Trust	0.13	0.07	0.06	0.07	0.05	0.06	0.08
	(.040)	(.039)	(.043)	(.042)	(.043)	(.043)	(.044)
Anti-Social Punishment	−0.23	−0.26	−0.17	−0.21	0.03	−0.01	0.02
	(.034)	(.032)	(.047)	(.043)	(.043)	(.047)	(.044)
Pro-Social Punishment	−0.01	−0.03	−0.07	−0.10	0.03	0.08	0.03
	(.040)	(.041)	(.045)	(.043)	(.046)	(.044)	(.044)
Dislike Having More	0.17	0.14	0.17	0.09	0.10	−0.11	0.02
	(.037)	(.043)	(.045)	(.046)	(.055)	(.049)	(.060)
Dislike Having Less	−0.09	−0.08	0.03	0.00	0.13	−0.18	0.03
	(.041)	(.048)	(.046)	(.044)	(.051)	(.042)	(.053)
WTA	−0.04	−0.09	−0.03	−0.10	−0.03	0.01	0.00
	(.041)	(.036)	(.052)	(.048)	(.049)	(.050)	(.052)
Risk Aversion: Gains	0.00	0.06	0.04	0.09	0.08	−0.01	−0.01
	(.043)	(.040)	(.052)	(.050)	(.050)	(.050)	(.054)
Risk Aversion: Losses	0.00	−0.03	−0.08	−0.03	0.01	−0.06	0.06
	(.045)	(.048)	(.056)	(.056)	(.058)	(.055)	(.062)
Risk Aversion: Gain/Loss	0.06	0.02	0.02	0.06	−0.01	−0.08	0.05
	(.043)	(.044)	(.050)	(.051)	(.054)	(.051)	(.057)
WTP	0.06	0.05	−0.08	0.04	−0.11	0.03	0.09
	(.046)	(.046)	(.047)	(.053)	(.046)	(.049)	(.051)
Risk Aversion: CR Certain	−0.01	0.03	0.08	0.02	0.16	0.03	0.03
	(.041)	(.040)	(.049)	(.047)	(.044)	(.048)	(.048)
Risk Aversion: CR Lottery	−0.11	−0.11	−0.05	−0.04	0.08	−0.04	−0.08
	(.047)	(.044)	(.048)	(.048)	(.051)	(.050)	(.051)
Ambiguity Aversion	0.05	0.01	0.02	0.00	−0.02	0.02	−0.03
	(.040)	(.041)	(.047)	(.047)	(.046)	(.049)	(.045)
Compound Lott. Aversion	0.01	0.03	0.00	0.00	−0.01	−0.06	−0.07
	(.040)	(.043)	(.050)	(.053)	(.054)	(.053)	(.046)
Overestimation	−0.62	−0.22	−0.09	−0.11	0.20	0.12	0.18
	(.061)	(.059)	(.063)	(.059)	(.060)	(.064)	(.062)
Overplacement	−0.04	−0.01	0.11	0.05	0.06	−0.07	0.10
	(.089)	(.067)	(.066)	(.075)	(.075)	(.077)	(.085)
Overprecision	0.03	−0.03	0.10	0.13	−0.07	0.16	0.10
	(.062)	(.060)	(.063)	(.064)	(.061)	(.061)	(.069)
Patience	0.23	0.21	0.16	0.16	0.04	−0.08	0.06
	(.038)	(.039)	(.051)	(.046)	(.050)	(.054)	(.056)

Table D.2: Principal Components Analysis of Risk Preferences with Common Ratio Measure

	Risk Aversion: WTA	Risk Aversion: WTP	Uncertainty	Unexplained
WTA	**−0.50**	0.09	0.11	0.33
Risk Aversion: Gains	**0.54**	0.08	−0.17	0.24
Risk Aversion: Losses	**0.37**	0.04	0.39	0.35
Risk Aversion: Gain/Loss	**0.52**	0.01	0.15	0.22
WTP	0.05	−0.09	**0.66**	0.39
Common Ratio	0.20	−0.09	**−0.59**	0.47
Ambiguity Aversion	0.00	**0.70**	0.02	0.27
Compound Lottery Aversion	0.01	**0.69**	−0.03	0.28
Percent of Variation	33%	19%	17%	32%

Notes: First three principal components using the varimax rotation. Weights greater than or equal to 0.25 in bold.

Table D.3: Adding an additional risk aversion measure switches the ranking of several components, while leave the structure largely unchanged.

	Risk Aversion: WTA	Generosity	Inequality Aversion/ WTP	Uncertainty	Overconfidence	Impulsivity	Unexplained
Reciprocity: Low	0.00	**0.50**	0.00	−0.02	0.02	0.05	0.29
Reciprocity: High	0.01	**0.51**	−0.02	0.04	−0.07	0.06	0.27
Altruism	0.05	**0.39**	0.03	−0.02	0.04	0.05	0.55
Trust	0.01	**0.47**	−0.06	−0.05	0.12	−0.04	0.33
Anti-social Punishment	0.00	−0.01	0.03	−0.02	0.05	**0.67**	0.26
Pro-social Punishment	−0.04	0.09	−0.04	0.02	−0.08	**0.59**	0.42
Dislike Having More	0.02	0.22	**0.27**	0.14	−0.18	−0.11	0.58
Dislike Having Less	−0.08	−0.06	**0.40**	0.05	−0.02	0.21	0.49
WTA	**−0.46**	−0.04	−0.05	0.06	0.06	0.04	0.30
Risk Aversion: Gains	**0.52**	0.04	0.10	0.09	0.04	−0.05	0.20
Risk Aversion: Losses	**0.32**	−0.08	−0.18	0.16	0.04	0.16	0.44
Risk Aversion: Gain/Loss	**0.45**	0.00	−0.12	0.10	−0.05	−0.01	0.27
Risk Aversion: Urn	**0.40**	−0.04	0.03	**−0.35**	0.01	0.00	0.18
WTP	0.00	0.00	**−0.47**	−0.01	0.14	0.01	0.45
Risk Aversion: CR Certain	0.09	−0.02	**0.52**	−0.07	0.07	−0.01	0.39
Risk Aversion: CR Lottery	−0.03	−0.03	**0.42**	0.05	0.11	−0.02	0.57
Ambiguity Aversion	0.03	0.04	−0.01	**0.65**	−0.04	0.05	0.29
Compound Lottery Aversion	0.03	−0.07	0.02	**0.61**	0.08	−0.07	0.30
Overestimation	−0.03	0.04	0.01	0.00	**0.51**	0.10	0.54
Overplacement	0.03	0.03	0.15	−0.03	**0.49**	−0.09	0.57
Overprecision	−0.01	−0.01	−0.07	0.04	**0.62**	−0.02	0.33
Patience	−0.17	0.19	−0.05	0.02	−0.05	**−0.28**	0.65
Percent of Variation	14%	13%	10%	8%	8%	7%	39%

Notes: First six principal components using the varimax rotation. Weights greater than or equal to 0.25 in bold.

Table D.4: Removing "Dislike Having More" leaves the overall structure of components unchanged.

	Generosity	Risk Aversion: WTA	Inequality Aversion/ WTP	Overconfidence	Impulsivity	Uncertainty	Unexplained
Negative Reciprocity	**0.52**	0.01	0.03	−0.01	0.04	0.03	0.26
Positive Reciprocity	**0.53**	0.02	0.01	−0.09	0.05	0.07	0.24
Altruism	**0.39**	0.05	0.02	0.06	0.05	−0.07	0.56
Trust	**0.47**	0.01	−0.06	0.13	−0.04	−0.07	0.34
Anti-Social Punishment	−0.01	0.00	0.03	0.05	**0.67**	−0.02	0.27
Pro-Social Punishment	0.08	−0.05	−0.05	−0.07	**0.60**	0.03	0.41
Dislike Having Less	−0.07	−0.11	**0.38**	−0.01	0.22	0.05	0.51
WTA	−0.04	**−0.50**	−0.07	0.06	0.04	0.10	**0.31**
Risk Aversion: Gains	0.04	**0.55**	0.12	0.03	−0.05	0.06	0.23
Risk Aversion: Losses	−0.11	**0.36**	−0.18	0.08	0.17	0.06	0.41
Risk Aversion: Gain/Loss	−0.01	**0.50**	−0.10	−0.04	−0.01	0.03	0.24
WTP	−0.03	0.01	**−0.49**	0.16	0.02	−0.05	0.44
Risk Aversion: CR Certain	0.00	0.11	**0.56**	0.03	−0.02	−0.07	0.35
Risk Aversion: CR Lottery	−0.01	−0.03	**0.44**	0.09	−0.03	0.05	0.55
Ambiguity Aversion	0.06	0.00	−0.01	−0.04	0.05	**0.72**	0.23
Compound Lottery Aversion	−0.06	0.01	0.02	0.08	−0.07	**0.65**	0.30
Overestimation	0.02	−0.03	0.00	**0.52**	0.11	−0.02	0.52
Overplacement	0.03	0.04	0.17	**0.48**	−0.09	−0.04	0.57
Overprecision	−0.02	−0.02	−0.08	**0.63**	−0.02	0.05	0.32
Patience	0.20	−0.18	−0.04	−0.06	**−0.28**	0.06	0.64
Percent of Variation	14%	13%	10%	8%	8%	7%	39%

Notes: First six principal components using the varimax rotation. Weights greater than or equal to 0.25 in bold.