FACULTAD DE CIENCIAS EMPRESARIALES Y ECONOMIA

Serie de documentos de trabajo del Departamento de Economía /
Department of Economics Working Papers Series

APPARENT OVERCONFIDENCE*

Jean- Pierre Benoit
Economics Area
London Buisness School
jpbenoit@london.edu

Juan Dubra
Departamento de Economía
Facultad de Ciencias Empresariales y Economía
Universidad de Montevideo
jdubra@um.edu.uy

Working paper UM_CEE_2011-06
http://www.um.edu.uy/cee/investigaciones/

Apparent Overconfidence*

Jean-Pierre Benoît
London Business School

Juan Dubra
Universidad de Montevideo

August 9, 2011

Abstract

It is common for a majority of people to rank themselves as better than average on simple tasks and worse than average on difficult tasks. The literature takes for granted that this apparent misconfidence is problematic. We argue, however, that this behaviour is consistent with purely rational Bayesian updaters. In fact, better-than-average data by itself cannot be used to show overconfidence; we indicate which type of data can be used. Our theory is consistent with empirical patterns found in the literature.

Keywords: Overconfidence; Better than Average; Experimental Economics; Irrationality; Signalling Models.
Journal of Economic Literature Classification Numbers: D11, D12, D82, D83

For a while, there was a consensus among researchers that overconfidence is rampant. Typical early comments include "Dozens of studies show that people ... are generally overconfident about their relative skills" (Camerer, 1997), "Perhaps the most robust finding in the psychology of judgment is that people are overconfident" (DeBondt and Thaler, 1995), and "The tendency to evaluate oneself more favorably than others is a staple finding in social psychology" (Alicke et al. 1995).[1] Recent work has yielded a more nuanced consensus: When the skill under consideration is an easy one to master, populations display overconfidence in their relative judgements, but when the skill is difficult they display underconfidence (see, for example, Kruger et al. (2008) and Moore (2007)). In this paper, we argue that both the earlier and the later consensus are misleading – much of the supposed evidence for misconfidence reveals only an *apparent*, not a true, overconfidence or underconfidence.

*We thank Stefano Sacchetto and Gabriel Illanes for their research assistance. We also thank Raphael Corbi, Rafael Di Tella, John Duffy, Federico Echenique, Emilio Espino, PJ Healy, Richard Lowery, Henry Moon, Don Moore, Nigel Nicholson, Madan Pilutlla, Matt Rabin, Ariel Rubinstein, Luís Santos-Pinto, Jack Stecher and Juan Xandri for their comments. Juan Dubra gratefully acknowledges the financial support of ANII.

[1] Papers on overconfidence in economics include Camerer and Lovallo (1999), Fang and Moscarini (2005), Garcia, Sangiorgi and Urosevic (2007), Hoelzl and Rustichini (2005), Kőszegi (2006), Menkhoff et al. (2006), Noth and Weber (2003), Sandroni and Squintani (2008), Van den Steen (2004), Zábojník (2004). In finance, recent (published) papers include Barber and Odean (2001), Biais et al. (2005), Bernardo and Welch (2001), Chuang and Lee (2006), Daniel, Hirshleifer and Subrahmanyam (2001), Kyle and Wang (1997), Malmendier and Tate (2005), Peng and Xiong (2006), Wang (2001).

While we consider the evidence on overconfidence and underconfidence, for expository purposes we emphasize overconfidence, since this is the bias that is better known among economists. Overconfidence has been reported in peoples' beliefs in the precision of their estimates, in their views of their absolute abilities, and in their appraisal of their relative skills and virtues. In this paper, we analyze the last form of overconfidence (or underconfidence), which has been termed overplacement (underplacement) by Larrick, Burson, and Soll (2007). Our analysis has implications for overconfidence in absolute abilities as well, but it is not directly applicable to overconfidence in the precision of estimates.

Myers (1999, p.57) cites research showing that most people perceive themselves as more intelligent than their average peer, most business managers rate their performance as better than that of the average manager, and most high school students rate themselves as more original than the average high-schooler. These findings, and others like them, are typically presented as evidence of overconfidence without further comment. Presumably, the reason for this lack of comment is that, since it is impossible for most people to be better than average, or, more accurately, better than the median, it is *obvious* that some people must have inflated self-appraisals. But the simple truism that most people cannot be better than the median does not imply that most people cannot rationally rate themselves above the median. Indeed, we show that median comparisons, like the ones cited above, can never demonstrate that people are overconfident. More detailed information, such as the percentage of people who believe they rank above each decile and the strengths of these beliefs, is needed.

As an illustration of our main point, consider a large population with three types of drivers, low-skilled, medium-skilled, and high-skilled, and suppose that the probabilities of any one of them causing an accident are $f_l = \frac{47}{80}$, $f_m = \frac{9}{16}$, and $f_h = \frac{1}{20}$, respectively. In period 0, nature chooses a skill level for each person with equal probability, so that the mean probability of an accident is $\frac{2}{5}$. Initially, no driver has information about his, or her, own particular skill level, and each person (rationally) evaluates himself as no better or worse than average. In period 1, everyone drives and learns something about his skill, based upon whether or not he has caused an accident. Each person is then asked how his driving skill compares to the rest of the population.

How does a driver who has not caused an accident reply? Using Bayes' rule, he evaluates his own skill level as follows: $p\,(\text{low skill} \mid \text{no accident}) = \frac{11}{48}$, $p\,(\text{medium skill} \mid \text{no accident}) = \frac{35}{144}$, $p\,(\text{high skill} \mid \text{no accident}) = \frac{19}{36}$. A driver who has not had an accident thinks there is over a $\frac{1}{2}$ chance (in fact, $\frac{19}{36}$) that his skill level is in the top third of all drivers, so that both the median and mode of his beliefs are well above average. His mean probability of an accident is about $\frac{3}{10}$, which is better than for $\frac{2}{3}$ of the drivers and better than the population mean. Moreover, his beliefs about himself strictly first order stochastically dominate (fosd) the population distribution. Any way he looks at it, a driver who has not had an accident should evaluate himself as better than average. Since $\frac{3}{5}$ of drivers have not had an accident, $\frac{3}{5}$ rank themselves as above average and the population of drivers seems overconfident on the whole. However, rather than being *overconfident*, which implies some error in judgement, the drivers are simply using the information available to them in the best possible manner.[2]

[2] For a suggestive calculation using real-world data, in 1990 there were 13,851,000 drivers in the U.S.A aged 16-19, who were involved in 1,381,167 accidents (Massie and Campbell (1993)). Invoking the so-called Pareto principle, let us suppose that 80% of the accidents were caused by 20% of the drivers, and, for simplicity, that there were two types of drivers, good and bad. The above data then yields that bad drivers

Although in this example a driver who has not had an accident considers himself to be above average whether he ranks himself by the mean, mode, or median of his beliefs, such uniformity is not always the case; in general it is important to consider exactly how a subject is placing himself. Note that the example can easily be flipped, by making accidents likely, to generate an apparent underconfidence instead of overconfidence.

The experimental literature uses two types of experiments, those in which subjects rank themselves relative to others, and those in which subjects place themselves on a scale. We show that, in contrast to ranking experiments, certain scale experiments should not have even the appearance of misconfidence.

There is a vast literature on overconfidence, both testing for it and providing explanations for it. On the explanatory side, most of the literature takes for granted that there is something amiss when a majority of people rank themselves above the median, and seeks to pinpoint the nature of the error. Mistakes are said to result from *egocentrism* (Kruger (1999)), *incompetence* (Kruger and Dunning (1999)), or *self-serving biases* (Greenwald (1980)), among other factors. Bénabou and Tirole (2002) introduce a behavioral bias that causes people to become overconfident.

A strand of the literature more closely related to ours involves purely rational Bayesian agents. In Zábojník (2004), agents who are uncertain of their abilities, which may be high or low, choose in each period to either consume or perform a test to learn about these abilities. Given technical assumptions on the agents' utilities, the optimal stopping rule of agents leads them to halt their learning in a biased fashion, and a disproportionate number end up ranking themselves as high in ability. Brocas and Carillo (2007) also have an optimal stopping model, which can be interpreted as leading to apparent overconfidence. In Kőszegi (2006), agents with a taste for positive self-image sample in a way that leads to overplacement. Moore and Healy(2008) have a model in which people are uncertain about both their abilities and the difficulty of the task they are undertaking. They show that people presented with a task that is easier than expected may simultaneously overplace their rankings and underestimate their absolute performances, while the opposite holds true for those presented with a task that is more difficult than expected.

1 Ranking Experiments

In a ranking experiment, a researcher asks each member of a population to rank his "skill" relative to the other members of the group by placing himself, through word or deed, into one of k equally-sized intervals, or *k-ciles* (defined formally below). Implicitly, the experimenter assumes that skills can be well-ordered – say by a one dimensional type – and that the distribution of actual types has a fraction $\frac{1}{k}$ in each k-cile. The experimenter assembles *population ranking data*: a vector $x \in \Delta^k \equiv \left\{x \in \mathbf{R}_+^k : \sum_1^k x_i = 1\right\}$, where x_i, $i = 1, ..., k$ is the fraction of people who rank themselves in the ith k-cile.

Svenson's (1981) work is a prototypical example of a ranking experiment and provides

have a $\theta_b = \frac{2}{5}$ chance of having an accident in a single year, while good drivers have a $\theta_g = \frac{1}{40}$ chance. Using only accidents as a gauge, Bayes' rule and some combinatorics yield that after 3 years, 79% of drivers will have beliefs about themselves that fosd the population distribution.

perhaps the most widely cited population ranking data. Svenson gathered subjects into a room and presented them with the following instructions (among others):

> We would like to know about what you think about how safely you drive an automobile. All drivers are not equally safe drivers. We want you to compare your own skill to the skills of the other people in this experiment. By definition, there is a least safe and a most safe driver in this room. We want you to indicate your own estimated position in this experimental group. Of course, this is a difficult question because you do not know all the people gathered here today, much less how safely they drive. But please make the most accurate estimate you can.

Each subject was then asked to place himself or herself into one of ten intervals, yielding population ranking data $x \in \Delta^{10}$. Svenson found that a large majority of subjects ranked themselves above the median. To determine if Svenson's data evinced bias, inconsistency, or irrationality in his subjects, we need a notion of what it means for data to be rational and consistent. We derive this notion using an approach based upon the Harsanyi common prior paradigm.

We first define a **rationalizing model** $\left(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta}\right)$, where $\Theta \subseteq \mathbf{R}$ is a type space, p is a prior probability distribution over Θ, S is a set of signals, and $\{f_\theta\}_{\theta\in\Theta}$ is a collection of likelihood functions: each f_θ is a probability distribution over S. We adopt the following interpretation of this model. There is a large population of individuals. In period 0 nature draws a skill level, or type, for each individual independently from p. Higher types correspond to higher skill levels. The prior p is common knowledge, but individuals are not informed directly of their own type. Rather, each agent receives information about himself from his personal experience. This information takes the form of a signal, with an individual of type $\theta \in \Theta$ receiving signal $s \in S$ with probability $f_\theta(s)$. Draws of signals are conditionally independent. Given his signal and the prior p, an agent updates his beliefs about his type using Bayes' rule whenever possible.

Our basic idea is that data is unproblematic if it can arise from a population whose beliefs are generated within a rationalizing model. Implementing this idea, however, is not completely straightforward. Recall that in Svenson's ranking experiment, his instructions state that he is asking "a difficult question because (the subjects) do not know all the people gathered, ... much less how safely they drive." But even if this difficulty were not present – for instance, if the subjects assumed that as a group they formed a representative draw from a well-known population – an issue would remain: Does a person know how safely she herself drives? Of course, a driver has more information about herself than about a stranger, but there is no reason to presume that she knows precisely how safe her driving is[3] (even assuming that she knows exactly what it means to drive "safely"[4]). Thus, a person may

[3] Several strands of the psychology literature, including Festinger's (1954) social comparison theory and Bem's (1967) self-perception theory stress that people are uncertain of their types. In the economics literature, a number of papers start from the premise that, as Bénabou and Tirole (2002) put it, "learning about oneself is an ongoing process".

[4] Dunning et al. (1989) argue that people may have different notions of what it means to drive safely, so that the data is not what it appears to be. Here, we give the best case for the data and assume that all subjects agree on the meaning of a safe driver.

consider herself to be quite a safe driver since she has never had an accident, rarely speeds, and generally manoeuvres well in traffic, but at the same time realize that the limited range of her experience restricts her ability to make a precise self-appraisal. In ranking herself, a subject must form beliefs about her own driving safety. Svenson ignores this issue and, in effect, asks each subject for a summary statistic of her beliefs, without specifying what this statistic should be. There is no way of knowing if subjects responded using the medians of their beliefs, the means, the modes, or some other statistic. As a result, it is unclear what to make of Svenson's data. Svenson's experiment is hardly unique in this respect: much of the overconfidence literature, and other literatures as well, share this feature that the meaning of responses is not clear.[5]

Not all experiments share this ambiguity, however. For instance, the design of Hoelzl and Rustichini (2005) induces subjects to place themselves according to their median beliefs, while Moore and Healy (2008) calculate mean beliefs from subjects' responses. In the interest of space, in this paper we analyze ranking data only under the assumption that subjects place themselves according to their median types – that is, that a subject places himself into a certain k-cile if he believes there is at least a probability $\frac{1}{2}$ that his actual type is in that k-cile or better and a probability $\frac{1}{2}$ that his type is in that k-cile or below – or, more generally, according to some specific quantile of their types. In Benoît and Dubra (2009), we establish that the analysis is similar under different assumptions about subjects' responses (see below also).

The next definition says that data can be *median-rationalized* when it corresponds to the medians of the posteriors of a rationalizing model. We start with some preliminary notation.

- Given Θ, p and k, for each $0 \leq i \leq k$, let Θ_i denote the ith k-cile: for $i \leq k-1$, $\Theta_i = \left\{\theta \in \Theta \mid \frac{i-1}{k} \leq p\left(\theta' < \theta\right) < \frac{i}{k}\right\}$ and $\Theta_k = \left\{\theta \in \Theta \mid \frac{k-1}{k} \leq p\left(\theta' < \theta\right)\right\}$. Note that a k-cile is a set of types, not a cutoff type, and that higher $k-$ciles correspond to higher types. We do not include the dependence of Θ_i on p and k, since this does not cause confusion.

- Given k and a rationalizing model $\left(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta}\right)$, for each $0 \leq i \leq k$, let S_i denote the set of signals that result in an updated median type in Θ_i:
$$S_i = \left\{ s \in S \mid p\left(\cup_{n=i}^{k}\Theta_n \mid s\right) \geq \frac{1}{2} \text{ and } p\left(\cup_{n=1}^{i}\Theta_n \mid s\right) \geq \frac{1}{2} \right\}.$$
Let F denote the marginal of signals over S: for each (measurable) $T \subset S$, $F\left(T\right) = \int_\Theta \int_T df_\theta\left(s\right) dp\left(\theta\right)$.

Thus, $F\left(S_i\right)$ is the (expected) fraction of people that will place their median types in decile i, when types are distributed according to p and signals are received according to f_θ.

Definition 1 *Given a type space $\Theta \subseteq \mathbf{R}$ and a distribution p over Θ, the population ranking data $x \in \Delta^k$ can be* ***median-rationalized for*** (Θ, p) *if there is a rationalizing model* $\left(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta}\right)$ *with $x_i = F\left(S_i\right)$, for $i = 1, ..., k$.*

[5] Dominitz (1998), in critiquing a British survey of expected earnings, writes "what feature of the subjective probability distribution determines the category selected by respondents? Is it the mean? Or perhaps it is the median or some other quantile. Or perhaps it is the category that contains the most probability mass."

The example in the introduction shows that $x = \left(0, \frac{2}{5}, \frac{3}{5}\right)$ can be median-rationalized for $\Theta = \{l, m, h\}$ and $p(h) = p(m) = p(l) = \frac{1}{3}$. When ranking data is median-rationalizable, it can arise from a Bayes-rational population working from a common prior, and there is no prima facie case for calling it biased.

The following theorem indicates when data can be median-rationalized. In effect, a rational population can appear to be twice as confident as reality would suggest, but no more. For instance, suppose that people place themselves into ten intervals ($k = 10$). Then apparently overconfident data in which up to $\frac{2}{10}$ of the people rank themselves in the top 10-cile, up to $\frac{4}{10}$ rank themselves in the top two 10-ciles, and up to $\frac{2i}{10}$ rank themselves in the top i 10-ciles for $i = 3, 4, 5$ can be rationalized. However, data in which $\frac{1}{2}$ of the population places itself in the top two 10-ciles cannot be explained as rational.

Theorem 1 *Suppose that $\Theta \subseteq \mathbf{R}$ and p is a distribution over Θ such that $p(\Theta_i) = 1/k$ for all i. Then the population ranking data $x \in \Delta^k$ can be median-rationalized for (Θ, p) if and only if, for $i = 1, ..., k$,*

$$\sum\nolimits_{j=i}^{k} x_j \;<\; \frac{2}{k}(k - i + 1), \qquad \textit{and} \tag{1}$$

$$\sum\nolimits_{j=1}^{i} x_j \;<\; \frac{2}{k} i. \tag{2}$$

Proof. All proofs are in the appendix. ■

Population ranking data x that satisfies the necessary conditions (1) and (2) can be generated from a rational population with any distribution of types p, provided only that p is legitimate in the sense that it partitions the population uniformly into k-ciles.[6] Conversely, if the necessary conditions are not satisfied,[7] then no legitimate prior p can yield the data x.

The necessary part of Theorem 1 comes from the fact that Bayesian beliefs must average out to the population distribution. When people self-evaluate by their median type, $\frac{1}{2} \times \sum_{j=i}^{k} x_j$ is a lower bound on the weight their beliefs put into the top $(k - i + 1)$ k-ciles. If (1) is violated for some i, then too much weight, i.e., more than $\frac{1}{k}(k - i + 1)$, is put into these k-ciles. Similarly for condition (2). The sufficiency part of the theorem is more involved, although it is straightforward for the special case where each $x_i < \frac{2}{k}$. For this case, set $S = (s_1, ..., s_k)$ and let types $\theta \in \Theta_i$ observe signal s_j with probability $f_{\theta \in \Theta_i}(s_j) = k\left(\frac{1}{k} - \frac{x_i}{2}\right) x_j$ for $i \neq j$, and $f_{\theta \in \Theta_i}(s_j) = k\left(\frac{1}{2} + \frac{1}{k} - \frac{x_i}{2}\right) x_j$ for $i = j$. Then $\left(\Theta, p, S, \{f_\theta\}_{\theta \in \Theta}\right)$ median-rationalizes x for (Θ, p).

Theorem 1 is a corollary of a more general theorem, formally stated and proved in the appendix, that covers the case in which person i places himself into a k-cile based upon an arbitrary quantile q of his beliefs. In that case, data x can be rationalized if and only if for all i,

$$\sum\nolimits_{j=i}^{k} x_j < \frac{1}{qk}(k - i + 1) \text{ and } \sum\nolimits_{j=1}^{i} x_j < \frac{1}{(1-q)k} i. \tag{3}$$

[6] This restriction on p avoids uninteresting trivialities. If, for instance, the distribution p were to assign all the weight to just the first two k-ciles, then even $\left(\frac{1}{k}, ..., \frac{1}{k}\right)$ could not be median-rationalized for any $k > 2$.

[7] In fact, conditions (1) only have bite for $i > \frac{k+1}{2}$, while conditions (2) only have bite for $i < \frac{k+1}{2}$ (Since $\frac{k+1}{2}$ may or may not be an integer, it is easiest to state the conditions as we do.)

As an application of this more general theorem to an experimental setting, if a group of subjects is offered the choice between a prize of M if their score on a quiz places them in the 8^{th} decile or above and M with probability 0.7, at most 43% should bet on their quiz placement.

Researchers often summarize population ranking data by the percentage of people who place themselves above the population median. However, such data cannot be used to show overconfidence. Theorem 1 shows that any fraction $r < 1$ of the population can rationally place itself in the top half when people self-evaluate using their median types. Benoît and Dubra (2009, 2011) establish that the same holds true when people self-evaluate using their mean or modal types.[8]

In Svenson's experiment, students in Sweden and the United States were questioned about their driving safety and driving skill relative to their respective groups. Swedish drivers placed themselves into ten- percent intervals in the following proportions when asked about their safety:

Interval	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Reports (%)	0.0	5.7	0.0	14.3	2.9	11.4	14.3	28.6	17.1	5.7

Note first that, although a majority of drivers rank themselves above the median, this population ranking data does not have an unambiguously overconfident appearance, as fewer than 10% of the drivers population place themselves in the top 10%. More importantly, Theorem 1 implies that this data can be median-rationalized, as can the Swedish responses on driving skill. On the other hand, on both safety and skill, Svenson's American data cannot be median-rationalized. For instance, 82% of Americans placed themselves in the top 30% on safety and 46% placed themselves in the top 20% on skill. Thus, Svenson does find some evidence of overconfidence, if his subjects based their answers on their median types, but this evidence is not as strong as is commonly believed. Note also that when 46% of the population place themselves in the top 20%, this is only 6% too many, not 26%.

Some researchers summarize their results by their subjects' mean k-cile placement, $\mu(x) = \sum_{i=1}^{k} ix_i$, and infer overconfidence if $\mu > \frac{k+1}{2}$. The following corollary to Theorem 1 shows that much of this overconfidence is only apparent.

Corollary 1 *Suppose that $\Theta \subseteq \mathbf{R}$ and p is a distribution over Θ such that $p(\Theta_i) = 1/k$ for all i. Then the mean k-cile placement μ can come from population ranking data that can be median-rationalized for (Θ, p) if and only if*

$$\begin{array}{ll} \left|\mu - \frac{k+1}{2}\right| < \frac{k}{4} & \text{for } k \text{ even} \\ \left|\mu - \frac{k+1}{2}\right| < \frac{\left(k - \frac{1}{k}\right)}{4} & \text{for } k \text{ odd} \end{array}$$

Thus, when subjects are asked to place themselves into 10-ciles, a mean placement of, say, 7.9 is not out of order.

We have modelled individuals who know the distribution of types in the population. It is easy to generalize beyond this, although a bit of care must be taken, as the following

[8] Benoît and Dubra (2009) also show how the ideas in this paper can be applied to game theoretic settings, such as the entry games in Camerer and Lovallo (1999).

example shows. Let the type space be $\Theta = [0,1]$. In Period 0, nature chooses one of two distributions, $p = U[0,1]$ with probability $\frac{4}{5} < q < 1$ or $p' = U\left[\frac{3}{4},1\right]$ with probability $1-q$. Nature then assigns each individual a type, using the chosen distribution. In Period 1, every one is informed exactly of his type , i.e., $f_\theta(\theta) = 1$, and then *median-ranks* himself. Although individuals know their own types, the median plays a role since individuals are not told which distribution nature used in assigning types. (Recall that a median placement in k-cile j means that a subject believes there is at least a $\frac{1}{2}$ chance that his type lies in k-cile j or above and a $\frac{1}{2}$ chance that it lies in j or below.[9])

Suppose that, as it happens, nature used the distribution p' in assigning types, so that all types lie in the interval $\left[\frac{3}{4},1\right]$. After being informed of his type, any individual believes there is a $\frac{q}{q+4(1-q)} > \frac{1}{2}$ chance that the distribution of types is p. Since the lowest type is $\frac{3}{4}$, *all* individuals median-place themselves somewhere in the top 25% of the population. In contrast, Theorem 1 does not allow more than half the individuals to place themselves in the top 25%. Note, however, that we have analysed the result of only one population distribution draw, namely p', and one draw is necessarily biased. If we consider a large number of draws, so that a fraction q of the time the population distribution is p, we will find that overall only the fraction $\left(q \times \frac{1}{4}\right) + ((1-q) \times 1) < \frac{2}{5}$ places itself in the top 25%, in line with the theorem.[10]

In general, we model a population of individuals who may be uncertain of both their own types and the overall distribution of types using a quadruplet $\left(\Theta, \pi, S, \{f_\theta\}_{\theta\in\Theta}\right)$, where π is a prior over a set of probability distributions over Θ, and making concomitant changes to the definitions of a k-cile, etc... . With this modelling , conditions (1) and (2) in Theorem 1 remain necessary and sufficient for median-rationalization. Indeed, (1) and (2) remain necessary in any environment in which Bayesian updaters start from a common and correct prior, since these conditions simply reflect the fact that beliefs average out to the population distribution.

1.1 Monotone Signals

Theorem 1 describes when population ranking data can be rationalized, without regard as to whether or not the collection of likelihood functions $\{f_\theta\}$ used is, in some sense, reasonable. While it may not be possible to specify exactly what constitutes a reasonable collection of likelihood functions, it is possible to identify some candidate reasonable properties. One such property is that better types should be more likely to receive better signals (for instance, a safe driver should expect to experience few adverse driving incidents) and, conversely, better signals should be indicative of better types. More precisely, given $\Theta \subseteq \mathbf{R}$ and $S \subseteq \mathbf{R}$, we say that the collection of likelihood functions $\{f_\theta\}_{\theta\in\Theta}$ satisfies the **monotone signal property** (msp) if i) $f_{\theta'}$ fosd f_θ for $\theta' > \theta \in \Theta$, and ii) for all $s' > s \in S$, the posterior after s' fosd the posterior after s, for all probability distributions p over Θ that assign probability $1/k$ to each $k-$cile. A rationalizing model $\left(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta}\right)$ satisfies msp if $\{f_\theta\}_{\theta\in\Theta}$ satisfies msp.

[9] Information on median-placements can be induced by offering subjects the choice between a prize with probability $\frac{1}{2}$, and the prize if their type is at least in k-cile j.

[10] Uncertainty about the distribution of types can be interpreted as uncertainty about the difficulty of the task, in the spirit of Moore and Healy (2008) (although our results differ from theirs).

A standard restriction found in the literature is that the collection of f_θ's should satisfy the monotone likelihood ratio property (mlrp): for all $\theta' > \theta$, $\frac{f_{\theta'}(s)}{f_\theta(s)}$ is increasing in s. Mlrp is equivalent to insisting that f satisfy properties i) and ii) for all p, not only those that assign probability $\frac{1}{k}$ to each k-cile (see Whitt (1980) and Milgrom (1981)). In our framework, the only priors p that are relevant are those that divide the type space evenly into k-ciles, so that msp can be seen as the appropriate version of mlrp for our context.

The following theorem, which has been formulated with overconfident looking data in mind, shows that msp imposes more stringent necessary conditions that population ranking data x must satisfy if it is to be rationalized. When $k \leq 4$ these conditions are also sufficient, while for $k > 4$ they are approximately sufficient in the following sense. Define $h = \min\{n \in \mathbf{N} : n > k/2\}$. Say that vector y is **comparable** to x if $y_i = x_i$ for $i = h, ..., k$ and $y_1 \leq x_1$. If x satisfies the necessary conditions, then there is a y comparable to x that can be rationalized. The comparable vector y matches x exactly in the components which are in the upper half, so that in this regard the overconfident aspect of the data is explained. Below the median, however, we may need to do some rearranging. However, we do not to do this by creating a large group of unconfident people who rank themselves in the bottom k-cile.

Theorem 2 *Suppose that $\Theta \subseteq \mathbf{R}$ and p is a distribution over Θ such that $p(\Theta_i) = 1/k$ for all i. The population ranking data $x \in \Delta^k$ can be median-rationalized for (Θ, p) by a rationalizing model that satisfies the monotone signal property only if*

$$\sum_{j=i}^{k} x_j \frac{2j-i-1}{j-1} < \frac{2}{k}(k-i+1), \text{ for } i = 2, ..., k \tag{4}$$

$$\sum_{j=1}^{i} x_j \frac{k+i-2j}{k-j} < \frac{2}{k} i, \text{ for } i = 1, ..., k-1 \tag{5}$$

Suppose $x \gg 0$. Then, for $k \leq 4$ the above inequalities are also sufficient. For $k > 4$, if x satisfies (4) and (5), then there exists a y comparable to x that can be median-rationalized for (Θ, p) with a rationalizing model that satisfies the monotone signal property.

For $i = k$, condition (4) yields $x_k < \frac{2}{k}$, the same necessary condition as in Theorem 1. For $2 \leq i < k$, however, the restrictions on the data are more severe. Thus, for $i = k-1$ we have $x_{k-1} + x_k \frac{k}{k-1} < \frac{4}{k}$, rather than $x_{k-1} + x_k < \frac{4}{k}$. To derive this tighter bound, suppose that data x is median-rationalized by a rationalizing model that satisfies msp. As is shown in the appendix, x is then also median-rationalized by a model with k signals – $S = (s_1, ..., s_k)$ – in which all agents within a k-cile j receive a signal s_i with the same probability $f_{\theta \in \Theta_j}(s_i)$. For each $i = 1, ..., k$ we have: (a) $\sum_{j \geq i}^{k} f_{\theta \in \Theta_j}(s_i) > \frac{k}{2} x_i$, so that an individual who sees signal s_i has unique median type in Θ_i, and (b) $\sum_{j=1}^{k} f_{\theta \in \Theta_j}(s_i) = k x_i$, so that the fraction x_i see signal s_i. Since msp is satisfied, $f_{\theta \in \Theta_j}(s_k)$ is increasing in j. Therefore, $\sum_{j=1}^{k-1} f_{\theta \in \Theta_j}(s_k) \leq (k-1) f_{\theta \in \Theta_{k-1}}(s_k)$ and, from (b), $f_{\theta \in \Theta_{k-1}}(s_k) \geq \frac{k x_k - f_{\theta \in \Theta_k}(s_k)}{(k-1)}$, so that $f_{\theta \in \Theta_{k-1}}(s_{k-1}) \leq 1 - \frac{k x_k - f_{\theta \in \Theta k}(s_k)}{(k-1)}$. Since $f_{\theta \in \Theta_k}(s_{k-1}) \leq (1 - f_{\theta \in \Theta_k}(s_k))$, we have $1 - \frac{k x_k - f_{\theta \in \Theta_k}(s_k)}{(k-1)} + (1 - f_{\theta \in \Theta_k}(s_k)) = 2 - \frac{f_{\theta \in \Theta_k}(k-2) + k x_k}{k-1} \geq f_{\theta \in \Theta_{k-1}}(s_{k-1}) + f_{\theta \in \Theta_k}(s_{k-1}) > \frac{k}{2} x_{k-1}$, where the last inequality follows from (a). Again from (a), we have $2 - \frac{\frac{k}{2} x_k (k-2) + k x_k}{k-1} > 2 - \frac{f_{\theta \in \Theta_k}(s_k)(k-2) + k x_k}{k-1} > \frac{k}{2} x_{k-1}$, as was to be shown.

Conditions (4) and (5) are not sufficient since, for instance, the data $\left(\frac{8}{25},\frac{1}{75},\frac{1}{75},\frac{1}{15},\frac{4}{15},\frac{8}{25}\right)$ cannot be rationalized with monotone signals, although the comparable vector $\left(\frac{9}{75},\frac{9}{75},\frac{8}{75},\frac{1}{15},\frac{4}{15},\frac{8}{25}\right)$ can be. In the appendix, we show by direct construction that such a counterexample cannot arise when $k \leq 4$. The reason is that msp places fewer demands on the likelihood functions when there are fewer signals, and fewer signals are needed when k is smaller.

While the monotone signal property imposes tighter bounds on population ranking data, plenty of scope for apparent overconfidence remains. In particular, msp still allows any fraction $r < 1$ of the population to place itself above the median and vectors comparable to Svenson's Swedish data.

1.2 Some Empirical Considerations

Kruger (1999) finds a "*below-average effect* in domains in which absolute skills tend to be low". Moore (2007), surveying current research, writes that "When the task is difficult or success is rare, people believe that they are below average", while the opposite is true for easy tasks. As suggested by the phrase "success is rare", some tasks are evaluated dichotomously: success or failure. Call an easy task one where more than half the people succeed and a difficult task one where more than half the people fail. Then, if people evaluate themselves primarily on the basis of their success or failure on the task– in the limit, if their only signal is whether or not they succeed – rational updating will lead to a better-than-average effect on easy tasks and a worse-than-average effect on difficult ones.[11] Formally, this situation is described by a rationalizing model $\left(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta}\right)$ with $S = \{0,1\}$ and $f_\theta(1)$ increasing in θ, which yields the fraction $F(1) = \int f_\theta(1)\,dp(\theta)$ of the population with median type above the population median.

Comparing two dichotomously evaluated tasks $\mathcal{F}$ and $\mathcal{G}$, with rationalizing models $(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta})$ and $\left(\Theta, p, S, \{g_\theta\}_{\theta\in\Theta}\right)$, if g_θ fosd f_θ for all θ, then $\mathcal{G}$ is an easier task on which to succeed and more people will rate themselves above the median on $\mathcal{G}$ than $\mathcal{F}$. While this observation is in keeping with the current wisdom on the effect of ease, this simple comparative static does not generalize to tasks that are not evaluated dichotomously.

Suppose that $\mathcal{F}$ and $\mathcal{G}$ are two tasks in which competence is evaluated on the basis of three signals. On each task fifty percent of the population is of type θ_L and fifty percent of type $\theta_H > \theta_L$. The likelihood functions for the tasks are $f_{\theta_L}(1) = \frac{2}{3}$, $f_{\theta_L}(2) = \frac{1}{3}$, $f_{\theta_H}(2) = \frac{1}{2}$, $f_{\theta_H}(3) = \frac{1}{2}$ on task $\mathcal{F}$, and $g_{\theta_L}(1) = \frac{1}{2}$, $g_{\theta_L}(2) = \frac{1}{2}$, $g_{\theta_H}(2) = \frac{1}{3}$, $g_{\theta_H}(3) = \frac{2}{3}$, on task $\mathcal{G}$. Both $\{f_\theta\}_{\theta\in\Theta}$ and $\{g_\theta\}_{\theta\in\Theta}$ satisfy the monotone signal property, so that a higher signal can be interpreted as a better performance. Since g_θ fosd f_θ for all θ, these better performances are easier to obtain on Task $\mathcal{G}$ than on Task $\mathcal{F}$. Nevertheless, only $\frac{1}{3}$ of the population will place itself in the top half on $\mathcal{G}$, while $\frac{2}{3}$ of the population will place itself in the top half of the population on $\mathcal{F}$. On the face of it, this example conflicts with the claim that easier tasks lead to more overconfidence; on reflection this is not so clear. Given the way that low and high types perform on the two tasks, a case can be made that "success" on Task $\mathcal{F}$ is a signal of 2 or above, while success on Task $\mathcal{G}$ is a signal of 3. Then, more people succeed

[11] Moore also notes that, "people believe that they are more likely than others to experience common events – such as living past age 70 – and less likely than others to experience rare events such as living past 100." By interpreting experiencing the event as a "success" (for instance, having a parent live past 70 would be a success), and not experiencing it as "failure", we obtain this prediction about people's beliefs.

on $\mathcal{F}$ than on $\mathcal{G}$, and Task $\mathcal{F}$ is the easier task. (In a more concrete vein, a judgement as to whether or not bowling is easier than skating depends on how one defines success in the two activities). This ambiguity cannot arise when there are only two signals.

At a theoretical level, it is unclear exactly how to define the ease of a task in general and establish a clear link between ease and apparent overconfidence[12]. This suggests that the current wisdom on the impact of ease needs to be refined and reexamined. In line with this suggestion, Grieco and Hogarth (2009) find no evidence of a hard/easy effect, and while Kruger (1999) does find such an effect, his data contains notable exceptions.[13] Moreover, even when our approach predicts a better-than-average effect on a task, it does not necessarily predict that too many people systematically place themselves in the upper k-ciles; that is, that the population ranking data fosd the uniform distribution. As far as we know, the current literature makes no claims in this regard, and it is worth recalling that while Svenson finds a better-than-average effect in his Swedish drivers, he also finds that too few people rank themselves in the top 10% on safety (and the top 20% on skill).

We turn now to some empirical evidence on how experience affects the degree of apparent overconfidence.

Generally speaking, as people gather more information about themselves, they derive tighter estimates of their types. A population with tight estimates can be captured in our framework by only allowing rationalizing models in which, after updating, individuals are at least $c\%$ sure of the k-cile in which their types lie, for some large c. As a corollary of conditions (3), as c increases the fraction of people that can rationally place themselves above the median gets closer and closer to $\frac{1}{2}$. This suggests that populations with considerable experience should exhibit little misconfidence. In keeping with this prediction, Walton (1999) interviews professional truck drivers, who each drive approximately $100,000$ kilometers a year, and finds no bias in their self-assessments of their relative skills. He does find that a majority claim to be safer drivers than average, however, it is quite possible that most of the truckers had only had safe driving experiences, so that a majority could rationally rank themselves highly. Experience also comes with age, and the evidence on age and overconfidence is mixed. While some researchers find that misplacement declines with age, others find no relation.[14]

Accidents and moving violations are, presumably, negative signals about a driver's safety. Despite this, Marotolli and Richardson (1998) find no difference between the confidence levels of drivers who have had adverse driving incidents and those who have not, which points against the hypothesis that they are making rational self-evaluations.[15] On the other hand, Groeger and Grande (1996) find that, although drivers' self-assessments are uncorrelated to the number of accidents they have had, their self-assessments are positively correlated to the average number of *accident-free miles* they have driven. The number of accident-free

[12] However, it is possible to establish such a link for specific cases beyond dichotomous tasks. In Benoît and Dubra (2011) we show that the findings of Hoelzl and Rustichini (2005) and Moore and Cain (2007) on ease can be generated within our framework.

[13] For instance, although Kruger categorizes *organizing for work* as a difficult task, this task also displays a large better-than-average effect.

[14] For instance, Mathews and Moran (1986), and Holland (1993) find that drivers' overplacement declines with age, while Marotolli and Richardson (1998) and Cooper (1990) find no such decline.

[15] Note that interpreting the evidence can be a bit tricky. For instance, an accident may lead a driver to conclude that he *used* to be an unsafe driver but that now, precisely because he has had an accident, he has become quite a safe driver.

miles seems to be the more relevant signal, as one would expect better drivers to drive more, raising their number of accidents.

2 Scale Experiments

In a scale experiment, a scale in the form of a real interval and a population average are specified (sometimes implicitly), and each subject is asked to place himself somewhere on the scale.[16] **Population scale data** is a triplet $(\Theta, m, \bar{a})$, where $\Theta \subset \mathbf{R}$ is a real interval, $m \in \Theta$ is a population average, and $\bar{a} \in \Theta$ is the average of the placements.

The idea underlying scale experiments is that, in a rational population, self-placements should average out to the population average. When the scale is a subjective one, this presumption is, at best, debatable, so let us restrict ourselves to experiments with an objective scale (for a brief discussion of the issues with a subjective scale, see Benoît and Dubra (2009)). As an example, Weinstein (1980) asks students how their chances of obtaining a good job offer before graduation compare to those of other students at their college, with choices ranging from 100% less than average to 5 times the average. Here there is no ambiguity in the meaning of the scale. However, two ambiguities remain; namely, what is meant by an average student, and what a subject means by a point estimate of his or her own type.

To illustrate, suppose for the sake of discussion that all of Weinstein's subjects agree that there are two types of students at their college, low and high, with job offer probabilities $p_L = 0.3$ and $p_H = 1$, and that 80% of the population are low type. A reasonable interpretation of an average student is one whose chance of obtaining an offer is 0.3. Consider a respondent who thinks that there is a 50% chance that she is a low type. Her probability of obtaining a good job offer is $(.5 \times 0.3) + (.5 \times 1) = 0.65$. A perfectly reasonable response to Weinstein's question is that her chances are 35% above average. Thus, *one* sensible way to answer the question uses the population median, or mode, in determining what an average student is, but the mean of own beliefs for self-evaluating.

Just considering medians and means, there are four ways to interpret answers to (unincentivized) scale questions. It is fairly obvious that in the three cases involving the median, apparent overconfidence will not imply overconfidence, since there is no particular reason for median calculations to average out. Theorem 3, which is a simple consequence of the fact that beliefs are a martingale, concerns the remaining case. It says that when a rational population reports their mean beliefs, these reports must average out to the actual population mean.

Definition 2 *The population scale data* $(\Theta, m, \bar{a},)$ *can be* ***rationalized*** *if there is a rationalizing model* $\left(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta}\right)$ *such that* $m = E(\theta)$ *and* $\bar{a} = \int_\Theta \theta dc$*, where* c *is the probability distribution defined by*

$$c(T) = F\{s : E(\theta \mid s) \in T\} \text{ for } T \subset \Theta.$$

Theorem 3 *Population scale data* $(\Theta, m, \bar{a})$ *can be rationalized if and only if* $\bar{a} = m$.

[16] In some experiments, the scale Θ is not an interval of real numbers, but, say, a set of integers. This may force some subjects to round off their answers, leading to uninteresting complications which we avoid.

Clark and Friesen (2008) reports on a scale experiment in which subjects are incentivized to, in effect, report their mean beliefs relative to the population mean. In keeping with Theorem 3, the experiment finds no apparent overconfidence or underconfidence.[17] Moore and Healy (2008) run a set of incentivized scale experiments which yield no misconfidence in some treatments, and misconfidence in others.

3 Conclusion

Early researchers found a universal tendency towards overplacement. Psychologists and economists developed theories to explain this overplacement and explore its implications. Implicit in these theories was the presumption that a rational population should not overplace itself. We have shown, however, that there is no particular reason for 50% of the population to place itself in the top 50%. At an abstract level, our theory implies that rational populations should display both overplacement and underplacement, and this is what more recent work has uncovered.

Many of the overplacement studies to date have involved experiments that are, in fact, of limited use in testing for overconfidence. Our results point to the type of experimental design that can provide useful data in this regard. In particular, experiments should yield information about the strengths of subjects' beliefs and information beyond rankings relative to the median.[18] If, say, 65% of subjects believe there is at least an 0.7 chance that they rank in the top 40%, the population displays (true) overconfidence. Note, however, that this does not demonstrate that 25% of the subjects are overconfident. In the extreme, as much as 57% of the population could rationally hold such a belief. Thus, the overconfidence of a few can produce quite overconfident looking data and it may be misleading to broadly characterize a population as overconfident. At the same time, 65% of subjects could rationally hold that there is an 0.6 chance they are in the top 40%, so that a slight degree of overconfidence can also lead to quite overconfident looking data.

For the sake of discussion, let us suppose that Svenson's subjects answered his questions using their median beliefs about themselves. Then we have shown that Svenson's Swedish data can be rationalized but that his American data cannot. On one interpretation, we have explained his Swedish data but not his American data. We prefer a different interpretation. Namely, that we have provided a proper framework with which to analyze Svenson's data. This framework shows that his American data displays overconfidence, but that his Swedish data does not.

Some psychologists and behavioural economists may be uneasy with our approach on the prior grounds that individuals do not use Bayes' rule and, for that matter, may not even understand simple probability. Even for these researchers, however, the basic challenge of this paper remains: To indicate why, and in what sense, a finding that a majority of people

[17] In one variant of their experiment, Clark and Friesen find that subjects underestimate their absolute performance.

[18] In line with these requirements, the recent experimental paper of Merkle and Weber (2010) asks for subjects' belief distributions, while Burks et al. (2010) provide incentives designed to elicit modal beliefs, which they then combine with information on actual performance. Prior work by Moore and Healy (2008) uses a quadratic scoring rule to elicit beliefs. Karni (2009) describes a different procedure for eliciting detailed information about subjects' beliefs.

rank themselves above the median is indicative of *overconfidence*. If such a finding does not show overconfidence in a Bayes' rational population, there can be no presumption that it indicates overconfidence in a less rational population. It is, of course, possible that people are not rational, but not overconfident either.

4 Appendix

As was noted before, Theorem 1 is a special case of a Theorem which we present after the following definitions. For each i, let S_i^q denote the set of signals that result in an updated q^{th} percentile in Θ_i:

$$S_i^q = \left\{ s \in S \mid p\left(\cup_{n=i}^k \Theta_n \mid s\right) \geq q \text{ and } p\left(\cup_{n=1}^i \Theta_n \mid s\right) \geq 1-q \right\}.$$

Given a type space $\Theta \subseteq \mathbf{R}$ and a distribution p over Θ, the population ranking data $x \in \Delta^k$ can be **q-rationalized for** (Θ, p) if there is a rationalizing model $\left(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta}\right)$ with $x_j = F\left(S_j^q\right)$, for $j = 1, ..., k$. Note that median-rationalizing is $q-$rationalizing for $q = \frac{1}{2}$.

Theorem 4 *Suppose that $\Theta \subseteq \mathbf{R}$ and p is a distribution over Θ such that $p(\Theta_i) = 1/k$ for all i. For $q \in (0,1)$, the population ranking data $x \in \Delta^k$ can be q-rationalized for (Θ, p) if and only if, for $i = 1, ..., k$,*

$$\sum\nolimits_{j=i}^{k} x_j \;<\; \frac{k-i+1}{qk}, \qquad \textit{and} \tag{6}$$

$$\sum\nolimits_{j=1}^{i} x_j \;<\; \frac{i}{(1-q)\,k}. \tag{7}$$

The proof of Theorem 4 proceeds as follows. Given a type space Θ and prior p, we construct likelihood functions such that every type in a given $k-$ cile i observes signals with the same probability. This allows us to identify every $\theta \in \Theta_i$ with one type in Θ_i, and w.l.o.g. work with a type space $\{\theta_1, ..., \theta_k\}$. The key to $q-$rationalizing a vector x is finding a non-negative matrix $A = (A_{ji})_{j,i=1}^k$ such that $xA = \left(\frac{1}{k}, ..., \frac{1}{k}\right)$, $\sum_{i=1}^k A_{ji} = 1$, and $\sum_{i=1}^j A_{ji} > 1-q$ and $\sum_{i=j}^k A_{ji} > q$, for all j. Then, the matrix A can be interpreted as the rationalizing model that q-rationalizes x as follows. Nature picks (in an iid fashion) for each individual a type θ_i and a signal s_j with probability $x_j A_{ji}$. Each $k-$cile Θ_i then has probability $1/k$ since $xA = \left(\frac{1}{k}, ..., \frac{1}{k}\right)$. The likelihood functions are given by $f_{\theta_i}(s_j) = kx_j A_{ji}$ and row j of A is then the posterior belief after signal s_j. Since $\sum_{i=1}^j A_{ji} > 1-q$ and $\sum_{i=j}^k A_{ji} > q$, and the number of people observing s_j is x_j, the rationalizing model $q-$rationalizes x.

Proof of Theorem 4. Sufficiency for $q-$rationalization.

Step 1. Suppose $q \in (0,1)$ and that $x \in \Delta^k$ is such that inequalities (6) and (7) hold. We show that there exists a non-negative $k \times k$ matrix $A = (A_{ji})_{j,i=1}^k$ such that $xA = \left(\frac{1}{k}, ..., \frac{1}{k}\right)$, and for all j, $\sum_{i=1}^k A_{ji} = 1$, $\sum_{i=1}^j A_{ji} > 1-q$, and $\sum_{i=j}^k A_{ji} > q$.

Pick d such that $\min\left\{\frac{1}{q}, \frac{1}{1-q}, \frac{k+1}{k}\right\} > d > 1$ and for all i,

$$\sum_{j=i}^{k} x_j \leq \frac{k-i+1}{qdk} \qquad \text{and} \qquad \sum_{j=1}^{i} x_j \leq \frac{i}{(1-q)\,dk}. \tag{8}$$

We say that $r \in \Delta^k$ can be justified if there exists a non-negative $k \times k$ matrix R, such that $xR = r$, and for all i, $\sum_{i=1}^{k} R_{ji} = 1$, $\sum_{i=1}^{j} R_{ji} \geq (1-q)\, d$, and $\sum_{i=j}^{k} R_{ji} \geq qd$. Let $\mathcal{R}$ be the set of distributions that can be justified. Note that $\mathcal{R}$ is non-empty, since x itself can be justified by the identity matrix. Furthermore, $\mathcal{R}$ is closed and convex. We now show that $\left(\frac{1}{k}, ..., \frac{1}{k}\right) \in \mathcal{R}$.

Assume all inequalities in (6) and (7) hold, but that $\left(\frac{1}{k}, ..., \frac{1}{k}\right) \notin \mathcal{R}$. Then, since $f(t) = \left\| t - \left(\frac{1}{k}, ..., \frac{1}{k}\right)\right\|^2$ is a strictly convex function, there is a unique r such that $\left(\frac{1}{k}, ..., \frac{1}{k}\right) \neq r = \arg\min_{t \in \mathcal{R}} f(t)$. Let R be a matrix that justifies r.

Since $r \neq \left(\frac{1}{k}, ..., \frac{1}{k}\right)$ there exists some $r_i \neq \frac{1}{k}$, and since $r \in \Delta^k$, there must be some i for which $r_i > \frac{1}{k}$, and some i for which $r_i < \frac{1}{k}$. Let $i^* = \max\left\{i : r_i \neq \frac{1}{k}\right\}$ and $i_* = \min\left\{i : r_i \neq \frac{1}{k}\right\}$.

Part A: We prove that $r_{i^*}, r_{i_*} < \frac{1}{k}$.

Suppose instead that $r_{i^*} > \frac{1}{k}$ (a similar argument establishes that $r_{i_*} < \frac{1}{k}$). Then, for all $i > i^*$, $r_i = \frac{1}{k}$ and for some $i < i^*$, $r_i < \frac{1}{k}$. Let $\tilde{\imath} = \max\left\{i : r_i < \frac{1}{k}\right\}$. We show that for all $i > \tilde{\imath}$ (a) for any j such that $j \leq \tilde{\imath}$ or $j > i$, either $x_j = 0$ or $R_{ji} = 0$; (b) either $x_i = 0$ or $\sum_{g=i}^{k} R_{ig} = dq$.

To see (a) fix an $i' > \tilde{\imath}$ and suppose $x_{j'} > 0$ and $R_{j'i'} > 0$ for some $j' \leq \tilde{\imath}$ or $j' > i'$. Define the matrix $\widetilde{R}$ by $\widetilde{R}_{j'\tilde{\imath}} = R_{j'\tilde{\imath}} + \varepsilon R_{j'i'}$, $\widetilde{R}_{j'i'} = (1-\varepsilon)\, R_{j'i'}$, and for all $(j,i) \notin \{(j',i'), (j',\tilde{\imath})\}$, $\widetilde{R}_{ji} = R_{ji}$. We have

$$\left.\begin{array}{l}
\text{For } j \neq j',\ \sum_{i=1}^{j} \widetilde{R}_{ji} = \sum_{i=1}^{j} R_{ji} \geq d(1-q) \text{ and } \sum_{i=j}^{k} \widetilde{R}_{ji} = \sum_{i=j}^{k} R_{ji} \geq dq \\
\text{If } j' \leq \tilde{\imath},\ \sum_{i=1}^{j'} \widetilde{R}_{j'i} \geq \sum_{i=1}^{j'} R_{j'i} \geq d(1-q) \text{ and } \sum_{i=j'}^{k} \widetilde{R}_{j'i} = \sum_{i=j'}^{k} R_{j'i} + \varepsilon R_{j'i'} - \varepsilon R_{j'i'} \geq dq \\
\text{If } i' < j',\ \sum_{i=1}^{j'} \widetilde{R}_{j'i} = \sum_{i=1}^{j'} R_{j'i} + \varepsilon R_{j'i'} - \varepsilon R_{j'i'} \geq d(1-q) \text{ and } \sum_{i=j'}^{k} \widetilde{R}_{j'i} = \sum_{i=j'}^{k} R_{j'i} \geq dq
\end{array}\right\} (i)$$

For ε sufficiently small, define $\widetilde{r} = x\widetilde{R}$.

We have $\widetilde{r}_{\tilde{\imath}} = r_{\tilde{\imath}} + x_{j'}\varepsilon R_{j'i'}$, $\widetilde{r}_{i'} = r_{i'} - x_{j'}\varepsilon R_{j'i'}$, and for $i \notin \{i', \tilde{\imath}\}$, $\widetilde{r}_i = r_i$. Therefore $\sum_{i=1}^{k} \widetilde{r}_i = \sum_{i=1}^{k} r_i = 1$. For small enough ε, $1 \geq \widetilde{r}_i \geq 0$ for all i, since $x_{j'}, R_{j'i'} > 0$ implies that $r_{i'} > 0$. Hence $\widetilde{r} \in \Delta^k$ and, given (i), $\widetilde{r} \in \mathcal{R}$.

We now show that $f(\tilde{r}) < f(r)$.

$$\begin{aligned}
f(\tilde{r}) - f(r) &= \left(r_{\tilde{\imath}} + x_{j'}\varepsilon R_{j'i'} - \frac{1}{k}\right)^2 - \left(r_{\tilde{\imath}} - \frac{1}{k}\right)^2 + \left(r_{i'} - x_{j'}\varepsilon R_{j'i'} - \frac{1}{k}\right)^2 - \left(r_{i'} - \frac{1}{k}\right)^2 \quad (9)\\
&= (x_{j'}\varepsilon R_{j'i'})^2 + 2x_{j'}\varepsilon R_{j'i'}\left(r_{\tilde{\imath}} - \frac{1}{k}\right) + (x_{j'}\varepsilon R_{j'i'})^2 - 2x_{j'}\varepsilon R_{j'i'}\left(r_{i'} - \frac{1}{k}\right) \\
&= 2(x_{j'}\varepsilon R_{j'i'})\left[x_{j'}\varepsilon R_{j'i'} + r_{\tilde{\imath}} - \frac{1}{k} - r_{i'} + \frac{1}{k}\right] = 2(x_{j'}\varepsilon R_{j'i'})\left[x_{j'}\varepsilon R_{j'i'} + r_{\tilde{\imath}} - r_{i'}\right]
\end{aligned}$$

Recall that $r_{\tilde{\imath}} < \frac{1}{k}$, and since $i' > \tilde{\imath}$, $r_{i'} \geq \frac{1}{k}$. Hence, for ε sufficiently small, $[x_{j'}\varepsilon R_{j'i'} + r_{\tilde{\imath}} - r_{i'}] < 0$. We have a contradiction, since, by definition $r = \arg\min_{t \in \mathcal{R}} f(t)$.

To see (b), suppose that for some $j' > \tilde{\imath}$ we have $x_{j'} > 0$ and $\sum_{g=j'}^{k} R_{j'g} > dq$. Pick some $i' \geq j'$ with $R_{j'i'} > 0$. For ε sufficiently small, define $\widetilde{R}$ by $\widetilde{R}_{j'\tilde{\imath}} = R_{j'\tilde{\imath}} + \varepsilon R_{j'i'}$, $\widetilde{R}_{j'i'} = (1-\varepsilon)\, R_{j'i'}$, and for all $(j,i) \notin \{(j',i'), (j',i')\}$, $\widetilde{R}_{ji} = R_{ji}$. Define $\widetilde{r} = x\widetilde{R}$. As before, $\widetilde{r} \in \mathcal{R}$ and $f(\tilde{r}) < f(r)$, a contradiction.

Given (a) and (b), and recalling the definition of $\tilde{\imath}$, we have

$$\begin{aligned}
\frac{k-\tilde{\imath}}{k} &< \textstyle\sum_{t=\tilde{\imath}+1}^{k} r_t = \sum_{t=\tilde{\imath}+1}^{k}\sum_{j=1}^{k} x_j R_{jt} \\
&= \textstyle\sum_{t=\tilde{\imath}+1}^{k}\sum_{j=\tilde{\imath}+1}^{k} x_j R_{jt} \text{ (by (a), } j \le \tilde{\imath} \text{ implies } x_j = 0 \text{, or } R_{jt} = 0) \\
&= \textstyle\sum_{j=\tilde{\imath}+1}^{k} x_j \sum_{t=\tilde{\imath}+1}^{k} R_{jt} = \sum_{j=\tilde{\imath}+1}^{k} x_j \sum_{t=\tilde{\imath}+1}^{k} R_{jt} \text{ (by (a) } j > t > \tilde{\imath} \Rightarrow x_j = 0 \text{ or } R_{jt} = 0) \\
&= \textstyle\sum_{j=\tilde{\imath}+1}^{k} x_j dq \text{ (by (b) either } x_j = 0 \text{ or } \sum_{t=j}^{k} R_{jt} = dq) \\
&\le \frac{k-\tilde{\imath}}{k} \text{ (by definition of } d \text{ and assumption of the Theorem)}
\end{aligned}$$

Thus, we have a contradiction.

Part B: From Part A, there exists an $\widehat{\imath}$, $i_* < \widehat{\imath} < i^*$, such that $r_{\widehat{\imath}} > \frac{1}{k}$. Since $r_{\widehat{\imath}} = \sum_{j=1}^{k} x_j R_{j\widehat{\imath}}$, for some j^* we must have $R_{j^*\widehat{\imath}} > 0$. We now show that this leads to a contradiction.

Consider a small enough ε.

• Suppose first that for all $j \neq \widehat{\imath}$, $R_{j\widehat{\imath}} = 0$ so that $j^* = \widehat{\imath}$ and $R_{\widehat{\imath}\widehat{\imath}} > 0$. Then, we know that $R_{\widehat{\imath}\widehat{\imath}} x_{\widehat{\imath}} = r_{\widehat{\imath}} > \frac{1}{k} \Rightarrow R_{\widehat{\imath}\widehat{\imath}} > \frac{1}{k}$. If $\sum_{i=1}^{j^*} R_{j^*i} = (1-q)\,d$ and $\sum_{i=j^*}^{k} R_{j^*i} = qd$, we get

$$d = \sum_{i=1}^{j^*} R_{j^*j} + \sum_{i=j^*}^{k} R_{j^*j} = 1 + R_{j^*j^*} = 1 + R_{\widehat{\imath}\widehat{\imath}} > 1 + \frac{1}{k} > d$$

which is a contradiction. Hence we must have $\sum_{i=1}^{j^*} R_{j^*j} > (1-q)\,d$ or $\sum_{i=j^*}^{k} R_{j^*j} > qd$. Suppose therefore that $\sum_{i=1}^{j^*} R_{j^*i} > (1-q)\,d$ (an analogous argument can be made if $\sum_{i=j^*}^{k} R_{j^*i} > qd$). Define $\widetilde{R}$ by $\widetilde{R}_{j^*i^*} = R_{j^*i^*} + \varepsilon R_{j^*j^*}$, $\widetilde{R}_{j^*j^*} = (1-\varepsilon)\,R_{j^*j^*}$, and for all $(j,i) \notin \{(j^*,j^*),\ (j^*,i^*)\}$, $\widetilde{R}_{ji} = R_{ji}$. One can then verify that for small enough ε, for all j, $\sum_{i=1}^{j} \widetilde{R}_{ji} \ge (1-q)\,d$ and $\sum_{i=j}^{k} \widetilde{R}_{ji} \ge qd$. Defining $\tilde{r} = x\widetilde{R}$, we obtain $f(\widetilde{r}) < f(r)$ – a contradiction.

• Suppose instead that $j^* \neq \widehat{\imath}$. If $j^* < \widehat{\imath}$, define $\widetilde{R}$ by $\widetilde{R}_{j^*\widehat{\imath}} = (1-\varepsilon)\,R_{j^*\widehat{\imath}}$, $\widetilde{R}_{j^*i^*} = R_{j^*i^*} + \varepsilon R_{j^*\widehat{\imath}}$, and $\widetilde{R}_{ji} = R_{ji}$ for all $(j,i) \notin \left\{\left(j^*,\widehat{\imath}\right), (j^*,i^*)\right\}$. If $j^* > \widehat{\imath}$, define $\widetilde{R}$ by: $\widetilde{R}_{j^*i_*} = R_{j^*i_*} + \varepsilon R_{j^*\widehat{\imath}}$, $\widetilde{R}_{j^*\widehat{\imath}} = (1-\varepsilon)\,R_{j^*\widehat{\imath}}$, and $\widetilde{R}_{ji} = R_{ji}$ for all $(j,i) \notin \left\{\left(j^*,\widehat{\imath}\right), (j^*,i_*)\right\}$. In either case, for all $j \neq j^*$, $\widetilde{R}_{ji} = R_{ji}$ so $\sum_{i=1}^{j} \widetilde{R}_{ji} \ge d\,(1-q)$ and $\sum_{i=j}^{k} \widetilde{R}_{ji} \ge dq$; for $j = j^*$ if $j^* < \widehat{\imath}$, $\sum_{i=1}^{j} \widetilde{R}_{ji} = \sum_{i=1}^{j} R_{ji} \ge (1-q)\,d$ and $\sum_{i=j}^{k} \widetilde{R}_{ji} = \sum_{i=j}^{k} R_{ji} - \varepsilon R_{j^*\widehat{\imath}} + \varepsilon R_{j^*\widehat{\imath}} \ge dq$; for $i = j^*$ if $j^* > \widehat{\imath}$, $\sum_{i=1}^{j} \widetilde{R}_{ji} = \sum_{i=1}^{j} R_{ji} - \varepsilon R_{j^*\widehat{\imath}} + \varepsilon R_{j^*\widehat{\imath}} \ge (1-q)\,d$ and $\sum_{i=j}^{k} \widetilde{R}_{ji} = \sum_{i=j}^{k} R_{ji} \ge dq$. For $\tilde{r} = x\widetilde{R}$ it is easy to show (as in 9) that $f(\widetilde{r}) < f(r)$ – a contradiction.

Parts A and B show that $\left(\frac{1}{k},...,\frac{1}{k}\right) \in \mathcal{R}$. Let A be the matrix that justifies $\left(\frac{1}{k},...,\frac{1}{k}\right)$.

Step 2. Suppose that q, x, and A are as in Step 1. Given any Θ and p such that $p(\Theta_i) = \frac{1}{k}$ for each i, let $S = \{1,2,...,k\}$ and $f_\theta(j) = kx_jA_{ji}$, for $\theta \in \Theta_i$, $i,j = 1,...,k$. To complete the proof of sufficiency, we show that (Θ, S, f, p) q-rationalizes x for (Θ, p); that is, $x_j = F(S_j)$.

1) $x_j = F(j)$, since

$$F(j) = \left(\textstyle\sum_{i=1}^{k} kx_jA_{ji}\right)\frac{1}{k} = \sum_{i=1}^{k} x_jA_{ji} = x_j\sum_{i=1}^{k} A_{ji} = x_j$$

2) $j \in S_j$ since

$$p(\Theta_i \mid j) = \frac{k x_j A_{ji} \frac{1}{k}}{x_j} = A_{ji},$$

$\sum_{i=1}^{j} A_{ji} > 1-q$, and $\sum_{i=j}^{k} A_{ji} > q$.

3) $g \neq j \Rightarrow g \notin S_j$. Suppose $g > j$. We have, $\sum_{i=g}^{k} A_{gi} > q \Rightarrow \sum_{i=1}^{g-1} A_{gi} < 1-q \Rightarrow \sum_{i=1}^{j} A_{gi} < 1-q$, so that $g \notin S_j$. Similarly, $j > g$ implies $g \notin S_j$.

(1), (2) and (3) establish that $x_j = F(j)$.

Necessity. Suppose data x can be $q-$rationalized for some (Θ, p), and let $\left(\Theta, p, S, \{f_\theta\}_{\theta\in\Theta}\right)$ be the rationalizing model. Fix an i, and let $S^i = \cup_{g=i}^{k} S_g^q$. For each signal $s \in S^i$, $p\left(\cup_{g=i}^{k}\Theta_g \mid s\right) \geq q$. If $\sum_{j=i}^{k} x_j = 0$, inequalities (6) (and 7) hold trivially, so suppose $\sum_{j=i}^{k} x_j > 0$. Since $x_j = F\left(S_j^q\right)$ for all j, we have $F\left(S_j^q\right) > 0$, for some $i \leq j \leq k$.

For any j, let $T_j = \left\{s \in S_j^q : p\left(\cup_{i=j}^{k}\Theta_i \mid s\right) = q\right\}$. For any $s \in T_1$, we have $1 = p(\Theta \mid s) = p\left(\cup_{i=1}^{k}\Theta_i \mid s\right) = q < 1$, so we must have $T_1 = \emptyset$. Assume now $j \geq 2$. For any $s \in T_j$, $p\left(\cup_{i=1}^{j-1}\Theta_i \mid s\right) = 1-q$, so that $s \in S_{j-1}^q$. Thus, $s \in T_j$ implies $s \in S_j^q$ and $s \in S_{j-1}^q$. If $F(T_j) > 0$, then $F\left(S_j^q \cup S_{j-1}^q\right) < F\left(S_j^q\right) + F\left(S_{j-1}^q\right)$ so that

$$\begin{aligned}
1 &= F(S) \leq F\left(\cup_{g\neq j,j-1} S_g^q\right) + F\left(S_j^q \cup S_{j-1}^q\right) < F\left(\cup_{g\neq j,j-1} S_g^q\right) + F\left(S_j^q\right) + F\left(S_{j-1}^q\right) \\
&\leq \sum_{g\neq j,j-1} x_g + x_j + x_{j-1} = 1, \text{ a contradiction.}
\end{aligned}$$

Thus, for all j, $F(T_j) = 0$, and for almost every $s \in S^i$, $p\left(\cup_{g=i}^{k}\Theta_g \mid s\right) > q$. Hence,

$$\begin{aligned}
\frac{k-i+1}{k} &= p\left(\cup_{g=i}^{k}\Theta_g\right) = \int p\left(\cup_{g=i}^{k}\Theta_g \mid s\right) dF(s) \\
&\geq \int_{S^i} p\left(\cup_{g=i}^{k}\Theta_g \mid s\right) dF(s) > \int_{S^i} q\, dF(s) = q\textstyle\sum_{j=i}^{k} x_j
\end{aligned}$$

A similar argument applies for inequalities in (7). ■

Proof of the corollary. First note that, for any $x', x \in \Delta^k$, if x' fosd $x \neq x'$, then $\mu(x') > \mu(x)$. Let M be the set of median-rationalizable vectors. By Theorem 1, M is characterized by a set of linear inequalities, so M is a convex set. Suppose k is even. From Theorem 1 $\sup_{x\in M} \mu(x) = \mu\left(0,..,0,\frac{2}{k},...,\frac{2}{k}\right) = \sum_{i=\frac{k}{2}+1}^{k} \frac{2}{k} i = \frac{3}{4}k + \frac{1}{2}$ and $\inf_{x\in M} \mu(x) = \mu\left(\frac{2}{k},..,\frac{2}{k},0,...,0\right) = \frac{1}{4}k + 2$. Moreover, these bounds are not attained because neither $\left(0,..,0,\frac{2}{k},...,\frac{2}{k}\right)$ nor $\left(\frac{2}{k},..,\frac{2}{k},0,...,0\right)$ are in M. Since M is convex, for any $t \in \left(\frac{1}{4}k+2, \frac{3}{4}k+\frac{1}{2}\right)$, there exists an $x \in M$ such that $\mu(x) = t$. Similar reasoning applies to k odd. ■

The proof of Theorem 2 is constructive. For each $x \in \Delta^k$, $x \gg 0$ that satisfies (4), we construct $\widetilde{x} = \frac{1}{a}x - \frac{1-a}{a}\left(\frac{1}{k},...,\frac{1}{k}\right)$ for a arbitrarily close to 1. Then $\widetilde{x} \in \Delta^k, \widetilde{x} \gg 0$, and $\widetilde{x}$ satisfies the inequalities in (4). Then we find a z comparable to $\widetilde{x}$ and a non-negative matrix $P = (P_{ji})_{j,i=1}^{k}$ such that for $i, j = 1,...,k$, $\sum_{j=1}^{k} P_{ji} = \frac{1}{k}$, $\sum_{i=1}^{k} P_{ji} = z_j$, $\frac{1}{z_j}\sum_{i=1}^{j} P_{ji} \geq \frac{1}{2}$, and $\frac{1}{z_j}\sum_{i=j}^{k} P_{ji} \geq \frac{1}{2}$. Moreover, we construct P so that it satisfies certain dominance relations. As in the proof of Theorem 1, the matrix P embodies the likelihood functions f, through $f_\theta(S_j) = kP_{ji}$ for $\theta \in \Theta_i$ and it yields a rationalizing model which almost rationalizes

z; *almost* because for rationalization we would need *strict* inequalities $\frac{1}{z_j}\sum_{i=1}^{j}P_{ji}>\frac{1}{2}$, and $\frac{1}{z_j}\sum_{i=j}^{k}P_{ji}>\frac{1}{2}$. Since z is comparable to $\widetilde{x}$, the vector $y=az+(1-a)\left(\frac{1}{k},...,\frac{1}{k}\right)$ is comparable to x and is rationalized by the model yielded by $Q=aP+(1-a)\frac{1}{k}I$ (where $\frac{1}{y_j}\sum_{i=1}^{j}Q_{ji}>\frac{1}{2}$ and $\frac{1}{y_j}\sum_{i=j}^{k}Q_{ji}>\frac{1}{2}$). From the dominance relations that P satisfies, the model has monotone signals.

Proof of Theorem 2. **Sufficiency.** For any matrix P, let P^i denote the ith column and P_j denote the j^{th} row. Claim 1: For any $\widetilde{x}\in\Delta^k$, $\widetilde{x}\gg 0$ that satisfies (4), there exists a comparable z and a non-negative matrix $P=(P_{ji})_{j,i=1}^{k}$ such that for $i,j=1,...,k$, $\sum_{j=1}^{k}P_{ji}=\frac{1}{k}$, $\sum_{i=1}^{k}P_{ji}=z_j$, $\frac{1}{z_j}\sum_{i=1}^{j}P_{ji}\geq\frac{1}{2}$, and $\frac{1}{z_j}\sum_{i=j}^{k}P_{ji}\geq\frac{1}{2}$. Moreover, for all i,j, kP^{i+1} fosd kP^i, and

$$\sum_{r=i+1}^{k}\frac{P_{jr}}{z_j}\leq\sum_{r=i+1}^{k}\frac{P_{j+1,r}}{z_{j+1}} \tag{10}$$

with strict inequality if $\sum_{r=i+1}^{k}P_{j+1,r}>0$.

We prove the claim for k even; the proof for k odd is similar. In Part A, we build rows $h,...,k$ of P and in Part B we build rows $1,...,h-1$. In Part C, we verify that the matrix P satisfies the dominance relations.

Part A. $k\geq j\geq h$.

Let $z_j=\widetilde{x}_j$, and define P_k by $P_{kk}=z_k/2$, and $P_{ki}=z_k/2\,(k-1)$ for $i=1,...,k-1$.

We now build recursively P_{k-1} through P_h. Whenever rows $P_{j+1},...,P_k$ have been defined, define the "slack" vector $s^j\in\mathbf{R}^k$ by $s_i^j=\frac{1}{k}-\sum_{f=j+1}^{k}P_{fi}$. Note that $s_i^j=s_i^{j+1}-P_{ji}$.

Suppose that for every $r\geq j+1$: (i) $\sum_{i=r}^{k}P_{ri}=\frac{z_r}{2}$; (ii) for $1\leq i<r$, $P_{ri}=\frac{z_r}{2(r-1)}$, (iii) $P_{ri}\geq 0$ for all i and P_{ri} is decreasing in i for $i\geq r$, and (iv) $s_i^{r-1}\geq 0$ for all i, $s_i^{r-1}=s_{r-1}^{r-1}$ for $i<r$, and s_i^{r-1} decreasing in i for $i\geq r-1$. Notice that for $j=k-1$ (i)-(iv) are satisfied. We now build P_j in such a way that (i)-(iv) are satisfied for $r=j$.

For $i<j$, set $P_{ji}=z_j/2\,(j-1)$. We turn now to $i\geq j$.

Let $i_j=\max\left\{i\geq j:\frac{z_j}{2}-\sum_{f=i+1}^{k}s_f^j\leq s_i^j\,(i-j+1)\right\}$. We first establish that i_j is well-defined. By the induction hypothesis, for every $r\geq j+1$, $\sum_{i=j}^{k}P_{ri}=\frac{z_r}{2}+\frac{z_r}{2(r-1)}(r-j)=\frac{z_r}{2}\frac{2r-j-1}{r-1}$. We have

$$\begin{aligned}\sum_{i=j}^{k}s_i^j &= \frac{1}{k}(k-j+1)-\sum_{i=j}^{k}\sum_{r=j+1}^{k}P_{ri}=\frac{1}{k}(k-j+1)-\sum_{r=j+1}^{k}\sum_{i=j}^{k}P_{ri}\\ &= \frac{1}{k}(k-j+1)-\sum_{r=j+1}^{k}\frac{z_r}{2}\frac{2r-j-1}{r-1}>\frac{z_j}{2}\end{aligned} \tag{11}$$

where the inequality holds since, by (4), $\frac{1}{k}(k-j+1)>\sum_{r=j}^{k}\frac{z_r}{2}\frac{2r-j-1}{f-1}$. Therefore, $i=j$ satisfies the conditions in the definition of i_j and i_j is well-defined. Define

$$P_{ji}=\left\{\begin{array}{cc}\frac{\frac{z_j}{2}-\sum_{f=i_j+1}^{k}s_f^j}{i_j-j+1} & \text{for } i_j\geq i\geq j\\ s_i^j & \text{for } i>i_j\end{array}\right.. \tag{12}$$

Items (i) and (ii) are satisfied by construction. We now check (iii). Clearly $P_{ji}\geq 0$. Notice that P_{ji} is constant in i, for $j\leq i\leq i_j$, and decreasing in i for $i>i_j$ since s_i^j is decreasing, so that we only need to check that $P_{ji_j} \geq P_{j,i_j+1}$. But $P_{ji_j} < P_{j,i_j+1}$ would imply

$$\frac{\frac{z_j}{2} - \sum_{f=i_j+1}^{k} s_f^j}{i_j - j + 1} < s_{i_j+1}^j \Leftrightarrow \frac{z_j}{2} - \sum_{f=i_j+2}^{k} s_f^j < s_{i_j+1}^j \left(i_j - j + 2\right)$$

which violates the definition of i_j.

To establish (iv), we first show $s_i^{j-1} \geq 0$ for all i. By definition of P_{ji}, (a) $s_i^{j-1} = 0$ for all $i > i_j$. By definition of i_j, $P_{ji_j} \leq s_{i_j}^j$ so that (b) $s_{i_j}^{j-1} \geq 0$. Next, $P_{ji} = P_{ji_j}$ for $i_j \geq i \geq j$ and s_i^j decreasing in i establishes (c) $s_i^{j-1} \geq s_{i_j}^{j-1} \geq 0$ for $i_j \geq i \geq j$. Consider now $i < j$. Since $j \geq h = 1 + k/2$, we obtain $P_{jj} \geq z_j/2\left(k - j + 1\right) \geq z_j/2\left(j - 1\right) = P_{ji}$ for all $i < j$. Then, $s_i^j = s_j^j$ for all $i < j$, and $s_j^{j-1} \geq 0$ (established in (c)) imply (d) for all $i < j$

$$s_i^{j-1} = s_{j-1}^{j-1} \geq s_j^{j-1} \geq 0 \text{ and } s_i^{j-1} = s_{j-1}^{j-1} > s_j^{j-1} \geq 0 \text{ if } j > h. \tag{13}$$

Next, notice that $s_i^j = s_j^j$ and $P_{ji} = P_{j,j-1}$ for all $i < j$ establish $s_i^{j-1} = s_{j-1}^{j-1}$ for $i < j$.

We finally show that s_i^{j-1} is decreasing in i for $i \geq j-1$. Equation 13 established that $s_{j-1}^{j-1} \geq s_j^{j-1}$, so we only need to show that $s_i^{j-1} \geq s_{i'}^{j-1}$ for any $i' > i \geq j$. If $P_{ji} > P_{ji'}$ for some $i' > i \geq j$, then from 12 $P_{ji'} = s_{i'}^j$ and therefore $s_{i'}^{j-1} = 0 \leq s_i^{j-1}$. If $P_{ji} \leq P_{ji'}$, then since $s_i^j \geq s_{i'}^j$,

$$s_i^{j-1} = s_i^j - P_{ji} \geq s_{i'}^j - P_{ji'} = s_{i'}^{j-1}.$$

This completes the proof of Part A.

We now establish a property of P that will be used in Part B. Suppose that for some j and $i \geq j$, we have $P_{ji} = 0$. If $P_{ji} \neq s_i^j$, (12) implies $P_{ji'} = P_{ji} = 0$ for all $j \leq i' \leq i_j$. Also, (iii) implies that for all $i' \geq i$, $P_{ji'} = 0$. Then, (i) implies $z_j = 0$ which is a contradiction. Therefore $P_{ji} = s_i^j = 0$. From (iii) and (iv) we obtain $P_{ji'} = s_{i'}^j = 0$ for all $i' \geq i$ so that $s_{i'}^{j-1} = s_{i'}^j - P_{ji'} = 0$. Since $s_{i'}^{j-2} \geq 0$, we have $P_{j-1,i'} = 0$. Repeating the reasoning we obtain

$$P_{ji} = 0 \Rightarrow s_i^j = 0 \Rightarrow P_{j'i'} = s_{i'}^{j'} = 0 \text{ for all } i' \geq i \text{ and all } j' \leq j. \tag{14}$$

Part B. $j < h$.

For $j \geq h$, $z_j = \widetilde{x}_j$. For $j < h$, z_j can be different from $\widetilde{x}_j$; we proceed by defining P_j, and then setting $z_j = \sum_{i=1}^{k} P_{ji}$.

Let $P_{h-1,i} = s_i^{h-1}$ for $i \geq h$ and for $i \leq h-1$, define $P_{h-1,i}^1 = \sum_{g=h}^{k} s_g^{h-1}/\left(h-1\right)$.

We now show that $s_i^{h-1} - P_{h-1,i}^1 > 0$ for $i \leq h-1$. We will establish $s_{h-1}^{h-1} - P_{h-1,h-1}^1 > 0$, since $P_{h-1,i}^1 = P_{h-1,h-1}^1$ and $s_{h-1}^{h-1} = s_i^{h-1}$ for all $i \leq h-1$.

$P_{hh} > P_{h,h+1}$ implies $i_h = h$ and $P_{h,h+1} = s_{h+1}^h$, which ensures $s_{h+1}^{h-1} = 0$. In equation (11) we proved $\sum_{i=j}^{k} s_i^j > \frac{z_j}{2}$ for all j, so from equation (12) and $h = i_h$ we obtain

$$\sum_{i=h}^{k} s_i^h > \frac{z_h}{2} \Leftrightarrow s_h^h > \frac{z_h}{2} - \sum_{f=h+1}^{k} s_f^h = P_{hh}$$

and therefore $s_h^{h-1} = s_h^h - P_{hh} > 0 = s_{h+1}^{h-1}$. Hence $P_{hh} > P_{h,h+1}$ implies $s_h^{h-1} > s_{h+1}^{h-1}$. Also, $P_{hh} = P_{h,h+1}$, implies $s_h^{h-1} > s_{h+1}^{h-1}$ because by the strict inequality in (13), $s_h^h > s_{h+1}^h$. Since $P_{hh} \geq P_{h,h+1}$, we obtain $s_h^{h-1} > s_{h+1}^{h-1}$. Since s_i^{h-1} is decreasing in i,

$$P_{h-1,i}^1 = \sum_{g=h}^{k} \frac{s_g^{h-1}}{h-1} = \sum_{g=h}^{k} \frac{s_g^{h-1}}{k-h+1} < s_h^{h-1}$$

for all $i \leq h-1$, and therefore $s_h^{h-1} > s_{h+1}^{h-1} \Rightarrow s_h^{h-1} - P_{h-1,i}^1 > s_{h+1}^{h-1} - s_h^{h-1} = 0$ as was to be shown.

Define $P_{h-1,h-1}^2 = s_{h-1}^{h-1} - P_{h-1,h-1}^1 > 0$ and $P_{h-1,j}^2 = P_{h-1,h-1}^2 / (h-2)$ for all $j < h-1$ and let $P_{h-1,i} = P_{h-1,i}^1 + P_{h-1,i}^2$ for $i \leq h-1$.

Let $z_{h-1} = \sum_{i=1}^k P_{h-1,i} = 2\sum_{i=h}^k s_i^{h-1} + 2P_{h-1,h-1}^2$. We have,

$$\frac{\sum_{i=1}^{h-1} P_{h-1,i}}{z_{h-1}} = \frac{\sum_{j=h}^k s_j^{h-1} + 2P_{h-1,h-1}^2}{2\sum_{j=h}^k s_j^{h-1} + 2P_{h-1,h-1}^2} \geq \frac{1}{2}$$

and

$$\frac{\sum_{i=h-1}^{k} P_{h-1,i}}{z_{h-1}} = \frac{P_{h-1,h-1}^1 + P_{h-1,h-1}^2 + \sum_{i=h}^k s_i^{h-1}}{z_{h-1}} \geq \frac{1}{2}.$$

For $j < h-1$, set $P_{ji} = 0$ for $i > j$, $P_{jj} = s_j^j$ and $P_{ji} = P_{jj}/(j-1)$ for $i < j$, and $z_j = \sum_{i=1}^k P_{ji}$.

Part C. Checking dominance relations.

We first check inequality (10).

Case 1. P_j and P_{j+1}, $j \geq h$.

(I) For $i < j$, we have

$$\frac{\sum_{r=1}^i P_{jr}}{z_j} = \frac{i\frac{z_j}{2(j-1)}}{z_j} = \frac{i}{2(j-1)} > \frac{i}{2j} = \frac{i\frac{z_{j+1}}{2j}}{z_{j+1}} = \frac{\sum_{r=1}^i P_{j+1,r}}{z_{j+1}}. \tag{15}$$

(II) For $i = j$, since $P_{jj} > P_{jr} = \frac{z_j}{2(j-1)}$ and $P_{j+1,j} = P_{j+1,r} = \frac{z_{j+1}}{2j}$ for $r < j$, we have

$$\frac{\sum_{r=1}^i P_{jr}}{z_j} = \frac{\sum_{r=1}^j P_{jr}}{z_j} > \frac{j\frac{z_j}{2(j-1)}}{z_j} > \frac{1}{2} = \frac{\sum_{r=1}^j P_{j+1,r}}{z_{j+1}} = \frac{\sum_{r=1}^i P_{j+1,r}}{z_{j+1}}. \tag{16}$$

(III.a) Pick $i \geq j+1$ and suppose $P_{j+1,i+1} = 0$. By (14) we have that for $r \geq i+1$, $P_{jr} = 0$, and therefore

$$\frac{\sum_{r=1}^i P_{jr}}{z_j} = 1 \geq \frac{\sum_{r=1}^i P_{j+1,r}}{z_{j+1}}. \tag{17}$$

(III.b) Pick $i \geq j+1$ and suppose $P_{j+1,i+1} > 0$. If $i+1 > i_{j+1}$ then $P_{j+1,i+1} = s_{i+1}^{j+1}$, so that $s_{i+1}^j = 0$. By (14) $P_{jr} = 0$ for all $r \geq i+1$, so that

$$\sum_{r=i+1}^k \frac{P_{jr}}{z_j} = 0 < P_{j+1,i+1} \leq \sum_{r=i+1}^k \frac{P_{j+1,r}}{z_{j+1}}.$$

If $i+1 \leq i_{j+1}$ then, since P_{jr} and $P_{j+1,r}$ are decreasing in r, we have the following ordering between distributions: the distribution $2(P_{j+1,j+1}, ..., P_{j+1,k})/z_{j+1}$ fosd the uniform distribution on $j+1$ to i_{j+1}; the uniform from j to i_{j+1} fosd the distribution $2(P_{jj}, ..., P_{jk})/z_j$. Therefore,

$$\frac{2\sum_{r=i+1}^k P_{j+1,r}}{z_{j+1}} \geq \frac{i_{j+1} - i}{i_{j+1} - j} > \frac{i_{j+1} - i}{i_{j+1} - j + 1} \geq \frac{2\sum_{r=i+1}^k P_{jr}}{z_j}.$$

Note that because $P_{j+1,i}$ is decreasing in i for $i \geq j+1$, $\sum_{r=i+1}^{k} P_{j+1,r} > 0 \Leftrightarrow P_{j+1,i+1} > 0$. Therefore, (I), (II), (III.a) and (III.b) show that (10) holds, and that the inequality is strict if $P_{j+1,i+1} > 0$.

Case 2. P_j and P_{j+1}, $j = h-1$

(I) For $i < j$, recall from Part B that for all $i < h-1$,

$$P_{h-1,i} = P_{h-1,i}^{1} + P_{h-1,i}^{2} = \frac{\sum_{r=h}^{k} s_r^{h-1}}{h-1} + \frac{P_{h-1,h-1}^{2}}{h-2}.$$

Then, letting $a = \sum_{r=h}^{k} s_r^{h-1}$ and $b = P_{h-1,h-1}^{2}$, and recalling that $z_{h-1} = 2a + 2b$, we have that for $i < h-1$,

$$\frac{P_{h-1,i}}{z_{h-1}} = \frac{\frac{a}{h-1} + \frac{b}{h-2}}{2a+2b}. \tag{18}$$

Since, for $i < h-1$, $P_{hi} = z_h/2\,(h-1)$, and $b = P_{h-1,h-1}^{2} > 0$, we have

$$\frac{\sum_{r=1}^{i} P_{h-1,r}}{z_{h-1}} = i\frac{\frac{a}{h-1} + \frac{b}{h-2}}{2a+2b} > i\frac{\frac{a}{h-1} + \frac{b}{h-1}}{2a+2b} = \frac{i}{2\,(h-1)} = \frac{\sum_{r=1}^{i} P_{h,r}}{z_h}.$$

(II) For $i = j$,

$$\frac{\sum_{r=1}^{i} P_{h-1,r}}{z_{h-1}} = \frac{\sum_{r=1}^{h-1} P_{h-1,r}}{z_{h-1}} = \frac{a+2b}{2a+2b} > \frac{1}{2} = \frac{\sum_{r=1}^{h-1} P_{hr}}{z_h} = \frac{\sum_{f=1}^{i} P_{hr}}{z_h}$$

(III) Fix $i > j$. If $P_{j+1,i+1} = 0$, repeat step (III.a) of Case 1 to show $\sum_{r=1}^{i} \frac{P_{jr}}{z_j} = 1 \geq \sum_{r=1}^{i} \frac{P_{j+1,r}}{z_{j+1}}$. If $P_{j+1,i+1} > 0$, repeat step (III.b) of Case 1 to show that $\sum_{f=i+1}^{k} P_{h-1,f}/z_{h-1} < \sum_{f=i+1}^{k} P_{hr}/z_h$ for $i > h-1$.

(I), (II) and (III) show that (10) holds, and that the inequality is strict if $\sum_{r=i+1}^{k} P_{j+1,r} > 0$.

Case 3. P_j and P_{j+1}, for $j = h-2$.

(I) For $i < j$, recall equation (18) and that $P_{h-2,r}/z_{h-2} = 1/2\,(h-3)$ for $r \leq i$, so that

$$\sum_{r=1}^{i} \frac{P_{h-2,r}}{z_{h-2}} = \frac{i}{2\,(h-3)} > \frac{i}{2\,(h-2)} > i\frac{\frac{a}{h-1} + \frac{b}{h-2}}{2a+2b} = \sum_{r=1}^{i} \frac{P_{h-1,r}}{z_{h-1}}.$$

(II) for $i = j$,

$$\sum_{r=1}^{i} \frac{P_{h-2,r}}{z_{h-2}} = 1 > 1 - P_{h-1,h-1} \geq 1 - \sum_{r=h-1}^{k} \frac{P_{h-1,r}}{z_{h-1}} = \sum_{r=1}^{i} \frac{P_{h-1,r}}{z_{h-1}}$$

(III) for $i > j$, $P_{h-2,i} = 0$ for all $i \geq h-1$, so $\sum_{r=i+1}^{k} \frac{P_{h-2,r}}{z_{h-2}} = 0$. If $\sum_{r=i+1}^{k} P_{h-1,r} > 0$, we have $\sum_{r=i+1}^{k} \frac{P_{h-1,r}}{z_{h-1}} > \sum_{r=i+1}^{k} \frac{P_{h-2,r}}{z_{h-2}}$.

(I), (II) and (III) show that (10) holds, and that the inequality is strict if $\sum_{r=i+1}^{k} P_{j+1,r} > 0$.

Case 4. P_j and P_{j+1}, for $j < h-2$.

These cases are trivial, since $P_{gi} = \frac{z_g}{2(g-1)}$ for $i < g$, $P_{gg} = \frac{z_g}{2}$ and $P_{gi} = 0$ for $i > g$, for $g = j, j+1$.

We now check that P^{i+1} fosd P^i.

Suppose $i \geq h$. Since, $P_{kk} \geq P_{kc} = P_{kc'}$ for all $c, c' < k$, we have $\sum_{r=j}^{k} P_{r,i+1} \geq \sum_{r=j}^{k} P_{ri}$ for $j = k$. Fix then $h - 1 \leq j < k$.

• If $P_{j'i} > P_{j',i+1}$, for any $j \leq j' < k$, we know $P_{j',i+1} = s_{i+1}^{j'}$ and therefore $s_{i+1}^{j'-1} = 0$, and by (14) $P_{r,i+1} = 0$ for all $r \leq j' - 1$. This implies $\sum_{r=j}^{k} P_{r,i+1} \geq \sum_{r=j'}^{k} P_{r,i+1} = 1 \geq \sum_{r=j}^{k} P_{r,i}$.

• If $P_{j'i} \leq P_{j',i+1}$ for all $j \leq j' < k$, since $P_{kk} \geq P_{kc} = P_{kc'}$ for all $c, c' < k$, we have $\sum_{r=j}^{k} P_{r,i+1} \geq \sum_{r=j}^{k} P_{ri}$.

Fix $j < h - 1$. Since $P_{ji} = P_{j,i+1} = 0$, we have $1 = \sum_{r=j}^{k} P_{r,i+1} \geq \sum_{r=j}^{k} P_{ri}$.

Suppose $i = h - 1$. For all $j > h$, $P_{jh} = P_{j,h-1}$, and for $j = h$ we have $P_{hh} \geq P_{h,h-1}$, so that for all $j \geq h$, $\sum_{r=j}^{k} P_{rh} \geq \sum_{r=j}^{k} P_{r,h-1}$. Also, since $P_{h-1,h} = s_h^{h-1}$ and $P_{jh} = 0$ for $j < h - 1$ we have that for all $j < h$, $\sum_{r=j}^{k} P_{rh} = \frac{1}{k} \geq \sum_{r=j}^{k} P_{r,h-1}$.

Suppose $i \leq h - 2$. For all $j > i + 1$, $P_{j,i+1} = P_{ji}$, and for $j = i + 1$ we have $P_{j,i+1} \geq P_{ji}$, so that for all $j \geq i + 1$, $\sum_{r=j}^{k} P_{r,i+1} \geq \sum_{r=j}^{k} P_{ri}$. Also, since $P_{i+1,i+1} = s_{i+1}^{i+1}$ we have, for all $j < i + 1$, $\sum_{r=j}^{k} P_{r,i+1} = \frac{1}{k} \geq \sum_{r=j}^{k} P_{ri}$.

This establishes Claim 1.

Now take any vector $x \in \Delta^k, x \gg 0$, that satisfies (4), and for the $k \times k$ identity matrix I, let $J = \frac{1}{k} I$ and $K = \left(\frac{1}{k}, ..., \frac{1}{k}\right)$. We find a y comparable to x that can be rationalized with monotone signals.

Let $\widetilde{x} = \frac{1}{a} x - \frac{1-a}{a} K$ for a arbitrarily close to 1. Then $\widetilde{x} \in \Delta^k, \widetilde{x} \gg 0$, and $\widetilde{x}$ satisfies the inequalities in (4). Find a z and P as in Claim 1.

Let $y = az + (1-a) K$ and $Q = aP + (1-a) J$. Define the matrix A by $A_j = Q_j / \sum_{i=1}^{k} Q_{ji}$. Since $\sum_{i=1}^{k} P_{ji} = z_j$ and $\sum_{i=1}^{k} J_{ji} = \frac{1}{k}$, we have that $\sum_{i=1}^{k} Q_{ji} = y_j = az_j + (1-a) \frac{1}{k}$ and $A_j = Q_j / y_j$.

Note that y is comparable to x, $yA = \left(\frac{1}{k}, ..., \frac{1}{k}\right)$, and for all j, $\sum_{i=1}^{k} A_{ji} = 1$. For $1 \leq j \leq k$

$$\begin{aligned}
\textstyle\sum_{i=1}^{j} A_{ji} &= \frac{\sum_{i=1}^{j} Q_{ji}}{y_j} = \frac{\sum_{i=1}^{j} \left(aP_{ji} + (1-a) J_{ji}\right)}{az_j + (1-a)\frac{1}{k}} = \frac{az_j}{az_j + (1-a)\frac{1}{k}} \frac{\sum_1^j P_{ji}}{z_j} + \frac{(1-a)\frac{1}{k}}{az_j + (1-a)\frac{1}{k}} \frac{\sum_1^j J_{ji}}{\frac{1}{k}} \\
&= \frac{az_j}{az_j + (1-a)\frac{1}{k}} \frac{\sum_1^j P_{ji}}{z_j} + \frac{(1-a)\frac{1}{k}}{az_j + (1-a)\frac{1}{k}} \geq \frac{az_j}{az_j + (1-a)\frac{1}{k}} \frac{1}{2} + \frac{(1-a)\frac{1}{k}}{az_j + (1-a)\frac{1}{k}} > \frac{1}{2}.
\end{aligned}$$

Similarly, $\sum_{i=j}^{k} A_{ji} > y_j/2$.

Given any Θ and p such that $p(\Theta_i) = \frac{1}{k}$, for all i, let $S = \{1, 2, ..., k\}$ and $f_\theta(j) = kA_{ji} y_j$ for $\theta \in \Theta_i$, $i, j = 1, ..., k$. From Step 2 in the proof of Theorem 1, (Θ, S, f, p) q-rationalizes y for $q = \frac{1}{2}$.

We now verify that (Θ, S, f, p) satisfies msp. It is immediate that $f_{\theta'}$ fosd f_θ for $\theta' > \theta$, since kP^{i+1} fosd kP^i for all i. We need to show that for all i, $j < k$, $p\left(\cup_{g=1}^{i} \Theta_g \mid j\right) \geq p\left(\cup_{g=1}^{i} \Theta_g \mid j+1\right)$, which is true if and only if $\sum_{g=1}^{i} A_{jg} \geq \sum_{r=1}^{i} A_{j+1,g}$.

If $\sum_{r=i+1}^{k} P_{j+1,r} = 0$, we have $\sum_{r=1}^{i} P_{j+1,r} = z_{j+1}$ and by (10) $\sum_{r=1}^{i} P_{jr} = z_j$. If $\sum_{r=i+1}^{k} P_{j+1,r} = 0$ we must also have $i \geq j+1$, since for $i < j+1$, $\sum_{r=1}^{i} P_{j+1,r} = z_{j+1}$

would imply

$$0 = \sum\nolimits_{r=i+1}^{k} \frac{P_{j+1,r}}{z_{j+1}} \geq \sum\nolimits_{r=j+1}^{k} \frac{P_{j+1,r}}{z_{j+1}} \geq \frac{1}{2}.$$

We therefore have

$$\begin{aligned}
\sum\nolimits_{r=1}^{i} A_{jr} &= \sum\nolimits_{r=1}^{i} \frac{Q_{jr}}{y_j} = \frac{az_j}{az_j + (1-a)\frac{1}{k}} \sum\nolimits_{r=1}^{i} \frac{P_{jr}}{z_j} + \frac{(1-a)\frac{1}{k}}{az_k + (1-a)\frac{1}{k}} \frac{\sum_{r=1}^{i} J_{jr}}{\frac{1}{k}} \\
&= \frac{az_j}{az_j + (1-a)\frac{1}{k}} + \frac{(1-a)\frac{1}{k}}{az_j + (1-a)\frac{1}{k}} = 1 \\
&= \frac{az_{j+1}}{az_{j+1} + (1-a)\frac{1}{k}} + \frac{(1-a)\frac{1}{k}}{az_j + (1-a)\frac{1}{k}} = \sum\nolimits_{r=1}^{i} A_{j+1,r}
\end{aligned}$$

If $\sum_{r=i+1}^{k} P_{j+1,r} > 0$, then by (10) $\sum_{r=1}^{i} \frac{P_{jr}}{z_j} > \sum_{r=1}^{i} \frac{P_{j+1,r}}{z_{j+1}}$, and for a sufficiently close to 1, $\sum_{r=1}^{i} A_{jr} > \sum_{r=1}^{i} A_{j+1,r}$.

This completes the proof of sufficiency for $k > 4$.

For $k = 4$, suppose that $x \in \Delta^4, x \gg 0$ satisfies (4). Then $x_3 + \frac{4}{3}x_4 < 1$. Assume wlog that $x_3 + x_4 \geq x_1 + x_2$ and define the matrices P and P':

$$P \qquad\qquad\qquad\qquad\qquad\qquad\qquad P'$$

$$\begin{array}{cccc}
\frac{x_1}{2} + \varepsilon & \frac{x_1}{2} - \varepsilon & 0 & 0 \\
\frac{x_2}{2} - \varepsilon & \frac{x_2}{2} + \varepsilon & 0 & 0 \\
\frac{x_3}{4} & \frac{x_3}{4} - \varepsilon & \frac{1-2x_1-2x_2}{4} + 2\varepsilon & \frac{1-2x_4}{4} - \varepsilon \\
\frac{x_4-x_1-x_2}{4} & \frac{x_4-x_1-x_2}{4} + \varepsilon & \frac{x_1+x_2}{2} - 2\varepsilon & \frac{x_4}{2} + \varepsilon
\end{array}
\qquad
\begin{array}{cccc}
\frac{x_1}{2} + \varepsilon & \frac{x_1}{2} - \varepsilon & 0 & 0 \\
\frac{x_2}{2} - \varepsilon & \frac{x_2}{2} + \varepsilon & 0 & 0 \\
\frac{1-2x_1-2x_2}{4} & \frac{1-2x_1-2x_2}{4} - \varepsilon & \frac{1-2x_4}{4} + 2\varepsilon & \frac{1-2x_4}{4} - \varepsilon \\
0 & \varepsilon & \frac{x_4}{2} - 2\varepsilon & \frac{x_4}{2} + \varepsilon
\end{array}$$

Given any Θ and p such that $p(\Theta_i) = \frac{1}{4}$, for all i, let $S = \{1, 2, ..., 4\}$, $f_\theta(j) = kP_{ji}, f'_\theta(j) = 4P'_{ji}$ for $\theta \in \Theta_i$, $i, j = 1, ..., 4$. It is easily verified that for ε arbitrarily small, (θ, S, f, p) median-rationalizes x if $x_4 > x_1 + x_2$, and (θ, S, f', p) median-rationalizes x if $x_4 \leq x_1 + x_2$. Furthermore, both (θ, S, f, p) and (θ, S, f', p) satisfy msp.

Necessity. Suppose that x can be median-rationalized with a model $\left(\Theta, \widetilde{S}, \widetilde{f}, p\right)$ with monotone signals. We show that equation (4) holds; the argument for (5) is symmetric.

Let $S = \{1, .., k\}$ and $f_\theta(j) = \int_{\widetilde{S}_j} d\widetilde{f}_\theta(s)$. Then (Θ, S, f, p) is a model with monotone signals which also median-rationalizes x. Let P be the $k \times k$ matrix defined by

$$P_{ji} = \int_{\Theta_i} f_\theta(j)\, dp(\theta) = F(j \mid \Theta_i)\, p(\Theta_i).$$

In order to show necessity, we first establish three facts about P and (Θ, S, f, p). As in the proof of necessity of Theorem (1), for $q = \frac{1}{2}$ we have that, (i) for all j, such that $x_j > 0$, $p\left(\cup_{i=j}^{k} \Theta_i \mid j\right) > \frac{1}{2}$. Also, since (Θ, S, f, p) rationalizes x, $x_j = F(S_j) = F(j)$ and therefore, (ii) for all i and all j such that $x_j > 0$, $p(\Theta_i \mid j) = \frac{F(j|\Theta_i)}{F(j)} p(\Theta_i) = \frac{P_{ji}}{x_j}$.

Since (Θ, S, f, p) satisfies msp, $\sum_{n=j}^{k} f_\theta(n) \geq \sum_{n=j}^{k} f_{\theta'}(n)$, for $\theta \in \Theta_{i-1}$ and $\theta' \in \Theta_{i'}$, $i' \leq i - 1$. We have

$$L \equiv \inf_{\theta \in \Theta_{i-1}} \sum\nolimits_{n=j}^{k} f_\theta(n) \geq \sup_{\theta' \in \Theta_{i'}} \sum\nolimits_{n=j}^{k} f_{\theta'}(n) \equiv U$$

so that for all j and all i, i' with $i' \leq i-1$,

$$\begin{aligned}\textstyle\sum_{n=j}^{k} P_{ni'} &= \textstyle\sum_{n=j}^{k} \int_{\Theta_{i'}} f_{\theta'}(n)\, dp(\theta') = \int_{\Theta_{i'}} \textstyle\sum_{n=j}^{k} f_{\theta'}(n)\, dp(\theta') \leq \int_{\Theta_{i'}} U dp(\theta') \\ &\leq \int_{\Theta_{i'}} L dp(\theta') = \int_{\Theta_{i-1}} L dp(\theta) \leq \int_{\Theta_{i-1}} \textstyle\sum_{n=j}^{k} f_{\theta}(n)\, dp(\theta) = \textstyle\sum_{n=j}^{k} P_{n,i-1}.\end{aligned}$$

Therefore, (iii) for all j and all i, i' with $i' \leq i-1$, $\sum_{n=j}^{k} P_{ni'} \leq \sum_{n=j}^{k} P_{n,i-1}$.

Let $\widehat{j}$ be the largest j for which $x_j > 0$.For $i > \widehat{j}$, inequality (4) holds trivially. Let

$$C_i = \textstyle\sum_{g=i}^{k} \sum_{j=i}^{k} P_{jg},\ \overline{P}_j(i) = \sum_{m=i}^{k} P_{jm}, \text{ and } \overline{F}_i(j) = \sum_{m=j}^{k} P_{mi}. \tag{19}$$

Since $C_i \leq \sum_{g=i}^{k} \sum_{j=1}^{k} P_{jg} = \frac{k-i+1}{k}$, it suffices to show that

$$C_i > \textstyle\sum_{j=i}^{k} \frac{x_j}{2} \frac{2j-i-1}{j-1}. \tag{20}$$

for $i \leq \widehat{j}$. We proceed inductively. From (i) and (ii),

$$C_{\widehat{j}} = \textstyle\sum_{i=\widehat{j}}^{k} P_{\widehat{j}i} = \sum_{i=\widehat{j}}^{k} p\left(\Theta_i \mid \widehat{j}\right) x_{\widehat{j}} > \frac{x_{\widehat{j}}}{2} = \sum_{j=\widehat{j}}^{k} \frac{x_j}{2} \frac{2j-\widehat{j}-1}{j-1}$$

which establishes (20) for $i = \widehat{j}$.

Suppose that (20) holds for $i = t \leq \widehat{j}$. If $x_{t-1} = 0$ then $P_{t-1,g} = 0$ for all g, and $\overline{P}_{t-1}(t-1) = 0 = \frac{x_{t-1}}{2}$. If $x_{t-1} > 0$, from (i) and (ii) we have,

$$\frac{1}{2} < p\left(\cup_{m=t-1}^{k} \Theta_m \mid t-1\right) = \textstyle\sum_{m=t-1}^{k} p(\Theta_m \mid t-1) = \sum_{m=t-1}^{k} \frac{P_{t-1,m}}{x_{t-1}} = \frac{\overline{P}_{t-1}(t-1)}{x_{t-1}}.$$

Hence, $\overline{P}_{t-1}(t-1) \geq \frac{x_{t-1}}{2}$. From (iii), for all $i' \leq t-1$, we have $\overline{F}_{t-1}(t) \geq \overline{F}_{i'}(t)$ and therefore $\overline{F}_{t-1}(t) \geq \sum_{i'=1}^{t-1} \frac{\overline{F}_{i'}(t)}{t-1}$. Also, since $\sum_{i=1}^{k} P_{ji} = \sum_{i=1}^{k} p(\Theta_i \mid j)\, x_j = x_j$ we have

$$\begin{aligned} C_{t-1} &= \overline{F}_{t-1}(t) + \overline{P}_{t-1}(t-1) + C_t \geq \frac{\sum_{i'=t}^{k} x_{i'} - C_t}{t-1} + \overline{P}_{t-1}(t-1) + C_t \\ &= \frac{\sum_{i'=t}^{k} x_{i'}}{t-1} + \overline{P}_{t-1}(t-1) + C_t \frac{t-2}{t-1} \geq \frac{\sum_{i'=t}^{k} x_{i'}}{t-1} + \frac{x_{t-1}}{2} + C_t \frac{t-2}{t-1}. \\ &> \frac{\sum_{i'=t}^{k} x_{i'}}{t-1} + \frac{x_{t-1}}{2} + \left(\textstyle\sum_{i'=t}^{k} \frac{x_{i'}}{2} \frac{2i'-t-1}{i'-1}\right) \frac{t-2}{t-1} \\ &= \frac{x_{t-1}}{2} + \textstyle\sum_{i'=t}^{k} \frac{x_{i'}}{2} \frac{2i'-t}{i'-1} = \sum_{i'=t-1}^{k} \frac{x_{i'}}{2} \frac{2i'-t}{i'-1} \end{aligned}$$

so that (20) holds for $t-1$ as well. Hence, (20) holds for $1, ..., \widehat{j}$. ■

References

Alicke, M. D., M. L. Klotz, D. L. Breitenbecher, T. J. Yurak, and D.S. Vredenburg (1995), "Personal contact, individuation, and the better-than-average effect," *Journal of Personality and Social Psychology*, **68(5)**, 804-825.

Barber, B. and T. Odean (2001), "Boys Will Be Boys: Gender, Overconfidence, And Common Stock Investment," *Quarterly Journal of Economics*, 116(1), 261-92.
Bem, D.J. (1967), "Self-perception theory: An alternative interpretation of cognitive dissonance phenomena," *Psychological Review*, **74(3)**, 183-200.
Bénabou, R. and J. Tirole (2002), "Self Confidence and Personal Motivation," *Quarterly Journal of Economics*, **117(3)**, 871-915.
Benoît, J.P., and J. Dubra (2009), "Overconfidence?" at ssrn.com, abstract_id=1088746.
Benoît, J.P., and J. Dubra (2011), "Rationalizing Overconfident Data".
Bernardo, A. and I. Welch (2001), "On the Evolution of Overconfidence and Entrepreneurs," *Journal of Economics & Management Strategy*, **10(3)**, 301-330.
Biais, B., D. Hilton, K. Mazurier and S. Pouget, (2005), "Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market," *Review of Economic Studies*, **72(2)**, 287-312.
Brocas, I. and J. Carrillo (2007), "Systematic errors in decision-making," mimeo.
Burson, K., R. Larrick, and J. Soll (2005) , Social Comparison and Confidence: When Thinking You're Better than average Predicts Overconfidence, Ross School of Business Working Paper No. 1016.
Camerer, C. (1997), "Progress in Behavioral Game Theory," *Journal of Economic Perspectives*, **11(4)**. pp. 167-88.
Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: an experimental approach', *American Economic Review*, **89(1)**, pp. 306–18.
Chuang, W. and B. Lee, (2006), "An empirical evaluation of the overconfidence hypothesis," Journal of Banking & Finance, 30(9), 2489-515.
Clark, J. and L. Friesen (2008), "Rational Expectations of Own Performance: An Experimental Study," forthcoming *Economic Journal*.
Cooper, P. (1990), "Elderly drivers' views of self and driving in relation to the evidence of accident data," *Journal of Safety Research*, **21**, 103-13.
Daniel, K., D. Hirshleifer and A. Subrahmanyam (2001), "Overconfidence, Arbitrage, and Equilibrium Asset Pricing," *Journal of Finance*, **56(3)**, 921-65.
De Bondt, W. and R. H. Thaler, (1995), "Financial decision-making in markets and firms: a behavioral perspective', in (R. A. Jarrow, V. Maksimovic and W. T. Ziemba, eds), *Finance, Handbooks in Operations Research and Management Science*, vol. 9, pp. 385–410. Amsterdam: North Holland.
Dominitz, J., (1998) "Earnings expectations, revisions, and realizations," *Review of Economics and Statistics*, **80(3)**, 374-88.
Dunning, D., J. A. Meyerowitz, and A. D. Holzberg (1989) "Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability", *Journal of Personality and Social Psychology,* **57(6)**, 1082-1090.
Fang, H. and G. Moscarini, (2005) "Morale Hazard," *Journal of Monetary Economics*, **52(4)**, 749-777.
Festinger, L. (1954) "A Theory of Social Comparison Processes," *Human Relations*, **7(2)**, 117-140.
Garcia, D., F. Sangiorgi and B. Urosevic, (2007), "Overconfidence and Market Efficiency with Heterogeneous Agents," *Journal Economic Theory*, **30(2)**, 313-36.

Greenwald, A. (1980), "The Totalitarian Ego, Fabrication and Revision of Personal History," *American Psychologist*, **35(7)**, 603-13.

Grieco, D. and R. Hogarth (2009), "Overconfidence in absolute and relative performance: The regression hypothesis and Bayesian updating," *Journal of Economic Psychology*, **30**, 756–71.

Groeger, J. and G.E. Grande (1996) "Self-preserving assessments of skill?" *British Journal of Psychology* **87**, 61-79.

Hoelzl, E. and A. Rustichini, (2005), "Overconfident: do you put your money on it?" the *Economic Journal*, **115**, pp. 305-18.

Holland, C., (1993) "Self-bias in older drivers' judgments of accident likelihood", *Accident Analysis and Prevention*, **(25)(4)**, pp. 431-441.

Karni, E., (2009), "A Mechanism for Eliciting Probabilities," *Econometrica*, **77(2)**, 603–6.

Kőszegi, B., (2006), "Ego Utility, Overconfidence, and Task Choice," *Journal of the European Economic Association*, **4(4)**, 673-707.

Kruger, J. (1999), "Lake Wobegon Be Gone! The "Below-Average Effect" and the Egocentric Nature of Comparative Ability Judgements", *Journal of Personality and Social Psychology*, **77(2)**, 221-232.

Kruger, J., and D. Dunning (1999), "Unskilled and Unaware of it: How difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments," *Journal of Personality and Social Psychology*, **77(6)**, 1121-1134.

Kruger, J. P. Windschitl, J. Burrus, F. Fessel, J.R. Chambers (2008), "The rational side of egocentrism in social comparisons," *Journal of Experimental Social Psychology,* **44**, 220–32.

Kyle, A. and F.A. Wang, (1997), "Speculation Duopoly with Agreement to Disagree: Can Overconfidence Survive the Market Test?" *Journal of Finance*, **52(5)**, 2073-90.

Larrick, R. P., K. A. Burson and J.B. Soll, (2007), "Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not)," *Organizational Behavior and Human Decision Processes*, **102**, 76–94.

Malmendier, U. and G. Tate (2005), "CEO Overconfidence and Corporate Investment," *Journal of Finance*, **60(6)**, 2661-700.

Marottoli, R. and E. Richardson (1998), "Confidence in, and self-rating of, driving ability among older drivers," *Accident Analysis and Prevention*, **(30)(3)**, pp. 331-336.

Massie, D. and K. Campbell (1993), "Analysis of Accident Rates by Age, Gender, and Time of the Day, Based on the 1990 Nationwide Personal Transportation Survey," The University of Michigan Transportation Research Institute, report UMTRI-93-7.

Mathews, M., and A. Moran (1986), "Age differences in male driver's perception of accident risk: the role of perceived driving ability", *Accident Analysis and Prevention*, **(18)(4)**, pp.299-313.

Milgrom, P.R. (1981), "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, **12(2)**, pp. 380-91.

Moore, D. (2007), "Not so above average after all: When people believe they are worse than average and its impications for theories of bias in social comparison," Organizational Behavior and Human Decision Processes, **102(1)**, pp 42-58.

Moore, D. A., and P.J. Healy (2008), "The trouble with overconfidence," *Psychological Review*, **115(2)**, 502-517.

Myers (1999), Social Psychology, 6th edition

Noth, M. and M. Weber, (2003), "Information Aggregation with Random Ordering: Cascades and Overconfidence," *Economic Journal*, 113(484), 166-89.

Peng, L. and W. Xiong, (2006), "Investor attention, overconfidence and category learning," *Journal of Financial Economics*, **80(3)**, 563-602.

Sandroni, A. and F. Squintani, (2008), "Overconfidence, Insurance and Paternalism" forthcoming, *American Economic Review*.

Svenson, O., (1981), "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica*, **94**, pp 143-148.

Van den Steen, E. (2004), "Rational overoptimism," *American Economic Review,* **94(4)**, 1141–1151.

Walton, D., (1999), "Examining the self-enhancement bias: professional truck drivers' perceptions of speed, safety, skill and consideration," *Transportation Research Part F,* 91-113.

Wang, A. (2001), "Overconfidence, Investor Sentiment, and Evolution," *Journal of Financial Intermediation*, **10(2)**, 138-70.

Weinstein, N. (1980), "Unrealistic Optimism about Future Life Events," *Journal of Personality and Social Psychology*, **39(5)**, 806-20.

Whitt, W. (1980), "Uniform Conditional Stochastic Order," *Journal of Applied Probability*, **17(1)**, pp. 112-123.

Zábojník, J. (2004), "A Model of Rational Bias in Self-Assessments," *Economic Theory,* **23(2)**, 259–82.