Self-Centered Beliefs : An Empirical Approach

Eugenio Proto and Daniel Sgroi

No 978

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Self-Centered Beliefs: An Empirical Approach

Eugenio Proto and **Daniel Sgroi** [a] [*]

[a]University of Warwick and CAGE

February 2012

Abstract

We perform an experiment designed to assess the accuracy of beliefs about distributions. The beliefs relate to behavior (mobile phone purchasing decisions, hypothetical restaurant choices), attitudes (happiness, politics) and observable characteristics (height, weight) and are typically formed through real world experiences. We find a powerful and ubiquitous bias in perceptions that is "self-centered" in the sense that an individual's beliefs about the population distribution changes with their own position in the distribution. In particular, those at extremes tend to perceive themselves as closer to the middle of the distribution than is the case. We discuss possible explanations for this bias.

Keywords: subjective beliefs, attitudes, observable characteristics, self-centered bias.

JEL classification: D03, C83, D84.

[*]Daniel Sgroi (corresponding author) contact details: University of Warwick, Department of Economics, Gibbet Hill Road, Coventry UK CV4 7AL, UK. Email: daniel.sgroi@warwick.ac.uk. Telephone: +44 (0)2476 575557. Fax: +44 (0)2476 523032. Eugenio Proto contact details: University of Warwick, Department of Economics, Gibbet Hill Road, Coventry UK CV4 7AL, UK. Email: e.proto@warwick.ac.uk. The authors thank several coauthors and colleagues for discussions on related research, especially Robert Akerlof, Gordon Brown, David Gill, Peter Hammond, Graham Loomes, Heracles Polemarchakis, Andrew Oswald, Francesco Squintani and John Stovall. Funding was provided by the Leverhulme Trust. We are grateful to Alex Dobson, Stephen Lovelady, Andrew Taylor and Neil Gatti for expert research assistance.

"Let them eat cake." (Commonly attributed to) Marie Antoinette (1755 – 1793), Archduchess of Austria and Queen of France.

1 Introduction

Where do beliefs come from? Should everyone share the same prior beliefs? Should these beliefs be on average correct? Early work attempting to address these questions linked heterogeneous priors with information processing errors and individual bounded rationality.[1] Morris (1995) provides a survey of the considerable literature which followed. He argues that in order to avoid using subjective heterogeneous priors to justify any result *ex post* given the sensitivity of many economic models to assumptions over beliefs, it is important to look for the roots of differences in beliefs to provide more discipline for the use of heterogeneity in prior beliefs.[2] The aim of this paper is to consider instances when individuals are themselves highly heterogeneous (through different past choices or differences in personal characteristics) and to use this innate heterogeneity to investigate empirically the existence of a general rule governing subjective beliefs, without ruling out the possibility that such beliefs may be biased away from the truth.

We perform an incentivized experiment and consider three types of beliefs: beliefs concerning behavior (relating to mobile phone purchasing decisions, and hypothetical restaurant choices); beliefs concerning attitudes (specifically relating to happiness and political stance); and beliefs relating to observable characteristics (height and weight). In all but the hypothetical restaurant choice problem, the beliefs are formed through experiences in the real world, outside of the laboratory. As a result we have a measure of external validity for our findings which we fortified by designing the questions in our experiment to be more natural than those which make sense only in the context of an entirely laboratory-based experiment. Our questions are generally of two types. The first type of questions essentially ask what percentage of peers a subject thinks lie below him or her in a cumulative distribution. By asking this in an extremely simple way (for example "What percentage of students at Warwick of your gender do you think are shorter than you?") we avoid the need to explain cumulative distributions or require any knowledge of probability. In the second type of questions we ask our subjects to

[1] For example, Geanakoplos (1989).

[2] As Morris puts it "We should resort to unmodelled heterogeneities of prior beliefs only when we can imagine an origin for the differences in beliefs and we can perform comparative static exercises, comparing the predictions of heterogeneous prior models with alternative explanations."

estimate the averages for the characteristic, attitude or behavior from among their peers (for example "What do you think is the average weight for a Warwick student of your gender?") and again the focus is on keeping the language natural and simple.

What we find is a powerful and ubiquitous bias in perceptions that is "self-centered" in the sense that an individual's beliefs about the population distribution changes with their own position in that distribution. In particular, those at the extremes tend to perceive themselves as closer to the middle of the distribution than is the case. For example those who are left-wing see the population as more left-wing and feel themselves to be more typical, and those who own a particular type of mobile phone are likely to misperceive their own brand as more popular than is the case.

Using our data we characterize a distribution of "self-centered beliefs". We show that perceived cumulative distributions on average exceed the true cumulative distributions for individuals with lower values and are smaller for individuals with higher values, while individuals with values closer to the mean tend hold more accurate perceptions about the true cumulative distributions. In this way the average perceived distribution second order statistically dominates the real distribution. Accordingly, taller and heavier individuals think that there are more tall and heavy individuals in the population, individuals in the political fringes perceive themselves more representative, as do those who are very happy or sad. Individuals with less popular mobile phone brands think that there are more individuals using the same brand, but the ones with popular mobile brands correctly estimate its distribution. Finally, when asked to decide on a tie-breaking rule, in the hypothetical restaurant choice problem those who choose in a given way tend to see their choice as being more popular than is the case.[3] Put simply, individuals tend to see themselves as more "average" than is the case.

Consistent with this notion of self-centered beliefs, we also find that errors made by individuals when asked to estimate averages are a function of their own values. For example, a heavier male believes the population average weight to be greater than his lighter counterparts.

In terms of our findings, the closest related work in the psychology literature is the so called "false consensus bias" (Ross, Greene and House (1977)). In this experimental literature individuals are generally asked to indicate their attitude or behavior on a dichotomous measure (yes or no, agree or disagree) often to a matter of opinion. They are then asked to estimate the percentage of their peers who would respond one way

[3]For example, an individual who opts for a more deferential tie-breaking rule sees others as similarly deferential.

or the other. Individuals are said to perceive a false consensus when their estimate of consensus for their own position exceeds the estimate for it made by those who endorsed the opposite position.[4] On the contrary, in the current paper we are concerned with matters of fact, and hence we investigate how beliefs differ with respect to the true state of the world and not with respect to the opinion of others. The difference is fundamental for this paper since we are concerned with the accuracy of beliefs with respect to the truth (or departures from rational expectations).[5] Another concept usually associated with false consensus is "assumed similarity" (Cronbach, 1955). This is generally measured as the absolute difference between the position attributed to oneself and that ascribed to a benchmark individual known to everyone. An assumed similarity score is generally computed as the absolute difference between the position attributed to oneself and that ascribed to the target, and generally each individual tends to position target closer to themselves. Again, these procedures do not provide information on the accuracy of perception, the aim of the current paper, since they are instead interested in comparing individuals' relative perceptions.[6] Therefore, to the best of our knowledge, there are no contributions aimed at determining empirically a general rule on how subjective beliefs are related to true values.

There are many possible routes through which the self-centered beliefs and the resulting heterogeneous priors which we see in our results might emerge. In particular we identify a number of plausible causal factors from assortative matching or homophily, self-deception and self-serving biases, the availability heuristic, costly learning to self-projection. We will examine these in more detail in section 4 but we will give two possible routes here. Consider assortative matching (Becker, 1973) or homophily (McPherson,

[4]See page 612 in the Handbook of Experimental Economics (Kagel and Roth, 1995) for more on false consensus, including an example.

[5]In one seminal paper on false consensus, Marks and Miller (1987) note: "...The false consensus hypothesis has no direct bearing on whether subjects will overestimate, underestimate, or accurately estimate the actual consensus for their own behavior." For example, consider the view of a right-wing individual who is asked what percentage of people are right-wing. She might say 60%. A left-wing person who is asked the same questions might say 50%. The"false consensus bias" is concerned with the difference between 60% and 50%, while our concern is the difference between 60% and the true percentage of right-wingers and how this difference between beliefs and the truth changes as the type of individual changes. Notice that we can derive one observation from false consensus which closely links to our work: if individuals do differ in their opinions then it is immediate that at least one person must have an opinion that is at odds with the truth. However, this does not help us to see which individual is more likely have such a view.

[6]Also related to false consensus is "projection bias" in the Freudian tradition (the argument that individuals see others as reflections of themselves and so find it hard to envisage views that are too different from their own).

Smith-Lovin and Cook, 2001; Golub and Jackson, 2011) through which people may associate with those who are similar to themselves. This introduces a simple bias: when individuals attempt to learn about the true population by sampling those around them they will see characteristics and choices that are too similar to their own to be truly representative. Combined with the availability heuristic (Tversky and Kahneman, 1973), according to which individuals put *undue* weight on easily available data, they may then draw biased self-centered inferences. The failure to realize that inferences are biased by a non-representative sample can be categorized as irrational; however, Carrillo and Mariotti (2000) and Benabou and Tirole (2002) argue forcefully that such "strategic ignorance" or "self-serving biases" can be, and often are, rational. Hence, it may be utility-maximizing to self-delude. Finally we also consider the possibility that salience may be present in our results. To explain, it may be that those at the extremes are more salient to everyone, and so we would expect to see perceived distributions with fatter tails than the true distributions. We also discuss this in section 4 where we use other data from our experiment to show that while there is some evidence of salience for some variables it is clearly not the driving factor behind the ubiquitous bias we see throughout our results.

In the next section we summarize the experimental design and the key variables. Section 3 examines the experimental results both in terms of the perceived cumulative distributions and the errors made by different individuals when asked to estimate averages. Section 4 discusses possible explanations for our results and section 5 concludes. Tables and figures are provided in the Appendix, and the experimental instructions are reproduced in the Supplementary Information.

2 Experimental Design and Key Variables

Our data was collected using a series of computerized tasks and questions presented in a controlled experiment at the University of Warwick. The text of the questionnaires and the accompanying instructions are provided in the Supplementary Information. The experiment was conducted in a laboratory, however many of the choices and characteristics are drawn from the real-world experiences of our participants as well as information given to them in the laboratory. The participants were 154 students drawn from the university-wide experimental pool of over 1500 subjects.[7] The experiment took place in

[7]Ensuring a varied pool of experimental participants selected on as randomized a basis as possible was especially important for this experiment. The University of Warwick keeps a register of those

19 sessions with about 8 students per session, and was conducted on 27 May, 30 May and 29 June 2011.[8]

Subjects were given a £2.50 show-up fee, plus a bonus of £5 pounds if a randomly drawn answer was within 10% of the correct answer in rounds 2 to 6 as described below. For example, if participants were asked to state the average height of the student body in Warwick and this was the randomly allocated bonus question they received a £5 bonus if and only if their answer was within 10% of the true average.[9] The payment scheme was fully transparent to all participants and highlighted in the instructions during the experiment. No participant was allowed to participate more than once.[10] The experiment itself typically lasted 20 minutes and the average payment was a little over £5, producing an hourly rate of around 25 US dollars. The experimental time-line is summarized below but a full transcript of the instructions and tasks faced by the participants is provided in the Supplementary Information.

2.1 Experimental Time-line

Participants arrive at the laboratory, are registered and taken to a screened computer terminal where they receive on-screen instructions as detailed in the Supplementary Information. The payment scheme is explained and they are informed that everything they do is fully anonymous. To that end they receive randomly generated usernames and passwords to use as logins for the terminals. After entering usernames and passwords the experiment proper begins in round 1 by asking participants to report their gender (1a), height (1b), weight (1c), happiness (1d), political beliefs (1e), and current brand of mobile phone (1f). They are also given a hypothetical restaurant choice as follows: "Imagine that you have to decide between two restaurants in which to have dinner alone.

available for use as experimental participants and a research assistant (rather than the experimenters) drew from this large pool of potential applicants on a random basis. Participants were recruited without any knowledge of the nature of the experiment and so could not self-select into or out of the experiment. The times and dates of the sessions were varied to avoid discriminating against participants from any demographic, and in the event we had large variety in terms of subject, year-group and gender.

[8] There was also an earlier non-incentivized pilot experiment which consisted of 120 participants drawn from the same experimental pool, held on 17 March, 5 May and 11 May, 2010. The main results for this paper will be drawn from the fully-incentivized experiment, though the data from the pilot study will be used when calculating the average height, weight, happiness and political stance of the Warwick student body.

[9] For our purposes the true average was based on the numbers generated within this experiment and and from the earlier non-incentivized pilot experiment. For one question, denoted (5e) below the scheme was changed slightly as participants had to select from an interval and so they were told that an answer in the correct interval or the one to either side would be sufficient to win the prize.

[10] Participation in the pilot experiment also ruled out participation in the full experiment.

They are called restaurant A and B. You have some private information that A is better, but you know that an equally well-informed colleague has information suggesting that B is better. Would you choose to eat at A, B or are you indifferent?" (1g).

For (1b) and (1c) they were allowed to enter their heights and weights in metric or imperial measurements as they wished and this freedom was maintained in all rounds. For (1d) participants are instructed to use a 7-point Likert scale as follows: "Please use a 7-point scale where 1 is completely sad, 2 is very sad, 3 is fairly sad, 4 is neither happy nor sad, 5 is fairly happy, 6 is very happy and 7 is completely happy." For (1e) they are instructed to use a similar scale: "Please use a 7-point scale where 1 is far left, 2 is left, 3 is centre left, 4 is centre, 5 is centre right, 6 is right and 7 is far right." (1g) might be of special interest to those interested in rational herding and informational cascade literature (see Banerjee, 1992, and Bikhchandani, Hirshleifer and Welch, 1992) and essentially asks what participants would do in a situation of theoretical indifference when processing information. There is no clear right or wrong answer to (1g) though "A" points to a measure of confidence in the participant's private signal over that of their colleague, whereas "B" perhaps implies a measure of deference towards others (or a lack of confidence in the individual's own signal). Once they answered these questions and hit a "submit answers" button they were taken to the round 2 questions.

At this point they were informed that one question would be chosen at random in round 2 or the later rounds (rounds 3, 4, 5 and 6) as a "prize question" for which they would receive a bonus payment of £5 if their answer was within 10% of the correct answer. They were then asked to report the percentage of students at Warwick they thought were less happy than they were (2a), less right-wing (2b), shorter (2c) and lighter (2d). For (2c) and (2d) they were asked to consider only their own gender. In (2e) they were asked to consider the mobile phone brand listed in round 1 and asked what percentage of students at Warwick they thought also used the same brand of mobile phone as their main mobile phone. For (2f) they were asked to consider the hypothetical restaurant choice and report what percentage of their fellow Warwick students they thought chose the same answer that they did (they were reminded of the entirety of the question and the possible answers).

In round 3 they were asked to report the average height for someone in the 10% tallest Warwick students of their gender (3a), the average weight for someone who is in the 10% heaviest Warwick students of their gender (3b), the average happiness for someone who is in the 10% happiest students at Warwick (3c), and the average political belief for someone who is in the 10% most right-wing students at Warwick (3d). For (3c) and (3d) they were asked to use a 7-point Likert scale as before.

Round 4 was phrased identically to round 3 except that in each case in the four questions they were asked to report the average for the 10% shortest (4a), 10% lightest (4b), 10% most sad (4c) and 10% most left-wing (4d), again for the population of students at Warwick, using a 7-point Likert scale for (4c) and (4d), and considering only their own gender for (4a) and (4b).

Round 5 focused on overall averages rather than extremes in the distribution. They were asked to report the average height (5a) and weight (5b) for a Warwick student of their gender, and using a 7-point Likert scale the average happiness (5c) and political belief (5d) for a Warwick student. In question (5e) they were asked to estimate the percentage of their fellow Warwick students who used each of a selection of mobile phone brands. They were presented with a tabulated list of the most popular brands in the UK, and they were informed that the list was presented in alphabetical order (except for the "other" category which was presented last). They were asked to include an entry for every brand (including "other"). For (5f) they were asked again about the hypothetical restaurant choice: "Think again about the restaurant question you were asked earlier in the session. To remind you, you had to decide between two restaurants in which to have dinner alone. They were called restaurant A and B. You had some private information that A is better, but you knew that an equally well-informed colleague had information suggesting that B was better. What percentage of your fellow Warwick students do you think would have chosen to eat at restaurant A if they were asked the same question? Remember that the other options were indifferent and B."

For round 6 the participants were asked to answer a single question designed as a check on their ability to understand and manipulate expectations and probability: "Consider the following gamble. You have a 20% chance of winning £100, a 40% chance of winning £10 and a 40% chance of winning £0. If you played this gamble many times what would you expect to be your average winnings per gamble? (in pounds)"

Round 7 is a final questionnaire and, as is conventional, was not incentivized (participants were informed that the incentivized part of the experiment had ended) since there was no way of checking right or wrong answers. They were asked to report their age (7a), nationality (7b), degree subject (7c), whether they studied mathematics up to their final year at school (7d) and also comment on their methods, if any, during the incentivized parts of the experiment (7e).

Once each round was completed participants could not go back and change earlier answers, nor did they know the content of later rounds upon entering answers to earlier rounds. This was important as it prevented any attempt to retroactively alter their answers to make winning the bonus payment easier.

3 Results

Before moving on to the analysis we present the main variables in table 1 below. The variable "Happiness" is coded from completely sad (1) to completely happy (7) and taken from the answers to question (1d), and "Political Stance", from extreme left (1) to extreme right (7), taken from question (1e). "Weight" is converted to kg from the answers in question (1c) and "Height" to cm from the answers in question (1b).

In this section we first present an analysis of individuals' beliefs about their own positions in the true distribution, starting first with a model. Thereafter we analyze estimated averages. In both cases our findings indicate that beliefs about position in the distribution and about averages are a function of individual's own position in the distribution.

3.1 Modeling Beliefs

Let Θ be the set of characteristics, attitudes and choices considered in this paper, for example height, reported happiness or the market share of a brand of mobile phone. $\theta \in \Theta$ is a particular characteristic, attitude or choice, for example, height. We index the value of each characteristic for each individual i, so θ_i might be individual i's height. Let $F(\theta)$ be the cumulative distribution of characteristic θ. In a mild abuse of notation, denote $F(\theta_i)$ as individual i's true position in the cumulative distribution and then $E_i(F(\theta_i))$ is i's belief about his position in the cumulative distribution.[11] Given the nature of incentives a typical subject i solves the following programme:

$$Min_{E_i(F(\theta_i))} \{E_i(F(\theta_i)) - F(\theta_i)\} \tag{1}$$

We assume that individuals form their expectations by drawing from a sample M_i from the total population N. However we cannot rule out the possibility that the sample of size M_i may be biased. In particular is may be that the sample is taken from members

[11]We preferred to ask "less that", rather than "at least as" because we judged it a more natural question that individuals are likely to have faced in their everyday life. We calculated the CDFs using the same definition.

of the population who are similar to the individual i. An unbiased estimator would be

$$E_i\left(F\left(\theta_i\right)\right) = F\left(\theta_i\right) + error_i \tag{2}$$

where $error_i$ is i.i.d., with mean 0 across the population. In order to asses the existence of a systematic bias, we will therefore consider the model:

$$E_i\left(F\left(\theta_i\right)\right) = G(F\left(\theta_i\right)) + \sigma(\theta_i)\epsilon_i, \tag{3}$$

Where $\sigma(\theta_i)$ is a general function to allow for heteroscedasticity with respect to θ_i, and ϵ_i is a white noise error.

From the data, we can observe the beliefs $E_i\left(F\left(\theta_i\right)\right)$, the real distribution $F\left(\theta_i\right)$) hence we estimate $G(F\left(\theta_i\right))$ both non parametrically using local polynomial smoothing[12] and by making parametric assumptions on this functional form. This way we can directly examine the difference between $G(F\left(\theta_i\right))$ and the true distribution, $F\left(\theta_i\right)$ which represents the bias made by each individual, allowing for non-linearities in beliefs as we change the own values of each individual.

In the following subsection we present the empirical evidence with respect to the characteristics weight, height, happiness and political stance. We will also analyze the beliefs about mobile distribution and deference in information processing. These last are slightly different because $F\left(\theta_i\right)$ represent frequencies and not CDFs.

3.2 Empirical Evidence on Perceived Distributions

We present the CDF of 4 characteristics (weight, height, happiness and political stance) in figures 1 and 2.[13] Frequencies of answers to the restaurant choice question (1g) and the mobile choice question (1f) can be read from the histogram in figure 3.

We start by analyzing weight, height, happiness and political stance. From the right panels of figures 1 and 2. We can observe the real CDF, $F\left(\theta_i\right)$ and the scatterplot of the perceived positions, $G(F\left(\theta_i\right))$ with their locally weighted scatter-plot smoothing

[12] Kernel-Weighted Local polynomial smoothing involves fitting the response to a polynomial form of the regressor via locally weighted least squares. In the Kernel-weighted regression, $G(F\left(\theta_n\right))$ is calculated without assuming a functional form, as a constant term of a regression weighted by the kernel function of $E_i\left(F\left(\theta_i\right)\right)$ on the polynomial terms $G(F\left(\theta_n\right)) - G(F\left(\theta_i\right))$, $(G(F\left(\theta_n\right)) - G(F\left(\theta_i\right)))^2, ..., (G(F\left(\theta_n\right)) - G(F\left(\theta_i\right)))^p$, for each point $G(F\left(\theta_n\right))$. The definitive reference is Fan and Gijbels (1996), see also the Stata 12 base reference manual, p. 1001.

[13] When calculating the population averages we also used the 120 data from the pilot experiment to increase the size of the sample to 274.

or "Lowess".[14] The scatters in the right panels of figure 1 represent the real and the perceived CDFs in θ_i for each subject i, θ_i for height and weight, and in figure 2 the same observations for political stance and happiness. The dashed line represents the estimated function $G(F(\theta_i))$ with the 95% confidence interval; In order to compare $G(F(\theta_i))$ and $F(\theta_i)$, we also plot the 45 percent degree line. There is a clear pattern: individuals at the extreme tend to overestimate the number of individuals who are equal to or more extreme than themselves, while those at the center of the distribution seem better informed. This is true for happiness and political stance, female and male weight, and male height. For political stance and happiness we note that the two extreme numbers (1 and 7) seem to revert to the 45 degree line. This makes sense since these two characteristics are bounded above and below, therefore extreme individuals will be able to calculate more accurately their position (and have less scope for self-deception). This effect appears even more clearly in the parametric analysis to follow.

In the first 4 columns of table 2, where we estimated $G(F(\theta_i))$ parametrically for happiness, political stance, height and weight, we find comparable results. In this parametric analysis, we assume a linear model for height and weight, so that a positive and significant constant term would signal an overestimation of the shorter and lighter individuals and a coefficient less than 1 would signal a overestimation of the taller and heavier individuals. Given the bounded nature of the 7-point metric for Happiness and Political stance, we want to account for the fact that the least happy (most happy) and the most left-wing (right wing) can more accurately estimate their position in the CDFs.[15]

For female and male weight (column 3) and male height (column 4) we find coefficients less than 1 ($p - value < 0.01$) and a positive intercept. The coefficient of female weight is not significantly different from 0, a small and unique discrepancy between the non- parametric and parametric analysis. This indicates that females towards the lowest extremes are the most accurate in their perceptions about this characteristic. For hap-

[14] This displays for each value of the independent variable, θ_i a smoothed value of the dependent variable, $G(F(\theta_i))$. The smoothed values are obtained by running a linear regression using only the data $(x_i; y_i)$ and a subset of the data with the x values close to x_i. Data are weighted so that the central point $(x_i; y_i)$ receives the highest weight and points that are farther away receive smaller weight. The estimated regression line is then used to predict the smoothed value for y_i only; a separate weighted regression is performed for every point in the data.

[15] Note that there is a slight technical difference between those at the very top and bottom of the 7-point scale. Anyone who reports "1" is certain that there is no one below them in the CDF, while anyone reporting "7" is sure to be at the top but needs to consider how many others might report "7", when considering those who are below in the CDF.

piness and political stance the interpolating line is inverse s-shaped, with the center and the two extremes close to the 45 degree line. This makes sense as the characteristic is bounded so they can quite accurately calculate (perfectly for the lower extreme) $F(\theta_i)$.[16]

Mobile ownership and attitudes in the hypothetical restaurant choice are not ordered variables, it is therefore impossible to determine a perceived CDF and a real CDF. We therefore collected data on beliefs about frequencies and compared these with the true frequencies. Accordingly, we define our $F(\theta_i)$, as the true frequency of the choice θ_i and $G(F(\theta_i))$, the beliefs' function. In the left panels figure 3 we can compare the histogram representing the distribution on mobile phones and attitudes toward in the hypothetical restaurant choice question, $F(\theta_i)$, with the subjective beliefs $G(F(\theta_i))$, where the different θ_i have been ordered increasing with their frequencies. Although we are comparing now frequencies rather than CDFs, we note a similar pattern to the one emphasized in figures 1 and 2. The Lowess tends to stay above the real frequencies for less popular brands and for less frequent tendencies in the restaurant choice question.

As before we estimate the model 3, and present the estimated $G(F(\theta_i))$, in the right panels of figure 3. The reading of the panels is slightly different from before, with the average subjective beliefs above the real level for less frequent θ and are non-significantly different for more frequent θ. The interpretation of the results are the same as with happiness, political stance, height and weight: individuals with more common brands or with the most common attitude in the restaurant question have beliefs that are on average correct, on the contrary individuals making less common choices overestimate the frequency of their choices.

In the last two columns of table 2, we estimate parametrically $G(F(\theta_i))$ for the mobile phone distribution question (2e) and attitudes towards information processing from the restaurant question (2f). This confirms the main finding emphasized by figure 3.[17]

In essence subjects more at the extremes of the distributions think there are more individuals in the same position (or indeed in a more extreme position) than themselves, while more average subjects tend to have more correct beliefs. So, a very tall person really does perceive the world as taller. Perhaps more economically significant, someone

[16] Apart from the reason given in footnote 15 we might also argue that those at the very extreme points, for example extreme right-wingers, may have reason to wish to see themselves as different from the average and so may have more accurate beliefs. We come back to this in section 4.

[17] We also note a slight tendency for the owners of the very smallest market share brands to better understand their own-brand market-share as compared with those with slightly higher market share brands. This could be due to the fact that the owners of these least popular brands were looking for niche products with specific features and so spent more time and effort on research or may see themselves as different. We return to this in section 4.

who has purchased a less popular mobile phone believes it to be more popular than is the case. Note that this bias is true for different frames. Either when subject are asked about their position and when their asked about the frequency of their choices.

One remarkable point that stands out is the ubiquitous nature of this bias: only for females who are among the shortest and only in the parametric analysis is there any deviation from the simple rule that those at the extremes do not see themselves as being as extreme. We argue in section 4 that this possible exception may even help us to better understand the underlying cause of the bias.

More formally, we find that individuals with a θ_i closer to the average tend to estimate better their position than those with a θ_i closer to the extremes of the distribution. Denote E_i as the expectation operator for individual i. We can then express this as a simple conjecture: for each $\theta \in \Theta$, $\exists \theta^*$ s.t. $\theta_i < \theta^* \Rightarrow E_i\left(F\left(\theta_i\right)\right) > F\left(\theta_i\right)$ and $\theta_i > \theta^* \Rightarrow E_i\left(F\left(\theta_i\right)\right) < F\left(\theta_i\right)$. Put simply, we can consider a point in the cumulative distribution θ^* such that individual's with value $\theta_i < \theta^*$ believe they are positioned higher in the distribution than is the case, and for $\theta_i > \theta^*$ they believe they are lower. More compactly, $E_i\left(F\left(\theta_i\right)\right)$ stochastically dominates in the second order sense $F\left(\theta_i\right)$. This is a form of mean-reversion of beliefs. Since beliefs seem too heavily dependent on θ_i we label them "self-centered".

3.3 Empirical Evidence on Estimated Averages

In rounds 3, 4 and 5 we asked subjects to report their estimates about averages. Here we analyze their responses focussing in particular on the answers to questions (5a), (5b), (5c) and (5d) in which they were asked to estimate the average height, weight, happiness and political stance of their peers (controlled for gender).[18]

Since we are interested in the accuracy of beliefs with respect to the level of θ_i, we will analyze how errors are significantly affected by an individual's own values. Given our questions and the nature of incentives, individuals need to give an accurate number, facing a problem similar to the one presented in previous subsection, and in order to test for the existence of a bias that depends on their own-characteristic or own-attitude θ_i, we will consider the following expression:

$$E_i\left(\overline{\theta}\right) - \overline{\theta} = f\left(\theta_i\right) + error_i \tag{4}$$

[18]Given their non-ordered nature, the questions relating to mobile and restaurant choices are not amenable to the analysis carried out in this subsection.

Where $\overline{\theta}$ is the average for the characteristic or attitude being considered, and $E_i\left(\overline{\theta}\right)-\overline{\theta}$ therefore captures the error with respect to the truth made by individual i. $f\left(\theta_i\right)$ is a function of the individual's own value and we will explicitly consider linear and quadratic functional-forms.

We estimate $f\left(\theta_i\right)$ parametrically, using an OLS method with robust standard errors. Table 4 summarizes our findings. Notice that we see high levels of statistical significance across all four of the characteristics and attitudes we examine, indicating that own values are indeed important when individual's form estimates of averages and in particular indicating that an individual's errors with respect to the true averages, are a function of an individual's own values. We can express this result quite compactly as a simple conjecture: for each $\theta \in \Theta$, $E_i\left(\overline{\theta}\right)-\overline{\theta}$ depends upon θ_i.

Furthermore, from table 4, we note that the bias $f(\theta_i)$ is increasing in happiness, male and female weight and –with a slightly lower confidence– in male height. The way the biases on averages are related to θ is perfectly consistent with the bias highlighted in the previous subsection concerning distributions. If an individual with a low θ_i thinks that there are more individuals with even smaller values of θ_i it is natural that he believes the average of the θ_i distribution is lower than an individual with an high θ_i, who believes that there are more individuals with values even larger than his own θ_i.[19]

In conclusion, once again we have uncovered evidence of a bias, and one which can be labeled as "self-centered", since the bias we see is a function of something that we might think should not matter: the individual's own value.

4 Discussion: What Generates Self-centered Beliefs?

There are many possible routes through which self-centered beliefs might emerge. Assortative matching (following Becker, 1973) and the related phenomenon, homophily, might result in people associating with those who are similar to themselves giving rise

[19] An interesting puzzle is raised by the existence of possible hump shaped pattern for averages on political stance and female height. The peak for political stance is 4.6 and for female height is 161 cm, both significantly smaller than their respective upper bounds. This effect may be due to the fact that extremists by their very nature may have sought more (accurate) information about the mean individual, so for example a right-wing extremist student at Warwick may be more active in canvassing political views from others, and may therefore have a greater awareness of the average political stance in his locality (in this case the University of warwick), and females who are significantly shorter than average may have attempted to discover just how far from average they may be.

to a feeling of being more average than is the case.[20] The bias here is two-fold, first by associating with those who are similar, your beliefs may reflect the biased sample of the population chosen to be like yourself, but second, individuals must remain unaware of their own biased sampling process if it is to cloud their beliefs. This failure to correct for an apparently biased sampling process may be seen as simple evidence of irrationality. However, Carrillo and Mariotti (2000) and Benabou and Tirole (2002) provide numerous mechanisms (including self-delusion and memory manipulation) to foster rational ignorance of the truth (indeed Carrillo and Mariotti call this type of behavior "strategic ignorance"). To help differentiate between rational and irrational interpretations let us consider an interesting observation that comes through the analysis of height. In our results we find that shorter than average females seem to be more aware of the true distribution than do shorter than average males. This observation seems to have little to do with irrationality, but may well be consistent with Carrillo and Mariotti or Benabou and Tirole-style self-deception, as it is possible that shorter females may *want* to perceive themselves as *petite*, while males would rather think of themselves as above average. Also consistent with this argument is the fact that both males and females who are just below average weight are the group most likely to have accurate perceptions.

The fact that the extremes in the distribution suffer the worst from the bias lends some support to the *availability heuristic* explanation identified by Tversky and Kahneman (1973), according to which individuals put *undue* weight on easily available data, and may then draw biased inferences.[21]

Notice also that extreme characteristics (for instance, the very tall or extreme left wing) are generally thought to be more salient than those with average values, which could provide another possible bias. We can reasonably rule out this possibility for mobile brands, in this case there is no reason why the extreme, i.e. the owners of a fringe brand, should be more salient than the others, and there is no evidence to support salience in this case. However, for Political Stance, happiness, weight and height this is a real possibility. To test this we collected data on perceived averages for the potentially most salient (the top and bottom 10%) in rounds 3 and 4 of the (incentivized) experiment as described in section 2. For example, we asked subjects in question (2b) to report

[20] Homophily has has been studied principally within sociology (see for instance, McPherson, Smith-Lovin and Cook, 2001) but has become a recent source of interest to economists interested in social networks (see for example Golub and Jackson, 2011 and Currarino et al., 2009).

[21] Linked to this, is the idea that observing others is one of the cheapest ways to acquire information and is the key form of learning in the social learning and herding literatures (see Banerjee, 1992, and Bikhchandani, Hirshleifer and Welch, 1992).

their average for the top 10% heaviest individuals (of their gender). In tables 5 and 6 we regress the error between the data arising from these questions and the true values against the true values and the perceived average over the total population,. The later acts as a control for the presence of a general bias for the entire distribution, and not just the the top or bottom 10% of the distribution. A positive coefficient for the perceived top 10% average in each regression would support the presence of salience for the top 10%, i.e. individuals systematically overestimate the height of the tallest 10% in the population, the weight of the 10% heaviest, the political stance of the 10% most right-wing and the happiness of those in the top 10% of the happiness scale. In the same vein a negative coefficient for the perceived bottom 10% average would signal salience for the bottom 10%, i.e. individuals systematically under-estimate the height, weight and the level of happiness of the bottom 10%, and the political stance of the 10% most left wing. From tables 5 and 6 we note that salience is a possibility in some contexts but not others. It may exist for those in the bottom 10% of the happiness scale, the shortest 10% and, with a lower level of confidence, for the 10% most left-wing and 10% most right-wing, but not for the remaining candidates. It is therefore at best a partial explanation.

Closely related to salience is the evidence that subjective probabilities are inverse S-shaped with respect to true probabilities, or more simply that individuals overweight small probabilities and underweight large probabilities.[22] Interestingly enough, this might provide evidence that individuals perceive the distributions like probabilities, therefore we could re-interpret the questions we ask in round 2 slightly differently. For example consider being asked what percentage of people might be shorter than yourself. This could be re-interpreted as asking for the probability that a random individual is shorter, and so would be subject to the inverse S-shape relationship. We would then expect to see lower probability events (for instance the chance that someone who is very short being drawn) being overweighted. This would imply a cumulative distribution function biased in a similar way to those we derive from our data and would provide grounds for an interesting characterization of subjective beliefs, but it is not consistent with our findings when we consider beliefs over averages rather than distributions. In particular the overweighting of small probabilities with respect to true probabilities would not be "self-centered" in the sense that the bias would be independent of individual's own values: this would be a general bias that would look the same for all individuals in the population (much like the salience bias). While this might explain our derived

[22]There is a large literature on this, but a summary can be found in Camerer, Loewenstein and Rabin (2003).

cumulative distribution functions it cannot explain the findings in table 4 which clearly indicate that errors made when estimating averages are highly significantly dependent upon individual's own values. Therefore, the overweighting of small probabilities is at best only a partial explanation of our findings.

Costly learning is another possibility which links in well with self-deception and homophily since we can argue that where falsely believing yourself to be average is useful, it might cost more (in terms of final utility) to learn otherwise.[23] To explain why our subjects fail to learn the truth, we might even appeal to the optimal experimentation literature, beginning with Robbins (1952) and brought to economics by Rothschild (1974). This literature demonstrates that optimizing decisions while simultaneously attempting to learn can result in learning stopping before the truth is known. Since our subjects need to make decisions on a daily basis using the beliefs they form, this process of decision-making and learning in a simultaneous setting may be appropriate. However, this does not explain why the subjects towards the middle of the distribution in our model have more accurate beliefs so this is another explanation that is at best partial.

Another possibility is that our subjects suffer from an overconfident belief that they are more typical than is the case.[24] Burks et al. (2010) provide an interesting twist, which might fit our findings, and analyze the theory that overconfidence is induced by the desire to send positive signals to others about oneself.[25] However, we might wonder why being "average" is seen as an important target for overconfidence and how this overconfidence might be sustained relative to the truth, which might return us to the topic of "self-deception".

5 Conclusions

Our results highlight a subtle form of self-deception: those at the extremes seem to perceive themselves as much closer to average than is the case. These innate desires and feelings may have little to do with the true distribution and so these individuals may need to find a way of self-deceiving themselves into biasing their beliefs. This may come through assortative matching or homophily and a failure to make the appropriate

[23] Coate and Loury (1993) and Farmer and Terrell (1996) look at costly learning in the context of discrimination.

[24] See Benoit and Dubra (2001) for more on overconfidence.

[25] A recent empirical paper related in terms of methodology to ours is Oswald (2008), based on individuals' perceived and real height and showing that individuals' reporting function from reality to feelings is concave.

correction to biased sampling, or through a variety of other mechanisms, and this may even be utility-maximizing. Our second consistent finding is that estimates concerning averages tend to be a function of individual's own values. The biases we uncover are both powerful and ubiquitous, they apply across a variety of different characteristics, choices and parameters, and apply regardless of the observability of these characteristics.

Our primary contribution is in examining the link between subjective beliefs and the truth, and finding and characterizing the bias we observe. Nevertheless our design allows us to meaningfully compare several possible theories which might give rise to the biases that arise. While we find that several theories can partially explain our data (such as salience or the underweighting of low probability events) a combination of the availability heuristic, assortative matching and homophily combined with a form of strategic ignorance or potentially rational self-deception can successfully explain our findings. We would not wish to rule out other possibilities at this stage and a further analysis of the underlying causes of the biases we observe would represent the obvious next step for future research in this area.

Putting aside the origins and focusing on the implications, our findings lend tentative support to the heterogeneity of subjective priors and provide a guide about when heterogeneity makes sense (we would echo Carrillo and Mariotti (2000) and Benabou and Tirole (2002) in arguing that where it is utility-improving for individuals to have different beliefs they might well arise) and we would point to the ability to self-select peers as an important mechanism for the persistence of biased sampling. Moreover, it has been suggested (e.g. Manski, 2004) that in order to perfectly identify a model it is necessary to have data on expectations and beliefs. This paper, on the one hand we demonstrate the existence of a systematic bias in beliefs. On the other hand we also show that the bias is predictable and therefore it might be possible to make adjustments to correct for such a bias at the level of policy of model-building. Finally, we believe we have highlighted several avenues for future research including more work on understanding the underlying causes of the biases we identify and incorporating the ubiquitous bias which we identify into economic theory.

References

[1] Banerjee, A. V. (1992). "A simple model of herd behavior", *Quarterly Journal of Economics* 107: 797-817.

[2] Becker, G. S. (1973). “A theory of marriage: part I”, *Journal of Political Economy* 81: 813-46.

[3] Benabou, R., and Tirole, J. (2002). “Self-confidence and personal motivation.” *Quarterly Journal of Economics* 117: 871-915.

[4] Benoit, J-P, and Dubra, J. (2011). “Apparent overconfidence.” *Econometrica* 79: 15911625.

[5] Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). “A theory of fads, fashion, custom and cultural change as information cascades”, *Journal of Political Economy* 100: 992-1026.

[6] Burks, S., Carpenter, J., Goette, L., and Rustichini, A. (2010). “Overconfidence is a social signaling bias.” Mimeo, University of Minnesota.

[7] Camerer, C.F., Loewenstain, G., and Rabin, M. (2003). *Advances in Behavioral Economics.* Princeton, New Jersey: Princeton University Press.

[8] Carrillo, J.D., and Mariotti, T. (2000). “Strategic Ignorance as a Self-Disciplining Device”, *Review of Economic Studies* 67: 529-544.

[9] Coate, S., and Loury, G.C. (1993). “Will affirmative-action policies eliminate negative stereotypes?”, *The American Economic Review* 83: 1220-1240.

[10] Cronbach, L.J. (1955). “Processes affecting scores on “understanding of others” and “assumed similarity.” *Psychological Bulletin* 52: 177-193.

[11] Currarino, S., Pin, P., and Jackson, M. (2009) “An economic model of friendship: homophily, minorities and segregation”, *Econometrica* 77: 1003-10045.

[12] Fan, J., and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications.* London: Chapman & Hall.

[13] Farmer, A. and Terrell, D. (1996). “Discrimination, Bayesian updating of employer beliefs and human capital accumulation”, *Economic Inquiry* 34: 204219.

[14] Geanakoplos, J. (1989). “Game theory without partitions, and applications to speculation and consensus”, Cowles Foundation Discussion Paper No.914. Available from http://cowles.econ.yale.edu

[15] Golub, B., and Jackson, M. (2011). "Network structure and the speed of learning: measuring homophily based on its consequences", *Annals of Economics and Statistics*, Forthcoming.

[16] Kagel, J.H., and Roth, A.E. (1995). *The Handbook of Experimental Economics.* Princeton, New Jersey: Princeton University Press.

[17] Manski, C. (2004). "Measuring expectations." *Econometrica* 72, 13291376.

[18] Marks, G. and Miller, N. (1987). "Ten years of research on the false-consensus effect: An empirical and theoretical review." *Psychological Bulletin* 102: 72-90

[19] McPherson, M., Smith-Lovin, L., and Cook, J. (2001). "Birds of a feather: homophily in social networks." *Annual Review of Sociology* 27, 415-44.

[20] Morris, S. (1995). "The common prior assumption in economic theory." *Economics and Philosophy* 11: 227-253.

[21] Oswald, A. (2008). "On the curvature of the reporting function from objective reality to subjective feelings." *Economics Letters* 100, 369-372.

[22] Robbins, H. (1952). "Some aspects of the sequential design of experiments." *Bulletin of the American Mathematical Society* 55, 52735.

[23] Ross, L., Greene, D., and House, P. (1977). "The 'false consensus effect': an egocentric bias in social Perception and Attribution Processes." *Journal of Experimental Social Psychology* 13, 279-301.

[24] Rothschild, M. (1974). "A Two-Armed Bandit Theory of Market Pricing." *Journal of Economic Theory* 9, 185202.

[25] Tversky, A., and Kahneman, D. (1973). "Availability: a heuristic for judging frequency and probability." *Cognitive Psychology* 5, 207-232.

Tables and Figures

Table 1: Main Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
Weight Female	55.71	7.007	43	73	69
Weight Male	72.427	12.481	45	109	82
Height Female	162.418	7.237	138	181	67
Height Male	179.613	9.858	151	208	80
Political Stance	3.805	1.166	1	7	154
Happiness	4.766	1.021	1	7	154
Male	0.545	0.5	0	1	154

Figure 1: **Weight and Height:** The left panels represent the real CDFs (continuous lines) and the beliefs over the distributions (dots) and their respective Lowess function (dashed lines). The right panels represents the 45 degree lines (solid lines) and the local polynomial interpolation of the perceived and the real distributions (dashed lines), the shadow represents the 95% confidence interval.

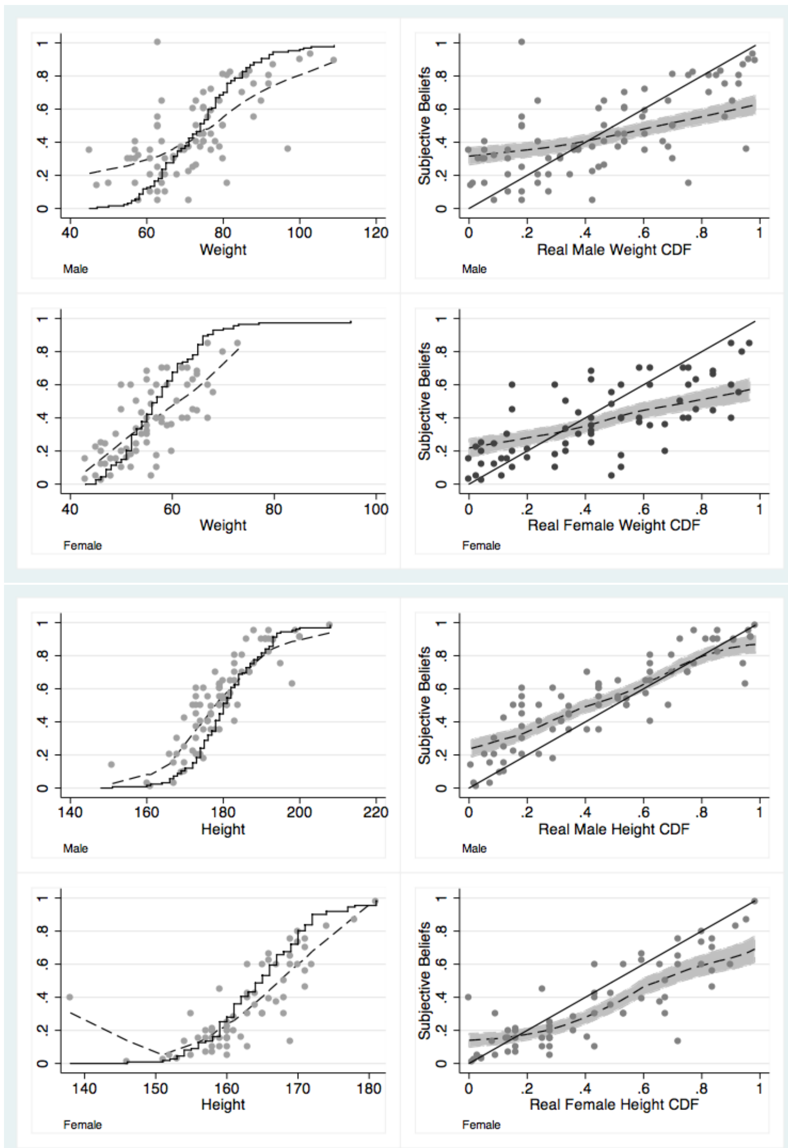

Figure 2: **Happiness and Political Stance:** The left panels represent the real CDFs (continuous lines) and the beliefs over the distributions (dots) and their respective Lowess function (dashed lines). The right panels represents the 45 degree lines (solid lines) and the local polynomial interpolation of the perceived and the real distributions (dashed lines), the shadow represent the 95% confidence interval.

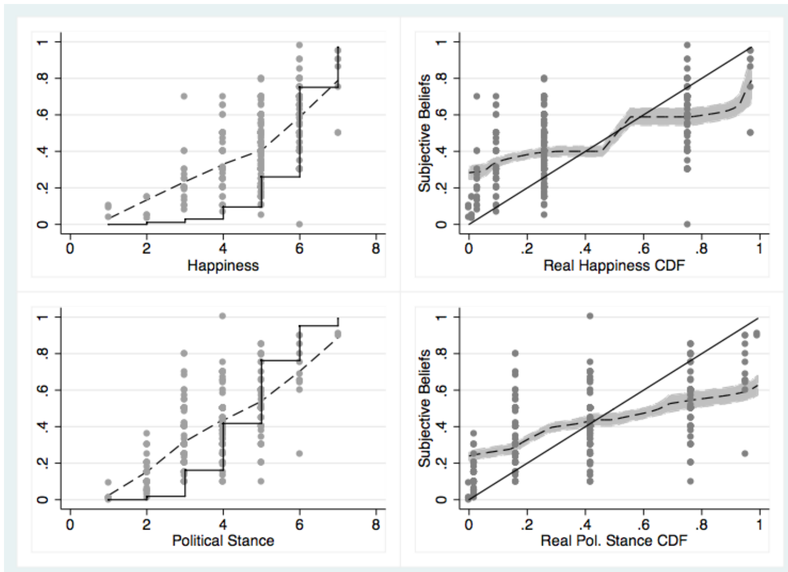

Table 2: **Determinants of Perceived Cumulative Distributions**. Dependent Variables: Perceived Share of Individuals below each individual Happiness, Political Stance, Weight and Height. Robust Standard Errors in Brackets.

	1	2	3	4
	Happiness CDF	Pol. Stance CDF	Weight CDF	Height CDF
	b/se	b/se	b/se	b/se
Happiness	1.2121***			
	(0.2682)			
Happiness^2	–2.0753***			
	(0.7412)			
Happiness^3	1.5031***			
	(0.5781)			
Political Stance		2.0570***		
		(0.2306)		
Political Stance^2		–4.1158***		
		(0.6878)		
Political Stance^3		2.7863***		
		(0.5154)		
Female*Weight			0.5502***	
			(0.0585)	
Male*Weight			0.5538***	
			(0.0682)	
Female*Height				0.7469***
				(0.0627)
Male*Height				0.7907***
				(0.0479)
Male	0.0056	0.0098	0.0612	0.1423***
	(0.0212)	(0.0192)	(0.0481)	(0.0394)
Constant	0.2042***	0.0973***	0.1327***	0.0245
	(0.0303)	(0.0187)	(0.0301)	(0.0277)
r2	0.355	0.471	0.484	0.783
N	273	267	151	147

Figure 3: **Mobile Phones and Restaurant Choices:** The left panels represent the histograms of the simple real distributions (with the different characteristics ordered by frequencies), the beliefs over the distribution (dots) and their respective Lowess function (dashed lines). The right panels represents the 45 degree lines (solid lines) and the local polynomial interpolations of the perceived and the real simple distributions (dashed lines) with the shadow representing the 95% confidence interval.

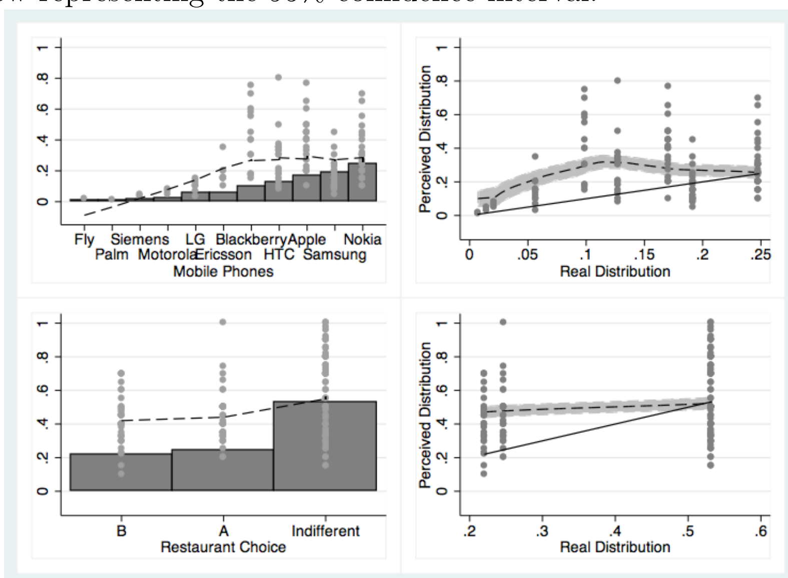

Table 3: **Determinants of Perceived Distributions**. Dependent Variables: Perceived Share of Individuals with the same mobile-phone brand and the same information processing behavior. Robust Standard Errors in Brackets.

	1	2
	Mobile Distribution	Deference Distribution
	b/se	b/se
Mobile	3.2592***	
	(0.6247)	
Mobile^2	–9.6402***	
	(2.3207)	
Deference		0.4330***
		(0.1060)
Male	–0.0536*	0.0141
	(0.0278)	(0.0324)
Constant	0.0652*	0.3217***
	(0.0381)	(0.0394)
r2	0.136	0.097
N	141	154

Table 4: **Errors when Estimating Averages Regressed against Own Values**. Errors are measured as the difference between reported estimates and the calculated average. Each regression includes either solely the own value for each individual or the own value and the square of the own value. Robust Standard Errors in Brackets.

	1	2	3	4
Errors in	Av. Happiness	Av. Pol. St.	Av. Weight	Av. Height
	b/se	b/se	b/se	b/se
Happiness	0.0455***			
	(0.0112)			
Political Stance		0.1644***		
		(0.0590)		
Political Stance^2		–0.0193**		
		(0.0077)		
Female*Weight			0.0045***	
			(0.0011)	
Male*Weight			0.0014***	
			(0.0003)	
Female*Height				0.0132**
				(0.0058)
Female*Height^2				–0.0000**
				(0.0000)
Male*Height				0.0006*
				(0.0003)
Male	0.0275	–0.0182	–0.4539***	0.9170*
	(0.0221)	(0.0345)	(0.0656)	(0.4693)
Constant	–0.2804***	–0.2363**	–0.2476***	–1.0415**
	(0.0562)	(0.1109)	(0.0609)	(0.4662)
r2	0.116	0.045	0.969	0.281
N	154	154	150	147

Table 5: **Determinant of Errors in the perception of Top and Bottom Beliefs**. Dependent variables: Differences between Real and Perceived Averages Happiness and Political Stance of the Top and Bottom 10% individuals. Robust Standard Errors in Brackets.

	1 Top Happ. Errors b/se	2 Bottom Happ. Errors b/se	3 Extreme Right Errors b/se	4 Extreme Left Errors b/se
Constant	–0.1559*	–0.7456***	0.0993*	–0.3470*
	(0.0813)	(0.1514)	(0.0588)	(0.1855)
Male	0.0035	–0.0165	0.0200	0.0089
	(0.0162)	(0.0458)	(0.0233)	(0.0712)
Happiness	0.0138	0.0459**		
	(0.0120)	(0.0208)		
Perceived Average Happiness	0.0160	0.0987***		
	(0.0128)	(0.0323)		
Political Stance			–0.0033	0.0152
			(0.0074)	(0.0299)
Perceived Average Political Stance			–0.0063	0.0961**
			(0.0128)	(0.0395)
r2	0.042	0.108	0.007	0.033
N	154	154	154	153

Table 6: **Determinant of Errors in the perception of Top and Bottom Beliefs**. Dependent variables: Differences between Real and Perceived Averages of Height and Weight of the Top and Bottom 10% individuals. Robust Standard Errors in Brackets.

	1	2	3	4
	Top	Bottom	Top	Bottom
	Weight Errors	Weight Errors	Height Errors	Height Errors
	b/se	b/se	b/se	b/se
Constant	–0.3571	–0.4631	–0.6288**	–0.5583***
	(0.2788)	(0.3008)	(0.2467)	(0.1348)
Male	0.1246	–0.1941	0.1937	–0.6332**
	(0.2945)	(0.3468)	(0.2815)	(0.2477)
Female*Weight			–0.0022	0.0014
			(0.0025)	(0.0019)
Male*Weight			0.0007	0.0017
			(0.0012)	(0.0012)
Female*Height	–0.0000	0.0009		
	(0.0005)	(0.0009)		
Male*Height	0.0007	0.0009		
	(0.0005)	(0.0006)		
Perceived Average Female Height	0.0022	0.0018		
	(0.0017)	(0.0013)		
Perceived Average Male Height	0.0004	0.0028**		
	(0.0006)	(0.0012)		
Perceived Average Female Weight			0.0144***	0.0081***
			(0.0039)	(0.0029)
Perceived Average Male Weight			0.0054**	0.0150***
			(0.0023)	(0.0032)
r2	0.312	0.131	0.174	0.378
N	147	147	149	150

Supplementary Information: Instructions

This Appendix presents the full on-screen instructions and questions for the entire experiment in the precise order observed by participants. The "session ID" box was automatically completed based on their initial login credentials. The questions marked with a "*" indicate required questions. Once they had answered each required question they could then proceed by hitting the "submit answers" button at the bottom of each web-page. After round 7 they were taken to a final page where they were asked to wait patiently until their fellow participants had finished before receiving payment.

ROUND 1

ID ENTRY

Session ID:*

All of your answers will be entirely anonymous. You have been allocated a Session ID and password, for example your ID might be "48576" or "60306". This ID and password are unique to you and will help us link the data you enter across several forms. However both the session ID and password are entirely anonymous and cannot be used to determine your personal identity or university ID.

ROUND 1 QUESTIONS

Please answer all of the following questions to the best of your ability and please scroll down through all of the questions until you reach the "SUBMIT ANSWERS" button.

What is your gender?

Your answer:*

What is your height?

The menu includes heights measured metric and imperial measures, so you can choose the units with which you are most comfortable.

Your answer:*

What is your weight at the moment?

The menu includes weights measured metric and imperial measures, so you can choose the units with which you are most comfortable.

Your answer:*

How would you rate your happiness at the moment?
Please use a 7-point scale where 1 is completely sad, 2 is very sad, 3 is fairly sad, 4 is neither happy nor sad, 5 is fairly happy, 6 is very happy and 7 is completely happy.
Your answer:*

How would you rate your political beliefs at the moment?
Please use a 7-point scale where 1 is far left, 2 is left, 3 is centre left, 4 is centre, 5 is centre right, 6 is right and 7 is far right.
Your answer:*

What is your current brand of mobile phone? If you have more than one enter the one that you use most often.
Your answer:*

Imagine that you have to decide between two restaurants in which to have dinner alone. They are called restaurant A and B. You have some private information that A is better, but you know that an equally well-informed colleague has information suggesting that B is better. Would you choose to eat at A, B or are you indifferent?
Your answer:*

END OF ROUND 1
Privacy statement
REMINDER: All of your answers today will be entirely anonymous. No data which personally identifies you is collected on the form. Your answers will only be linked to your session ID and not to your personal identity or your university ID. When you are ready to submit your answers please hit the "SUBMIT ANSWERS" button below.
(SUBMIT ANSWERS BUTTON)

ROUND 2

ID ENTRY
Session ID:*
All of your answers will be entirely anonymous. You have been allocated a Session ID

and password, for example your ID might be “48576” or “60306”. This ID and password are unique to you and will help us link the data you enter across several forms. However both the session ID and password are entirely anonymous and cannot be used to determine your personal identity or university ID.

VERY IMPORTANT - BONUS PAYMENT EXPLANATION - PLEASE READ!
You will receive a certain £2.50 for showing-up but by answering the next few questions you have the chance to win a bonus. One of the questions on this page and the ones to follow will be chosen at random and secretely allocated to be the “prize question”. If your answer is within 10% of the correct answer to the prize question you will win a bonus payment of £5.

ROUND 2 QUESTIONS
Please answer all of the following questions to the best of your ability as one of these may be the "prize question" and please scroll down through all of the questions until you reach the "SUBMIT ANSWERS" button.

What percentage of students at Warwick do you think are less happy than you are?
Your answer:*

What percentage of students at Warwick do you think are less right-wing than you?
Your answer:*

What percentage of students at Warwick of your gender do you think are shorter than you?
Your answer:*

What percentage of students at Warwick of your gender do you think are lighter than you?
Your answer:*

Think again about the mobile phone brand you listed in the last round. What percentage of students at Warwick do you think also use the same brand of mobile phone as their main mobile phone?
Your answer:*

Think again about the restaurant question in the last round. To remind you, you had to decide between two restaurants in which to have dinner alone. They were called restaurant A and B. You had some private information that A is better, but you knew that an equally well-informed colleague had information suggesting that B was better.
What percentage of your fellow Warwick students do you think would have chosen the same answer as you? Remember that the options were “indifferent”, “A” or “B”.
Your answer:*

END OF ROUND 2
Privacy statement
REMINDER: All of your answers today will be entirely anonymous. No data which personally identifies you is collected on the form. Your answers will only be linked to your session ID and not to your personal identity or your university ID. When you are ready to submit your answers please hit the “SUBMIT ANSWERS” button below.
(SUBMIT ANSWERS BUTTON)

ROUND 3

ID ENTRY
Session ID:*
All of your answers will be entirely anonymous. You have been allocated a Session ID and password, for example your ID might be “48576” or “60306”. This ID and password are unique to you and will help us link the data you enter across several forms. However both the session ID and password are entirely anonymous and cannot be used to determine your personal identity or university ID.

ROUND 3 QUESTIONS
Please answer all of the following questions to the best of your ability as one of these may be the ”prize question” and please scroll down through all of the questions until you reach the “SUBMIT ANSWERS” button.

What do you think is the average height for someone who is in the 10% tallest Warwick students of your gender?
Your answer:*

What do you think is the average weight for someone who is in the 10% heaviest Warwick students of your gender?
Your answer:*

What do you think is the average happiness for someone who is in the 10% happiest students at Warwick?
Please use a 7-point scale where 1 is completely sad, 2 is very sad, 3 is fairly sad, 4 is neither happy nor sad, 5 is fairly happy, 6 is very happy and 7 is completely happy.
Your answer:*

What do you think is the average political belief for someone who is in the 10% most right-wing students at Warwick?
Please use a 7-point scale where 1 is far left, 2 is left, 3 is centre left, 4 is centre, 5 is centre right, 6 is right and 7 is far right.
Your answer:*

END OF ROUND 3
Privacy statement
REMINDER: All of your answers today will be entirely anonymous. No data which personally identifies you is collected on the form. Your answers will only be linked to your session ID and not to your personal identity or your university ID. When you are ready to submit your answers please hit the "SUBMIT ANSWERS" button below.
(SUBMIT ANSWERS BUTTON)

ROUND 4

ID ENTRY
Session ID:*
All of your answers will be entirely anonymous. You have been allocated a Session ID and password, for example your ID might be "48576" or "60306". This ID and password are unique to you and will help us link the data you enter across several forms. However both the session ID and password are entirely anonymous and cannot be used to determine your personal identity or university ID.

ROUND 4 QUESTIONS

Please answer all of the following questions to the best of your ability as one of these may be the "prize question" and please scroll down through all of the questions until you reach the "SUBMIT ANSWERS" button.

What do you think is the average height for someone who is in the 10% shortest Warwick students of your gender?
Your answer:*

What do you think is the average weight for someone who is in the 10% lightest Warwick students of your gender?
Your answer:*

What do you think is the average happiness for someone who is in the 10% most sad at Warwick?
Please use a 7-point scale where 1 is completely sad, 2 is very sad, 3 is fairly sad, 4 is neither happy nor sad, 5 is fairly happy, 6 is very happy and 7 is completely happy.
Your answer:*

What do you think is the average political belief for someone who is in the 10% most left-wing students at Warwick? Please use a 7-point scale where 1 is far left, 2 is left, 3 is centre left, 4 is centre, 5 is centre right, 6 is right and 7 is far right.
Your answer:*

END OF ROUND 4

Privacy statement

REMINDER: All of your answers today will be entirely anonymous. No data which personally identifies you is collected on the form. Your answers will only be linked to your session ID and not to your personal identity or your university ID. When you are ready to submit your answers please hit the "SUBMIT ANSWERS" button below.
(SUBMIT ANSWERS BUTTON)

ROUND 5

ID ENTRY

Session ID:*

All of your answers will be entirely anonymous. You have been allocated a Session ID and password, for example your ID might be "48576" or "60306". This ID and password are unique to you and will help us link the data you enter across several forms. However both the session ID and password are entirely anonymous and cannot be used to determine your personal identity or university ID.

ROUND 5 QUESTIONS

Please answer all of the following questions to the best of your ability as one of these may be the prize question and please scroll down through all of the questions until you reach the SUBMIT ANSWERS button.

What do you think is the average height for a Warwick student of your gender?

The menu includes heights measured metric and imperial measures, so you can choose the units with which you are most comfortable.

Your answer:*

What do you think is the average weight for a Warwick student of your gender?

The menu includes weights measured metric and imperial measures, so you can choose the units with which you are most comfortable.

Your answer:*

What do you think is the average happiness at the moment for a Warwick student?

Please use a 7-point scale where 1 is completely sad, 2 is very sad, 3 is fairly sad, 4 is neither happy nor sad, 5 is fairly happy, 6 is very happy and 7 is completely happy.

Your answer:*

What do you think is the average political belief for a Warwick student?

Please use a 7-point scale where 1 is far left, 2 is left, 3 is centre left, 4 is centre, 5 is centre right, 6 is right and 7 is far right.

Your answer:*

Please estimate the percentage of your fellow Warwick students who use each of the following mobile phone brands as their primary mobile phone.

Note that the list is in alphabetical order (except for the "other" category) and you should insert an entry for each brand (including "other").

Note concerning payment: if this question is the one that is randomly chosen to be the payment question then a single brand will be chosen randomly and if you are within 15% of the correct answer you will receive the £5 bonus payment.

(Participants were then presented with a list of mobile phone brands in the following order: Apple Iphone, Blackbery, HTC, LG, Motorola, Nokia, Samsung, Siemens, Sony-Ericson, Other. For each one they were asked to select from a list of 5% intervals, beginning 0 to 5% and ending with 95 to 100%)

Think again about the restaurant question you were asked earlier in the session. To remind you, you had to decide between two restaurants in which to have dinner alone. They were called restaurant A and B. You had some private information that A is better, but you knew that an equally well-informed colleague had information suggesting that B was better.

What percentage of your fellow Warwick students do you think would have chosen to eat at restaurant A if they were asked the same question? Remember that the other options were indifferent and B.
Your answer:*

END OF ROUND 5
Privacy statement
REMINDER: All of your answers today will be entirely anonymous. No data which personally identifies you is collected on the form. Your answers will only be linked to your session ID and not to your personal identity or your university ID. When you are ready to submit your answers please hit the "SUBMIT ANSWERS" button below.
(SUBMIT ANSWERS BUTTON)

ROUND 6

ID ENTRY

Session ID:*

All of your answers will be entirely anonymous. You have been allocated a Session ID and password, for example your ID might be “48576” or “60306”. This ID and password are unique to you and will help us link the data you enter across several forms. However both the session ID and password are entirely anonymous and cannot be used to determine your personal identity or university ID.

ROUND 6 QUESTION

Please answer the following question to the best of your ability as it may be the “prize question” and please scroll down until you reach the SUBMIT ANSWERS button.

Consider the following gamble. You have a 20% chance of winning £100, a 40% chance of winning £10 and a 40% chance of winning £0. If you played this gamble many times what would you expect to be your average winnings per gamble? (in pounds)

Your answer:*

END OF ROUND 6

Privacy statement

REMINDER: All of your answers today will be entirely anonymous. No data which personally identifies you is collected on the form. Your answers will only be linked to your session ID and not to your personal identity or your university ID. When you are ready to submit your answers please hit the “SUBMIT ANSWERS” button below.

(SUBMIT ANSWERS BUTTON)

ROUND 7

ID ENTRY

Session ID:*

All of your answers will be entirely anonymous. You have been allocated a Session ID and password, for example your ID might be “48576” or “60306”. This ID and password are unique to you and will help us link the data you enter across several forms. However both the session ID and password are entirely anonymous and cannot be used to determine your personal identity or university ID.

ROUND 7 QUESTIONS

This is the final round. Please answer the questions to the best of your ability and please scroll down through all of the questions until you reach the "SUBMIT ANSWERS" button. Note that none of the following questions are possible "prize questions".

What is your age? (in years)
Your answer:*

What is your nationality?
Your answer:*

What is your degree subject?
Your answer:*

Did you study maths up to and including your final year at school? (e.g. to A-level or IB SL or HL?)
Your answer*

Please comment on the methods you used (if any) during the rounds when you had the opportunity to win bonus payments.
Your answer:

END OF ROUND 7

Privacy statement

REMINDER: All of your answers today will be entirely anonymous. No data which personally identifies you is collected on the form. Your answers will only be linked to your session ID and not to your personal identity or your university ID. When you are ready to submit your answers please hit the "SUBMIT ANSWERS" button below.

(SUBMIT ANSWERS BUTTON)