The Hyperplane Model of Survey Response

Thomas Demuynck
ECARES, Université Libre de Bruxelles

Leon Hennen
ECARES, Université Libre de Bruxelles

January 2026

ECARES working paper 2026-04

ECARES
ULB - CP 114/04
50, F.D. Roosevelt Ave., B-1050 Brussels BELGIUM
www.ecares.org

The Hyperplane Model of Survey Response

Thomas Demuynck León Hennen

January 20, 2026

Abstract

We present a model of survey response based on two consistency properties. We derive its empirical testable restrictions and illustrate our findings using data from the world values survey.

Keywords: Survey, Hyperplane Arrangements
JEL-codes:C14, C15, C44, C61

1 Introduction

Insitutions, governments and scientists routinely collect and analyse survey data to measure the prevalence of opinions, attitudes or beliefs within a population.[1] While such surveys offer a rich source of information, their use in economics has traditionally been limited.[2] Recently, however, interest in survey data has grown, as it offers access to information that is difficult to obtain through standard revealed-preference approaches (see, among others, Almas, Attanasio, and Pamela (2024); Stantcheva (2023); Manski (2004); Benjamin, Guzman, Fleurbaey, Heffetz, and Kimball (2003); Liu and Netzer (2023)). Despite this renewed attention, there remains little work aimed at developing formal decision-theoretic frameworks that try to explain how respondents answer to such surveys.

In this paper, we propose such a model based on two intuitive consistency properties. From these, we obtain what we call the *Hyperplane model of survey response*. We characterize this model's testable implications in the two-dimensional case and provide an illustration using data from the World Values Survey.

The hyperplane model of survey response We consider a setting where respondents are asked whether they agree or disagree with a finite set of statements. We will call these statements *survey questions*. Each survey question q is assumed to possess multiple latent attributes, and can therefore be represented by a non-zero vector in n-dimensional Euclidean space, $q = (a_1, \dots, a_n) \in \mathbb{R}^n \setminus \{0\}$. The notion of representing survey questions as vectors of attributes is not uncommon.[3] Often, it is assumed that there is only a single attribute, allowing the representation by a scalar value, i.e. $q \in \mathbb{R}$. Agreement or disagreement is then determined by the question's position on the real line (Ballester and Apesteguia, 2024).

[1]See for instance the European Eurobarometer or the American community survey.

[2]This hesitancy largely comes from concerns about various biases (Bertrand and Sendhil, 2001), as well as doubts regarding respondents' incentives to answer truthfully, given the non-incentivized nature of most surveys. However, see Prelec (2004) who employs Bayesian methods to construct a mechanism to elicit honest responses in surveys.

[3]See, for example, the literature on Item Response Theory in statistical psychology, which explores the relationship between latent traits (i.e. characteristics of respondents) and their observable manifestations (e.g. responses to questions or outcomes) (van der Linden, 1997).

We propose a model of survey response that allows for multiple attributes.[4] As an example, consider a survey on the topic of drug use. Questions in this survey encompass multiple attributes. One attribute might pertain to notions crime and safety (e.g. drug trafficking, violence, corruption, theft, etc.) while another attribute might relate to health and well-being (e.g. illness, neurological and psychological effects, needle usage, medical pain-relief, etc.).

We assume that a respondent's agreement or disagreement with a specific survey question is based on the values of the attributes of the vector q representing the survey question. We impose two consistency properties on these responses. The first property requires that if a respondent agrees with some survey question q, she should disagree with the inverse survey question $-q$, which is the question where the value of each attribute is negated. This condition reflects the principle that agreement with a question implies disagreement with its negation. Our second consistency property assumes that if a respondent agrees with two distinct questions q and q', she should also agree with any convex combination of the two, $q'' = \alpha q + (1-\alpha)q', \alpha \in [0,1]$. This captures the intuition that if a respondent agrees with two (more extreme) positions, she should also agree with a compromise between the two.

Using a simple separating hyperplane argument, we derive the following representation result: if the two consistency properties are satisfied, then each respondent can be associated with a non-zero *ideal point* $\theta \in \mathbb{R}^n \setminus \{0\}$ such that the respondent will agree with any question q for which the inner product $\langle q, \theta \rangle$ is strictly positive, and disagree when this inner product is strictly negative. In other words, by identifying respondents with their ideal points θ, the set of all respondents who agree with a given question q lies entirely on one side of the hyperplane through the origin whose slope is defined by q, while those who disagree lie on the opposite side. Given this, we call this the *Hyperplane model of survey response*.

In reality, we do not observe the questions' attributes nor the respondents' ideal points. Instead, we only observe a finite set of responses from a finite number of respondents to a finite number of questions. In other words, both q and θ are latent variables from the perspective of the econometrician. This leads to the following rationalizability question: given a dataset of observed survey responses, under what conditions can it be said to be consistent with answers obtained from the Hyperplane model, for some values of q or θ. Our paper addresses this question when there are two attributes.

Literature overview The paper most closely related to ours is the recent work of Ballester and Apesteguia (2024). They study a setting where questions are one-dimensional–that is, points on the real line–and opinions are modelled using a quasi-concave (single-peaked) utility function $u : \mathbb{R} \to \mathbb{R}$. In their framework, each individual has a threshold α and agrees with all questions for which the utility exceeds this threshold, and disagrees otherwise.

This setup is equivalent to assuming that each respondent has an ideal point $\theta \in \mathbb{R}$ and a *tolerance level* τ, such that they agree with a question if and only if $|\theta - q| \leq \tau$. A direct implication of this structure is that the set of statements that a respondent agrees with should appear consecutively, i.e. for any three questions $q < q' < q''$ if a respondent agrees with q and q'', then she should also agree with q'. For this reason, we refer to this framework as the Consecutive Ones (CO) model. As we will show in section 5, this CO model is empirically distinguishable from the Hyperplane model. We also present an alternative Circular Consecutive Ones (CCO) model and show that this CCO models (empirically) encompasses both the Hyperplane and the CO model.

Cavaillé, Chen, and van Der Straeten (2021) develop a model of survey response in which

[4] This bears some resemblance to spatial voting models, where political candidates are represented as points in multidimensional space, which each dimension corresponding to a distinct attribute (Enelow and Hinich, 1984).

individuals may give a biased opinion on Likert-type surveys due to auxiliary motives such as strategic reporting (e.g. to influence policy outcomes) or signalling (i.e. to express identity or conform to social norms). To address these biases, they propose a Quadratic Voting for Survey Research design aimed at eliciting more truthful responses. Next, Falk, Neuber, and Strack (2021) introduce a choice-theoretic model of survey response under the assumption that respondents possess imperfect self-knowledge regarding their own characteristics. Building on this framework, they develop a consistent and unbiased estimator for measuring self-knowledge. In contrast to these works, we start from a more normative benchmark for internal consistent survey behaviour while the models of Cavaillé, Chen, and van Der Straeten (2021) and Falk, Neuber, and Strack (2021) try to explain real-world deviations due to strategic motives or imperfect self-knowledge.

Outline Section 2 lays out the hyperplane model and the empirical setting. Section 3 defines our rationalizability notion. Section 4 provides a characterization of the survey datasets that are rationalizable by the hyperplane model when there are 2 attributes. Section 5 discusses the similarities and differences between our model and the CO and CCO models. Section 6 provides an illustration using data from the World Values Survey.

2 The Hyperplane Model of Survey Response

Consider a setting in which a respondent is presented with a number of survey questions. Each survey question asks whether the respondent agrees or disagrees with a given statement. We denote agreement by Y(es) and disagreement by N(o)

In practice, many surveys do not simply ask for agreement or disagreement with some statements, but rather use Likert-type scales to capture respondents' levels of agreement (Likert, 1932).[5] Likert scales allows respondents to express varying intensities of agreement or disagreement. While it is generally accepted that these responses convey ordinal information (e.g. "strongly agree" implies more agreement than "weakly agree"), there is less consensus on whether these intensities are comparable across questions–or between respondents.[6] For instance, what does it mean if one respondent "strongly agrees" with some statement while another respondent only "weakly agrees" with the same (or some other) statement? Given these interpretive challenges, we adopt a more conservative approach by collapsing the Likert-scale responses into two exhaustive categories ("Agree" (Y) and "Disagree" (N)) and assume these coarser categories do allow for meaningful comparison across questions and respondents.

We model a survey question as a non-zero vector $q = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n \setminus \{0\}$. Intuitively, this vector captures the values of n latent attributes a_k ($k \leq n$) that define the content or framing of the question. For any such question, we also define its negation $-q = (-a_1, -a_2, \ldots, -a_n)$. Conceptually, $-q$ represents the reverse or opposite of the original question q. For example if q asks the respondent whether she agrees with a particular statement, then $-q$ corresponds to asking whether the respondent disagrees with that same statement.

A respondent's answers can be modelled by a response function:

$$\psi : \mathbb{R}^n \setminus \{0\} \to \{\text{Y}, \text{N}\},$$

[5] Two other common scaling methods include the Thurstone scale (Thurstone, 1928), which records binary agreement or disagreement with individual statements, and the Guttman scale (Guttman, 1944), which presents a hierarchically ordered list of statements, where the level of agreement is inferred from the point at which the respondent switches from agreement to disagreement (or vice versa).

[6] See, among others, (Aldrich and McKelvey, 1977; Rossi, Gilula, and Allenby, 2001; Javaras and Ripley, 2007)

which assigns to each question $q = (a_1, \ldots, a_n)$ a response that can be either Yes (Agree) or No (Disagree).

We impose two consistency properties on the response function ψ. The first ensures logically coherence with respect to question negation:

Property 1. *For all* $q \in \mathbb{R}^n \setminus \{0\}$, $\psi(q) = \text{Y}$ *if and only if* $\psi(-q) = \text{N}$.

This condition captures the idea that agreement with a question implies disagreement with its exact opposite.

The second property imposes a convexity condition. If a respondent agrees with two distinct questions, she should also agree with any convex combination of those questions, which can be interpreted as a more intermediate version of the two originals.

Property 2. *For all* $q, q' \in \mathbb{R}^n \setminus \{0\}$ *if* $\psi(q) = \psi(q') = \text{Y}$, $\alpha \in [0, 1]$ *and* $\alpha q + (1 - \alpha)q' \neq 0$, *then* $\psi(\alpha q + (1 - \alpha)q') = \text{Y}$.

Note that both properties are invariant under coordinate-wise scaling of the latent attributes. However, Property 1 is not invariant under translations, as negation depends on the origin of the attribute space. We now provide a characterization of the response function ψ under the two consistency assumptions. All proofs are in the Appendix.

Theorem 1. *Let* $\psi : \mathbb{R}^n \setminus \{0\} \to \{\text{Y}, \text{N}\}$ *satisfy properties 1 and 2. Then, we can find a point* $\theta \in \mathbb{R}^n \setminus \{0\}$ *such that:*[7]

- *If* $\langle \theta, q \rangle > 0$ *then* $\psi(q) = \text{Y}$,
- *If* $\langle \theta, q \rangle < 0$ *then* $\psi(q) = \text{N}$.

The proof of Theorem 1 relies on a simply separating hyperplane argument. Intuitively, we can interpret θ as the respondent's *ideal point* and let q define the orientation of a hyperplane passing through the origin. Figure 1 illustrates this in the two dimensional case. The hyperplane associated with q divides the space $\mathbb{R}^n$ in two open half spaces: respondents whose ideal point θ lies on one side of the hyperplane will respond Y(es) to q, while those on the opposite side will respond N(o). We will indicate the half space corresponding to the Y responses by an arrow.

In this manner, each question acts as a separator, dividing respondents into those who agree and those who disagree.[8] From this point forwards, we will graphically represent questions by their corresponding hyperplanes and omit the axis that represent the latent attributes, as they are not essential to the analysis. See Figure 2 for an illustration with five questions.

Note that, in a dual formulation we could also have represented each respondent by a hyperplane with normal vector θ and each question q as a point on the plane. In this setting, a respondent will answer Y to all questions lying on one side of their hyperplane and N to those on the opposite side.

3 Rationalizability

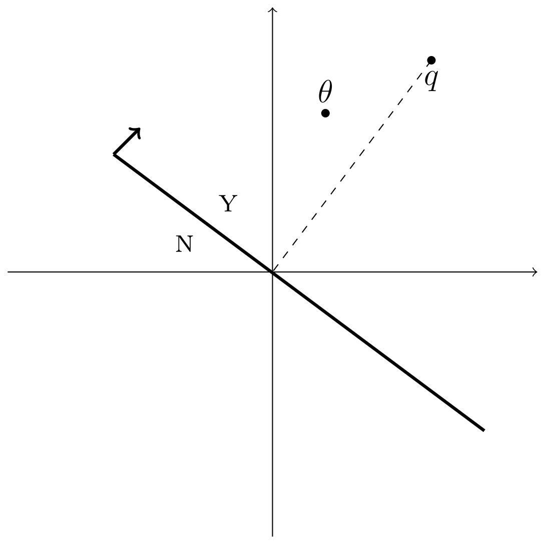
Consider an empirical setting where a finite number of respondents are asked to indicate agreement or disagreement with a finite set of J survey questions. A respondent's answers are represented by a response pattern, defined as a function:

$$r : \{1, \ldots, J\} \to \{\text{Y}, \text{N}\},$$

[7]For two vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ we use $\langle x, y \rangle$ to denote their scalar product $\sum_{k \le n} x_i y_i$.

[8]Note that the Theorem is not a full characterization as it says nothing about the response for a respondent where $\langle \theta, q \rangle = 0$.

Figure 1: A question, its hyperplane and an ideal point



where $r(j) \in \{\text{Y}, \text{N}\}$ denotes the response to question j. For convenience we often write a response pattern as a sequence of answers. For example, the response pattern $r = $ YNN indicates an agreement with question 1 and disagreement with questions 2 and 3. We define the negation of a response pattern's response to question j as:

$$\neg r(j) = \begin{cases} \text{N if } r(j) = \text{Y} \\ \text{Y if } r(j) = \text{N}. \end{cases} .$$

Extending this pointwise, the *negation* of a response pattern r, denoted $\neg r$, is the pattern obtained by negating the answer to each question $j \leq J$ in r. A *dataset* is defined as a finite collection of response patterns:

$$\mathcal{D} = (r^t)_{t \leq T}$$

where each $t \in \{1, \dots, T\}$ indexes a respondent and r^t is the response pattern of respondent t. We say that a dataset is rationalizable if every response pattern in $\mathcal{D}$ is consistent with the predictions of the Hyperplane model as stated by Theorem 1.

Definition 1. *A dataset $\mathcal{D} = (r^t)_{t \leq T}$ is rationalizable by the Hyperplane model if and only if for all respondents $t \in \{1, \dots, T\}$ there exists an ideal point $\theta^t \in \mathbb{R}^n \setminus \{0\}$ and for all questions $j \in \{1, \dots, J\}$ there is a vector $q^j \in \mathbb{R}^n \setminus \{0\}$ (non-proportional to each other) such that,*

$$r^t(j) = \text{Y} \rightarrow \langle \theta^t, q^j \rangle \geq 0,$$
$$r^t(j) = \text{N} \rightarrow \langle \theta^t, q^j \rangle \leq 0.$$

Note that Definition 1 requires that no two questions are proportional to each other. This means that different questions are represented by different hyperplane. This assumption is necessary for our rationalizability condition to be testable. Indeed, if we allow all questions to be represented by the same hyperplane, we could pick all vectors θ^t to lie on this common hyperplanes, implying that all datasets can be rationalized.

Remark that the rationalizability of a dataset critically depends on the number of latent attributes n. In particular, if we allow $n \geq J$ where J is the number of questions in the survey, then any dataset is rationalizable by the Hyperplane model. To see this, consider a construction where each question q^j is represented by the standard basis vector in $\mathbb{R}^n$ with a 1 in the jth position and zero elsewhere. Then, for each respondent t, we can define the ideal point

$\theta^t \in \mathbb{R}^n \setminus \{0\}$ by setting the jth coordinate to $+1$ if $r^t(j) = \text{Y}$ and to -1 if $r^t(j) = \text{N}$. If, so then if t answers Y to question j, we have that:

$$\langle \theta^t, q^j \rangle = 1 > 0$$

while if t answers N to question j,

$$\langle \theta^t, q^j \rangle = -1 < 0.$$

This construction therefore ensures that each response pattern is perfectly aligned with the Hyperplane model, thereby rationalizing the entire dataset.

On the other hand, if $n = 1$, i.e. there is only one attribute, then every hyperplane coincides with the point $\{0\}$, meaning that there are only two possible response profiles, where one is the negation of the other. The smallest non-trivial case, $n = 2$ is the one we will focus on in the remainder of this paper.

Hyperplane arrangements The general rationalizability problem (for arbitrary n) is closely related to the mathematical theory of *hyperplane arrangements*, which studies finite collections of hyperplanes in a vector space. When all such hyperplanes pass through the origin, such arrangement is referred to as a *central hyperplane arrangement*. From this perspective, our rationalizability problem aligns naturally with the study of central hyperplane arrangements, as each survey question corresponds to a hyperplane through the origin, and (rational) response patterns correspond to the regions defined by these arrangements.

To see this, note that every (central) hyperplane arrangement (in $\mathbb{R}^n$) partitions the space in a finite number of regions. Each region lies entirely within one of the two open half-spaces defined by each hyperplane. By labelling one side of the hyperplane as Y and the other as N, each region is therefore uniquely identified by a response profile. In the literature, these are called sign vectors. A central topic in the theory of hyperplane arrangements is to derive properties on the collection of sign-vectors that are derived from central hyperplane arrangements. This question is directly related to our notion of rationalizability. Indeed, a dataset $\mathcal{D}$ is a collection of response profiles, and it is rationalizable by the Hyperplane model if and only if this collection is contained within the set of all of sign vectors generated by some central hyperplane arrangement.

The set of sign vectors generated by a central hyperplane arrangement forms a combinatorial structure known as an *oriented matroid*. As such, analyzing the properties of oriented matroids should tells us something about the properties of hyperplane arrangements (Festa, 2024). However, not all oriented matroids are realizable in the sense that they can be represented by actual hyperplane arangements and deciding whether a particular oriented matroid is representable by such central hyperplane arrangement is a difficult problem.[9] This is relevant to our setting: verifying that the response profiles of a given dataset are part of an oriented matroid is therefore not sufficient for rationalizability, as the latter adds the additional requirement that this oriented matroid should also be realizable.

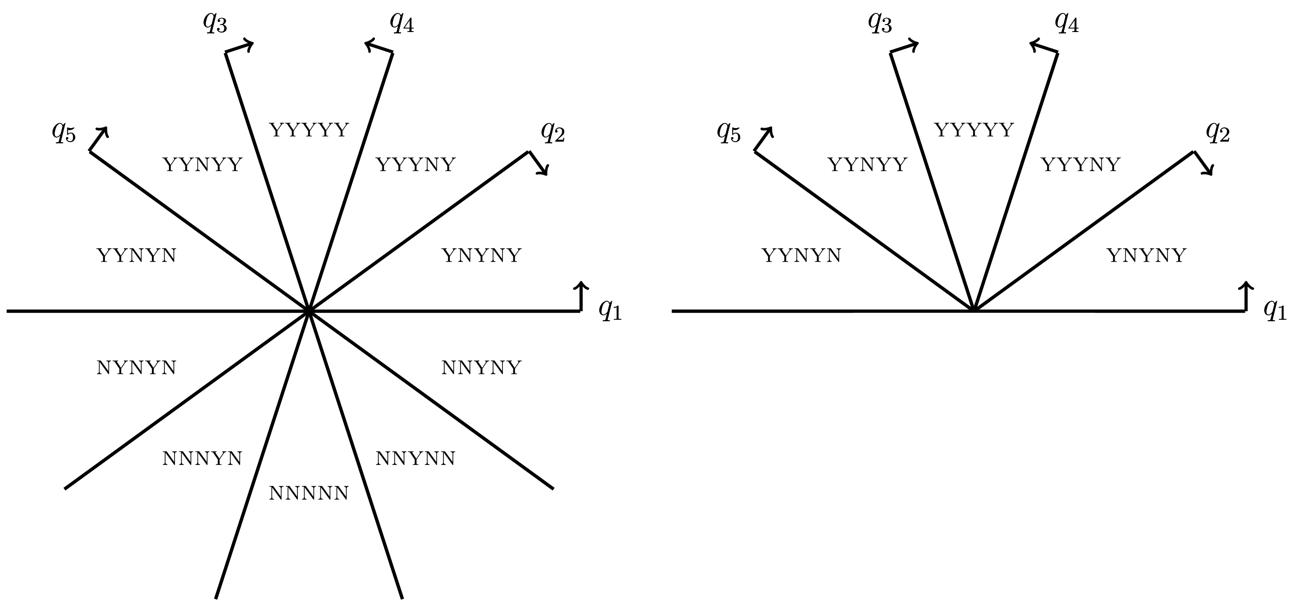
Given this complexity, we choose to restrict our attention in the remainder of the paper to the simplest non-trivial setting which is the case of two latent dimensions, $n = 2$.

The case $n = 2$ To analyze the setting with two attributes, i.e. $n = 2$, we start with an example involving 5 survey questions ($J = 5$). The left part of Figure 2 represents a possible hyperplane arrangement. One easily checks that the five hyperplanes give rise to 10 distinct regions and that each region corresponds to a unique response pattern. Each region contains the ideal points of all respondents who respond according to this particular response profile.

[9] In particular, it is NP-hard (Shor, 1991).

Moving from one region to a neighbouring region changes the response to the question whose hyperplane separates the two.

Figure 2: Example with 5 questions



The example allows us to draw several insights. First, if there are J questions, then the number of distinct response patterns–i.e. regions formed by the hyperplane arrangement–that can potentially be observed is at most $2J$ as each hyperplane adds an additional 2 regions. Of course, some regions may contain no respondents, so the actual number of observed patterns may be strictly less than $2J$. Importantly, however, for $J > 2$, the number of response patterns consistent with the hyperplane model is significantly smaller than the total number of theoretically possible patterns, which is 2^J. This reflects the model's restrictive structure.[10]

Second, it is important to note that not every possible collection of $2J$ response patterns is consistent with the Hyperplane model. A key constraint arises from the symmetry implied by Assumption 1: if some response pattern is possible, then its negation, i.e. the pattern obtained by negating all answers, must also be possible. As an example, consider a survey with $J = 3$ questions and a resulting dataset,

$$\mathcal{D} = \{\text{YYY}, \text{YNY}, \text{NNY}, \text{NYN}, \text{NYY}, \text{YNN}\},$$

which contains six distinct response patterns. This is the maximum number allowed by the Hyperplane model when there are three questions. However, it is not rationalizable, because the negation of YYY, namely NNN is absent from the dataset. This violates the requirement that possible response patterns must appear in complementary pairs.

Finally, it is important to note that the set of possible response patterns generated by the Hyperplane model is invariant under certain geometric transformations. Specifically, rotating the plane, reflecting the hyperplanes across an axis through the origin, or adjusting the spacing between the hyperplanes–so long their relative orientation remains unchanged–does not affect the set of possible response patterns.

Figure 2 also illustrates a method for generating all response patterns consistent with a given instance of the Hyperplane model. We begin by selecting a specific question, say question 1,

[10]For $J = 2$ one can easily see that any possible response pattern is consistent with the model as all four patterns YY, YN, NY, NN might occur.

and identify the half-space where respondents answer Y to that question. The right panel of Figure 2 shows this region. Due to the symmetry, the response patterns in the opposite half-space–that respond N to question 1–can be obtained by negating the patterns in the selected half-space.

We can now describe a systematic way to generate all response patterns within this half-space. Starting from, say, the rightmost response pattern r = YNYNY, we can construct all other patterns in the same half-space by sequentially negating the responses to selected questions one at a time. The order in which questions are negated corresponds to the order by which the hyperplanes are visited when traversing from the right to the left: first question 2, then question 4, then 3 and finally question 5. We denote this order by the list $\ell = 2, 4, 3, 5$.

This process yields the following sequence of response patterns and their negations.

$$\begin{array}{lccccccccc} \text{profile} & \text{YNYNY} & \xrightarrow{2} & \text{YYYNY} & \xrightarrow{4} & \text{YYYYY} & \xrightarrow{3} & \text{YYNYY} & \xrightarrow{5} & \text{YYNYN} \\ & \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\ \text{negation} & \text{NYNYN} & & \text{NNNYN} & & \text{NNNNN} & & \text{NNYNN} & & \text{NNYNY} \end{array}$$

At each step in the construction, a new response pattern is generated by negating exactly one answer corresponding to the next question in the specified list. This illustrates that, to construct a collection of response patterns consistent with the model, we need two components.

1. A starting response pattern r with $r(1)$ = Y.
2. An list $\ell = (q_1, \ldots, q_{J-1})$ of the remaining questions $\{2, \ldots, J\}$

Different starting response patterns r and different lists ℓ of $\{2, \ldots, J\}$ will correspond to different instances of the Hyperplane model. We call such combination (r, ℓ) a *response model.*

The general procedure for generating the J response patterns in the half-space defined by the response model (r, ℓ) is given by Algorithm 1. The full set of possible response patterns consistent with the hyperplane model is then given by all response profiles in the list $\mathcal{R}(r, \ell)$ generated by this algorithm and their negations. Definition 2 formalizes this.

Definition 2. *Given a set of J questions $\{1, \ldots, J\}$. A a response model (r, ℓ) consists of a response profile $r : \{1, \ldots, J\} \to \{\text{Y}, \text{N}\}$ with $r(1) = \text{Y}$, and an list $\ell = (q_1, \ldots, q_{J-1})$ of the $J - 1$ questions in $\{2, \ldots, J\}$. The collection of possible response profiles that can be generated by the response model (r, ℓ) is given by the list of profiles $\mathcal{R}(r, \ell)$, generated by Algorithm 1 and their negations.*

Algorithm 1 Algorithm to generate all response profiles for the response model (r, ℓ)

Require: a profile r with $r(1) = \text{Y}$, and an enumeration $\ell = (q_1, q_2, \ldots, q_{J-1})$ of the $J - 1$ questions $\{2, \ldots, J\}$.
  set $r_1 = r$.
  for $k = 2$ to J **do**
    Define $r_k(j) = \begin{cases} r_{k-1}(j) \text{ if } j \neq q_{k-1} \\ \neg r_{k-1}(j) \text{ if } j = q_{k-1} \end{cases}$.
  end for
  return $\mathcal{R}(r, \ell) = (r_1, \ldots, r_J)$

Given the definition of a response model, we can rephrase our rationalizability condition in the following way.

Definition 3. *We say that a dataset $\mathcal{D} = (r^t)_{t \leq T}$ is rationalizable by the Hyperplane model if and only if there exists a response model (r, ℓ) such that for all $r^t \in \mathcal{D}$, r^t or $\neg r^t$ is in the list $\mathcal{R}(r, \ell)$ generated by Algorithm 1.*

One way to verify whether a dataset is rationalizable by the Hyperplane model would be to directly verify Definition 3. This could be done by generating for all possible response models (r, ℓ), the list $\mathcal{R}(r, \ell)$ and to see if for one of them r^t or its negation occurs in $\mathcal{R}(r, \ell)$ for all $r^t \in \mathcal{D}$. This, however, is a quite inefficient procedure. For J questions, there are in total 2^J profiles r, half of which satisfy $r(1) = \text{Y}$. Additionally, there are $(J-1)!$ rankings of the questions $\{2, \ldots, J\}$ giving, in total $2^{J-1}(J-1)!$ possible response models (r, ℓ). The next section gives an easier way to verify whether a survey dataset is rationalizable.

4 Characterization

Given a dataset $\mathcal{D}$ and any profile $r \in \mathcal{D}$, define the transformed profile $\overline{r}$ as follows:

$$\overline{r} = \begin{cases} r & \text{if } r(1) = \text{Y} \\ \neg r & \text{if } r(1) = \text{N} \end{cases}$$

By construction: $\overline{r}(1) = \text{Y}$ for all response profiles r. For a dataset $\mathcal{D} = (r^t)_{t \leq T}$, define the transformed dataset $\overline{\mathcal{D}}$ as the set of all such transformed profiles:

$$\overline{\mathcal{D}} = \{\overline{r} | \exists t \leq T, r^t = r\}.$$

The modified dataset $\overline{\mathcal{D}}$ can easily be constructed given $\mathcal{D}$. Using this notation, we can now define a dataset $\mathcal{D}$ to be rationalizable by the Hyperplane model if there exists a response model (r, ℓ) such that any $\overline{r} \in \overline{\mathcal{D}}$ belongs to the set $\mathcal{R}(r, \ell)$.

For any two profiles $\overline{r}_0, \overline{r}_1 \in \overline{\mathcal{D}}$ we denote by $\Delta(\overline{r}_0, \overline{r}_1)$ the set of questions on which their responses differ:

$$\Delta(\overline{r}_0, \overline{r}_1) = \{j \in \{2, \ldots, J\} | \overline{r}_0(j) \neq \overline{r}_1(j)\}$$

For example, if $\overline{r}_0 = \text{YNYYN}$ and $\overline{r}_1 = \text{YYYNY}$ then $\Delta(\overline{r}_0, \overline{r}_1) = \{2, 4, 5\}$ as the two response profiles' answers differ on questions 2, 4 and 5. We call $\Delta(\overline{r}_0, \overline{r}_1)$ the *difference* between $\overline{r}_0$ and $\overline{r}_1$.

To motivate our first condition, suppose that $\mathcal{D}$ is rationalizable by the Hyperplane model. That is, there exists a response model (r, ℓ) such that every profile $\overline{r} \in \overline{\mathcal{D}}$ is somewhere in the list $\mathcal{R}(r, \ell) = (r_1, \ldots, r_J)$. Take 3 profiles $\overline{r}_0, \overline{r}_1, \overline{r}_2 \in \overline{\mathcal{D}}$. By construction of $\mathcal{R}(r, \ell)$ there must be indices $i, j, k \in \{1, \ldots, J\}$ such that $\overline{r}_0 = r_i, \overline{r}_1 = r_j$ and $\overline{r}_2 = r_k$, where r_i, r_j, r_k are the ith, jth and kth profile of the list $\mathcal{R}(r, \ell)$. Without loss of generality, let $i < j < k$. Let the ranking ℓ be given by the sequence of questions $(q_1, \ldots, q_{J-1})$. Given that $i < j < k$, we can characterize the difference between the profiles:

$$\begin{aligned} \Delta(\overline{r}_0, \overline{r}_1) &= \Delta(r_i, r_j) = \{q_i, \ldots, q_{j-1}\}, \\ \Delta(\overline{r}_1, \overline{r}_2) &= \Delta(r_j, r_k) = \{q_j, \ldots, q_{k-1}\}, \\ \Delta(\overline{r}_0, \overline{r}_2) &= \Delta(r_i, r_k) = \{q_i, \ldots, q_{k-1}\}. \end{aligned}$$

Observe that $\Delta(\overline{r}_0, \overline{r}_1) \cap \Delta(\overline{r}_1, \overline{r}_2) = \emptyset$. Consequentially, we have $\Delta(\overline{r}_0, \overline{r}_2) = \Delta(\overline{r}_0, \overline{r}_1) \cup \Delta(\overline{r}_1, \overline{r}_2)$. This implies the following structural property: among the three differences, $\Delta(\overline{r}_0, \overline{r}_1), \Delta(\overline{r}_1, \overline{r}_2)$ and $\Delta(\overline{r}_0, \overline{r}_2)$, two are disjoint, and the third is their union. Geometrically, referring back to Figure 2, this means that the region corresponding to the response profile $\overline{r}_1$ should be between the regions specified by $\overline{r}_0$ and $\overline{r}_2$. This motivates to our first necessary condition for rationalizability.

Condition 1. *For any three profiles $\overline{r}_0, \overline{r}_1, \overline{r}_2 \in \overline{\mathcal{D}}$, we have that among the three differences $\Delta(\overline{r}_0, \overline{r}_1)$, $\Delta(\overline{r}_1, \overline{r}_2)$, and $\Delta(\overline{r}_0, \overline{r}_2)$, two are disjoint (i.e. their intersection is empty), and the third is their union.*

While the previous argument shows that Condition 1 is necessary for rationalizability, it is not sufficient. To show this, consider the following example:

Example 1. *Let:*

$$\overline{\mathcal{D}} = \{\text{YYYY}, \text{YYYN}, \text{YNNN}, \text{YNNY}\}.$$

We compute the differences:

$$\Delta(\text{YYYY}, \text{YYYN}) = \{4\} \qquad \Delta(\text{YYYY}, \text{YNNN}) = \{2, 3, 4\},$$
$$\Delta(\text{YYYY}, \text{YNNY}) = \{2, 3\} \qquad \Delta(\text{YYYN}, \text{YNNN}) = \{2, 3\},$$
$$\Delta(\text{YYYN}, \text{YNNNY}) = \{2, 3, 4\} \qquad \Delta(\text{YNNN}, \text{YNNY}) = \{4\}.$$

A straightforward check confirms that all triples of profiles in $\overline{\mathcal{D}}$ satisfy Assumption 1.

However, this dataset is not rationalizable by the Hyperplane model. To see why, suppose for contradiction that it is. Then there exists a response model (r, ℓ) such that every profile in the dataset $\overline{\mathcal{D}}$ appears in the sequence $\mathcal{R}(r, \ell) = (r_1, \ldots, r_4)$. Since $\overline{\mathcal{D}}$ contains 4 elements, it must be that:

$$\{r_1, r_2, r_3, r_4\} = \{\text{YYYY}, \text{YYYN}, \text{YNNN}, \text{YNNY}\}.$$

If $r_1 = \text{YYYY}$ then as r_1 and r_2 differ on only one question, it must be that $r_2 = \text{YYYN}$. Now, r_3 must differ from r_2 also on exactly one question. However, neither YNNN *nor* YNNY *satisfies this condition relative to* YYYN *as both differ on more than one question. This gives a contradiction.*

If $r_1 = \text{YYYN}$ then a similar reasoning as before requires $r_2 = \text{YYYY}$ and again we have a contradiction as both YNNN *and* YNNY *differ on two question from* YYYY*.*

The cases where $r_1 = \text{YNNN}$ or $r_1 = \text{YNNY}$ create similar contradictions. This shows that $\mathcal{D}$ is not rationalizable.

Before introducing our second condition we first present some useful intermediate consequences of Condition 1. To do so, we introduce a few new concepts. Define the collection $\Delta_{\mathcal{D}}$ as the set of all differences between pairs of profiles in $\overline{\mathcal{D}}$.

$$\Delta_{\mathcal{D}} = \{\Delta(\overline{r}_0, \overline{r}_1) | \overline{r}_0, \overline{r}_1 \in \overline{\mathcal{D}}\}.$$

We say that $\delta \in \Delta_{\mathcal{D}}$ is an *atom* if it cannot be expressed as the union of other elements in $\Delta_{\mathcal{D}}$, each being a strict subset of δ. Intuitively, atoms are the minimal building blocks of the difference structure.

Definition 4. *A set $\delta \in \Delta_{\mathcal{D}}$ is an atom if there do not exist sets $\delta_1, \ldots, \delta_n \in \Delta_{\mathcal{D}}$ such that for all $i = 1, \ldots, n$, $\delta_i \subset \delta, \delta_i \neq \delta$ and $\cup_{i=1}^n \delta_i = \delta$. The set of atoms in $\Delta_{\mathcal{D}}$ is denoted by $\mathcal{A}_{\mathcal{D}}$.*

The following is a consequence of Condition 1.

Lemma 1. *If $\mathcal{D}$ satisfies Condition 1, then all atoms in $\mathcal{A}_{\mathcal{D}}$ are pairwise disjoint. In particular, if $\delta_1, \delta_2 \in \mathcal{A}_{\mathcal{D}}$ and $\delta_1 \neq \delta_2$ then $\delta_1 \cap \delta_2 = \emptyset$.*

Example 2. *Continuing our previous example, we had:*

$$\overline{\mathcal{D}} = \{\text{YYYY}, \text{YYYN}, \text{YNNN}, \text{YNNY}\}$$

This dataset satisfies Condition 1. We compute:

$$\Delta_{\mathcal{D}} = \{\{4\}, \{2, 3, 4\}, \{2, 3\}\}.$$

From this, we identify the atoms:

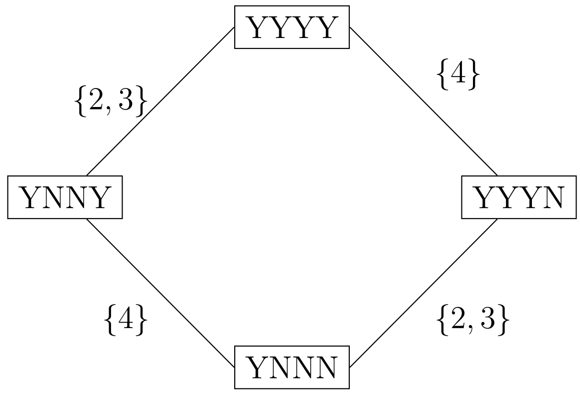
$$\mathcal{A}_{\mathcal{D}} = \{\{2, 3\}, \{4\}\}$$

These atoms are clearly disjoint, as required by Lemma 1.

We now use this structure to define an undirected graph $\mathcal{G}_{\mathcal{D}}$. The vertices of $\mathcal{G}_{\mathcal{D}}$ are the profiles in $\overline{\mathcal{D}}$ and we construct an undirected edge between two profiles $\overline{r}_0$ and $\overline{r}_1$ if and only if $\Delta(\overline{r}_0, \overline{r}_1) \in \mathcal{A}_{\mathcal{D}}$, that is, if their difference is an atom.

Example 3. *Continuing with our previous example, figure 6 illustrates the network $\mathcal{G}_{\mathcal{D}}$. Each vertex corresponds to a profile in $\mathcal{D}$ and each edge $(\overline{r}_0, \overline{r}_1)$ is labelled by the corresponding atom $\Delta(\overline{r}_0, \overline{r}_1)$. As shown, the resulting graph forms a cycle.*

Figure 3: A network $\mathcal{G}_{\mathcal{D}}$



Condition 1 imposes strong topological constraints on the network $\mathcal{G}_{\mathcal{D}}$, formalized in the following lemma:

Lemma 2. *If Condition 1 holds, then:*

1. *Every vertex in the network $\mathcal{G}_{\mathcal{D}}$ has at most two edges.*
2. *The network $\mathcal{G}_{\mathcal{D}}$ is connected, i.e. every two vertices are connected by a sequence of edges.*

Given Lemma 2, the structure of $\mathcal{G}_{\mathcal{D}}$ must be either a path or a cycle. Our next condition rules out the possibility of cycles.

Condition 2. *For all distinct edges $(\overline{r}_0, \overline{r}_1)$ and $(\overline{r}_2, \overline{r}_3)$ in $\mathcal{G}_{\mathcal{D}}$: $\Delta(\overline{r}_0, \overline{r}_1) \neq \Delta(\overline{r}_2, \overline{r}_3)$.*

Condition 2 ensures that each atom in $\mathcal{A}_{\mathcal{D}}$ labels at most one edge in the graph $\mathcal{G}_{\mathcal{D}}$. To see how this excludes cycles, suppose for contradiction that $\mathcal{G}_{\mathcal{D}}$ forms a cycle and that Condition 2 holds. Pick any profile $\overline{r}_0 \in \overline{\mathcal{D}}$ and consider an edge $(\overline{r}_0, \overline{r}_1)$ in the cycle. Let $j \in \Delta(\overline{r}_0, \overline{r}_1)$. As we traverse the cycle and eventually return to $\overline{r}_0$, there must exist another edge, say $(\overline{r}_2, \overline{r}_3)$, such that $j \in \Delta(\overline{r}_2, \overline{r}_3)$. However, since distinct atoms are disjoint, this implies: $\Delta(\overline{r}_0, \overline{r}_1) = \Delta(\overline{r}_2, \overline{r}_3)$, contradicting Condition 2.
Importantly, however, Condition 2 is stronger than merely excluding cycles as the following example illustrates.

Example 4. *Consider the dataset:*

$$\overline{\mathcal{D}} = \{\text{YYYYY}, \text{YNYYY}, \text{YNNYY}, \text{YNNNY}, \text{YYNNY}\}.$$

One can verify that Condition 1 is satisfied. The corresponding network $\mathcal{G}_{\mathcal{D}}$ is shown in Figure 4. Although $\mathcal{G}_{\mathcal{D}}$ is a path (not a cycle), the atom $\{2\}$ appears on two distinct edges, violating Condition 2. We now show that this dataset is indeed not rationalizable by the Hyperplane model. Assume, towards a contradiction, that it is. So there exists a response model (r, ℓ) such that $\mathcal{R}(r, \ell) = (r_1, \ldots, r_5)$ contain all five profiles in $\overline{\mathcal{D}}$:

$$\{r_1, \ldots, r_5\} = \{\text{YYYYY}, \text{YNYYY}, \text{YNNYY}, \text{YNNNY}, \text{YYNNY}\}$$

Figure 4: $\mathcal{G}_\mathcal{D}$ is a path but Condition 2 is not satisfied

Suppose $r_1 =$ YYYYY*. Then, since* r_2 *differs from* r_1 *on a single question, we must have* $r_2 =$ YNYYY*. Continuing* $r_3 =$ YNNYY*,* $r_4 =$ YNNNY *and* $r_5 =$ YYNNY*. This implies the ranking* $\ell = 2, 3, 4, 2$ *which is invalid due to the repetition of question* 2*. Trying other starting points (e.g.* $r_1 =$ YNYYY*) leads to similar contradictions.*

Condition 2 is also necessary for rationalizability. To see why, suppose that $\mathcal{D}$ is rationalizable. Then there is a decision model (r, ℓ) such that every profile in $\overline{\mathcal{D}}$ appears in $\mathcal{R}(r, \ell) = (r_1, \ldots, r_J)$. Consider the response profiles $\overline{r}_0, \overline{r}_1, \overline{r}_2, \overline{r}_3 \in \overline{\mathcal{D}}$ such that both $\Delta(\overline{r}_0, \overline{r}_1)$ and $\Delta(\overline{r}_2, \overline{r}_3)$ are atoms. Then $\overline{r}_0, \overline{r}_1, \overline{r}_2, \overline{r}_3 \in \{r_1, \ldots, r_J\}$. Suppose $\overline{r}_0 = r_i, \overline{r}_1 = r_j, \overline{r}_2 = r_k$ and $\overline{r}_3 = \overline{r}_s$ such that $i < j$ and $k < s$, and, without loss of generality, $i \leq k$. So:

$$\Delta(r_i, r_j) = \{q_i, \ldots, q_{j-1}\}, \qquad \Delta(r_k, r_s) = \{q_k, \ldots, q_{s-1}\}.$$

We now consider several cases:

- If $i < j \leq k < s$, then $\Delta(r_i, r_j) \cap \Delta(r_k, r_s) = \emptyset$ so Condition 2 is satisfied.
- If $i < k < j$, then,
$$\Delta(r_i, r_j) = \Delta(r_i, r_k) \cup \Delta(r_k, r_j),$$
contradicting the condition that $\Delta(r_i, r_j)$ is an atom.
- If $i = k < j < s$ then:
$$\Delta(r_k, r_s) = \Delta(r_k, r_j) \cup \Delta(r_j, r_s),$$
again contradicting the atomicity of $\Delta(r_k, r_s)$
- Finally if $i = k < s \leq j$, then
$$\Delta(r_i, r_j) = \Delta(r_i, r_s) \cup \Delta(r_s, r_j),$$
which also contradicts the assumption that $\Delta(r_i, r_j)$ is an atom.

We now state our main characterization result:

Theorem 2. *A dataset* $\mathcal{D}$ *is rationalizable by the Hyperplane model if and only if Conditions 1 and 2 are satisfied.*

Note that Conditions 1 and 2 also limits the size of $\overline{\mathcal{D}}$ to J or less. Indeed, if $\mathcal{D}$ has $m > J$ profiles and given that the network $\mathcal{G}_\mathcal{D}$ is connected (Lemma 2), there should be at least $m-1 \geq J$ edges. Every edge is labelled with an atom, which is a non-empty subset of $\{2, \ldots, J\}$. This means that $\mathcal{G}_\mathcal{D}$ should have at least two edges whose labels have a non-empty intersection. Lemma 1 then shows that these labels (begin atoms) should be equal, contradicting Condition 2.

To verify Conditions 1 and 2 a first check should be that $\overline{\mathcal{D}}$ has size J or less. Given this, constructing the differences and computing the atoms can be done in $O(J^2)$ steps. This allows us to verify Condition 1 in $O(J^2)$ steps. Next, constructing the graph $\mathcal{G}_\mathcal{D}$ and verifying Condition 2 also takes $O(J^2)$ steps so the total running time is at most quadratic in J.

5 The Consecutive Ones And Circular Consecutive Ones Models

The paper most closely related to ours in terms of topic is by Ballester and Apesteguia (2024). They propose a model in which each respondent $t \leq T$ can be represented by a one-dimensional ideal point $\theta^t \in \mathbb{R}$ and a threshold level $\tau^t \geq 0$. Each question $j \leq J$ is represented by a real number $q^j \in \mathbb{R}$. A respondent answers Y to question j if and only if the distance between θ^t and q^j is below this threshold, i.e. if $|\theta^t - q^j| \leq \tau^t$.

As shown in their paper, a dataset is rationalizable under this model if and only if the set of response profiles satisfies the Consecutive Ones (CO) property. This means that it is possible to give a complete ranking of the J questions such that, for every respondent, all questions to which they respond Y appear as a contiguous block. In other words, if a respondent answers Y to questions i and k, then she must also answer Y to any question j that lies between i and k in the ranking.

Given such ranking, the number of distinct response profiles consistent with the CO property is:[11]

$$1 + \frac{J(J+1)}{2}.$$

This is always greater or equal to $2J$ (for $J \geq 1$) and exceeds it for all $J > 2$, implying that there are many datasets rationalizable under the CO model that cannot be rationalized by our Hyperplane model. For instance, any dataset with more than $2J$ profiles satisfying the CO property falls outside the scope of our model.

The converse also holds: there are datasets rationalizable by the Hyperplane model that are not consistent with the CO model. As an example, consider the dataset:

$$\mathcal{D} = \{\text{YNY}, \text{YYN}, \text{NYY}\}.$$

This dataset fails the CO condition under any ordering of the questions. For example:

- Under the ordering $1 < 2 < 3$, the first profile violates the CO property
- Under $1 < 3 < 2$, the second profile violates the CO property
- Under $2 < 1 < 3$, the third does

All other permutations are inversions of these and yield similar violations. On the other hand, this dataset is rationalizable by the Hyperplane model. Consider the starting profile $r = \text{YNY}$ and the ranking $\ell = (3, 2)$. Algorithm 1 then generates the sequence:

$$\mathcal{R}(r, \ell) = (\text{YNY}, \text{YNN}, \text{YYN})$$

The first two profiles of $\mathcal{D}$ are directly included in $\mathcal{R}(r, \ell)$ while the third profile NYY is the negation of YNN. This demonstrates that the Hyperplane model and the CO model are not nested.

[11] One way to see this is to look at the first occurrence of Y in the ranking of the questions. There are J profiles with the CO property for which the the first Y occurs for the first question, $J - 1$ profiles with the CO property for which the first Y occurs for the second question in the ranking, and so on. Adding all these together gives $J(J - 1)/2$ profiles with the CO property. Adding the profile that answers N to all questions gives the final result.

Circular Consecutive ones There is a natural extension of the Ballester-Apesteguia CO model in which questions and ideal points are not placed on a line but on a circle. In this circular variant, each respondent t has a threshold τ^t and answers Y to question q^j if and only if the angular distance between θ^t and q^j is below τ^t.

The response profiles generated by this model are the ones that satisfy what is called the Circular Consecutive Ones (CCO) property: there exists a circular ordering of the questions such that, for each respondent, the questions answered Y form a contiguous arc. Equivalently, if we cut the circle at any point, one obtains a linear ordering of the questions in which either all Y answers or all N answers are consecutive. Note that this immediately shows that from an empirical point of view, the CCO model generalizes the CO model.

For a given linear ordering of the J questions, the number of distinct profiles satisfying the CCO property is:[12]

$$J(J-1)+2,$$

which is always greater or equal to $2J$ (for $J \geq 1$) and exceeds $2J$ for all $J > 2$. Thus, this model can rationalize a broader class of datasets than the Hyperplane and CO model. In fact, the Hyperplane model can also be seen as a special case of the CCO model. Suppose we represent ideal points θ^t and questions q^j as points in the plane and project them radially onto the unit circle. Then, each respondent answers Y to all questions within 90 degrees of their projected ideal point. This corresponds to a CCO model with a fixed threshold $\tau^t = 90°$, making our model isomorphic to a restricted version of the CCCO model.

In summary, the CCO model has weaker testable implications as it encompasses both the Hyperplane model and the Ballester-Apesteguia CO model as special cases. On the other hand, neither the Hyperplane model nor the CO model subsumes the other, highlighting that these frameworks capture distinct and non-nested patterns of rationalizability.

6 An Illustration

6.1 Goodness-of-fit And Predictive Success

In practice, it is unlikely that a given survey dataset $\mathcal{D}$ fully satisfies the necessary and sufficient Conditions 1 and 2 required for rationalizability by the Hyperplane model. For instance, the number of distinct observed response profiles in $\mathcal{D}$ often exceeds $2J$, especially when both the number of questions J and the number of respondents T are large. Nevertheless, it remains valuable to assess how well the model can account for a substantial portion of the observed responses. From this perspective a more valuable approach is to evaluate the degree to which a dataset deviates from being rationalizable. We'll do this by constructing a simple goodness-of-fit index that quantifies the largest fraction of the dataset that can be rationalized.

In particular, given a dataset $\mathcal{D} = (r^t)_{t \leq T}$, we define v^H to be the largest fraction of respondents whose response profiles are consistent with the Hyperplane model:

$$v^H = \max_{S \subset \{1,\dots,T\}} \frac{|S|}{T} \text{ s.t. } (r^t)_{t \in S} \text{ is rationalizable by the Hyperplane model.}$$

In a similar way we can define v^{CO} and v^{CCO} as the largest fraction of respondents whose choices are jointly consistent with the CO and CCO models. Computing v^H, v^{CO} and v^{CCO}

[12] There are $J(J+1)/2+1$ profiles that satisfy the CO property. Negating these gives profiles for which the N answers are consecutive. This gives in total $J(J+1)+2$ profiles. However, we need to correct for double counting, i.e. the J CO profiles that start with Y and the J CO profiles that end with a Y. So in total we have $J(J-1)+2$ profiles that satisfy the CCO property (for a given ranking of the questions).

is not trivial. Appendix B provides a dynamic programming approach for computing v^H, and two integer programming problems to compute v^{CO} and v^{CCO}.

Like our model, the CO and CCO models can be classified as area theories, meaning that they predict that a dataset should be found in a particular region of all possible data sets (Selten, 1991). As previously noted, the CCO model is the least restrictive of the three since it subsumes both the CO and Hyperplane models. Consequently, directly comparing the measures v^H, v^{CO} and v^{CCO} will inherently favor the CCO model due to its lower restrictiveness, i.e. for any dataset $v^{CCO} \geq v^H$ and $v^{CCO} \geq v^{CO}$. To correct for these differences in permissiveness, it is customary to compute a so called predictive success measure that penalizes models that are less stringent.[13] Usually, the restrictiveness of a model by seeing how it performs when applied to randomly generated, 'non-rational' data. To this end, we generate many synthetic (random) datasets and we calculate the (average) maximal proportion of responses that the three models can rationalize (i.e. the mean of v^H, v^{CO} and v^{CCO} over these randomly generated datasets). Denoting these averages by v^H_{sim}, v^{CO}_{sim} and v^{CCO}_{sim} we can then evaluate the models performance against what one could expect to be its performance if the true dataset would have been random itself by computing the differences $v^H - v^H_{sim}, v^{CO} - v^{CO}_{sim}$ and $v^{CCO} - v^{CCO}_{sim}$. These predictive success measures provide a more robust measure of the models' performances. We consider two ways of simulating random data.[14]

- **Sim 1: Independent Question-Wise Sampling.** For each question $j \leq J$, we compute the empirical probability p_j that a (random) respondent answers Y. This is done by taking for ach question j the fraction of respondents t for which $r^t(j) = \text{Y}$. Next, we generate synthetic response profiles r^t by independently sampling each respondent's answer $r^t(j)$ to question j according to the probability p_j. These simulations preserve the marginal distributions of responses for each question as in the data, but removes all correlations between the various answers for a given respondent as well as inter-respondent correlation for a given question.

- **Sim 2: Independent Respondent-Wise Sampling.** For each respondent $t \leq T$, we calculate the empirical probability p_t that the respondent t answers Y to a question. We do this by computing the fraction of Y responses over the various questions for this particular respondent. We then simulate for each respondent answers $r^t(j)$ by independently sampling each response according to p_t. This preserves individual response tendencies while it eliminates correlations across questions and the joint structure of responses over different respondents for a same question.

6.2 Data

We illustrate our model using data from the seventh wave of the World Values Survey (Inglehart et al., 2014). This dataset contains responses from 94,248 individuals across more than 100 countries. The survey consists of 259 questions covering a broad range of topics, including attitudes towards democracy, public security, and gender equality and other key social and political issues.

Most of the survey questions use Likert-type scales, asking respondents to indicate their level of agreement with a given statement. To adapt this to our binary-response framework, we convert these responses into dichotomous (yes/no) responses. For example, consider a typical survey item that asks respondents to express their agreement with the following statement:

[13] See Selten (1991) for the general framework and Beatty and Crawford (2011) who applied Selten's predictive success measure to revealed preference theory.

[14] The values v^H_S, v^{CO}_s and v^{CCO}_s are based on 1,000 simulated datasets.

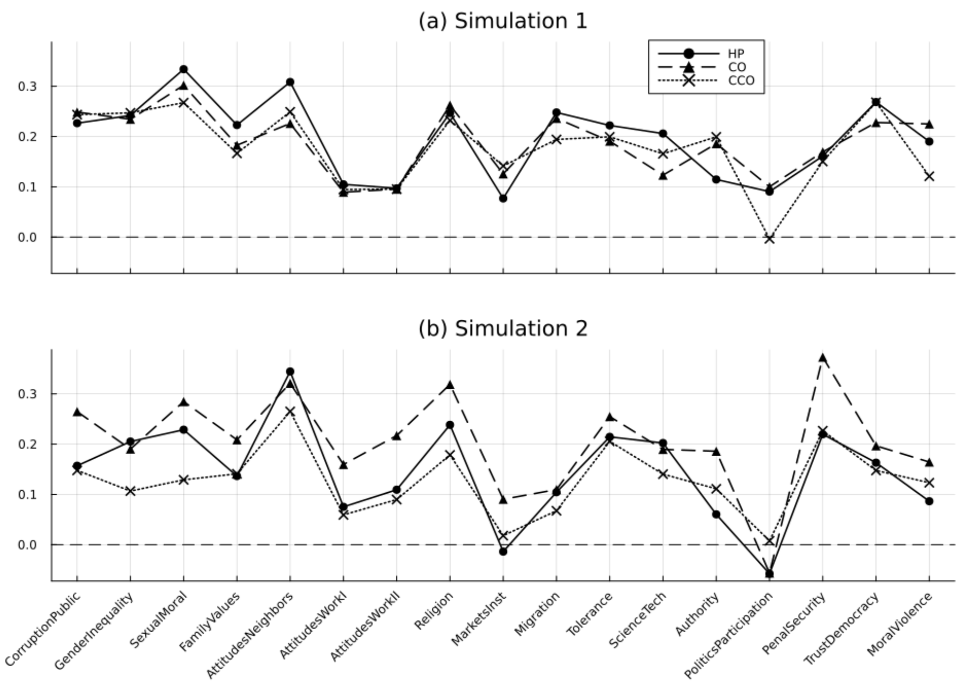
- *"A university education is more important for a boy than for a girl"*.

Respondents could choose from four options: *Strongly agree*, *Agree*, *Disagree*, and *Strongly Disagree*. We convert these by classifying both *Strongly Agree* and *Agree* as Y and *Disagree* and *Strongly Disagree* as N. For questions with an odd number of Likert-scale options, including a neutral midpoint (e.g. *Neither Agree nor Disagree*), we randomly assign the neutral responses to Y or N with equal probability. Survey items that could not be clearly or consistently transformed into a binary format were excluded from the analysis.

6.3 Overall Performance

To evaluate the performance of the Hyperplane model, we apply our methodology to various thematic sections of the World Values Survey. Some of these sections correspond to sections already defined in the original survey, while others combine questions that were originally dispersed across different parts of the questionnaire but shared a common theme. In total, we constructed 17 thematic groups, each consisting of 7 questions, which is the maximum that allows us to carry out the analysis while maintaining computational feasibility within a reasonable time frame.[15] The topics covered by these groups and the specific questions included for each group is given in Table 2 in Appendix C.[16]

Figure 5: Predictive success of Hyperplane CO and CCO models



The predictive success of the Hyperplane, CO and CCO models for the different topics are given in Figure 5.[17] The results reveal substantial variation in the model's performance across different topics. The results overall show that the predictive success of the various models is almost always greater than zero. This confirms that the models' performances are not

[15] The dynamic programming procedure for the hyperplane model can handle much larger sets of questions, but the integer programming (especially for the CCO model) is much more computationally intensive.

[16] As an additional exercise, we also asked an AI to group questions into thematic similar topics. The groupings obtained from this were almost identical to our categorization.

[17] Detailed figures are given in Table 3 in Appendix C.

attributable to random chance. In addition, the performance of the Hyperplane, CO and CCO are also highly correlated over the different sections, and over the two simulation exercises.

Concerning the ranking of the different models, the first predictive success measure (based on simulation 1) tends to favour the Hyperplane model for a majority of topics (8-9 out of 17) as the best performing model. The CO model is the second best as it dominates the other two for 5 topics. The CCO model is the best in only 3-4 topics (there is one case where the Hyperplane and CCO give identical predictive success). For the second predictive success measure (based on simulation 2) the CO model outperforms the other two for 13 topics, the Hyperplane model is best in 3 topics while the CCO model dominates the other 2 in only 1 topic. Overall, there does not seem to be one single model that overall provides a better fit compared to the other two based on the predictive success measure. On the other hand, for a large majority of the topics, the CCO model is seems to be dominated by either the CO or the Hyperplane model.

6.4 A Case Study

In this section, we present a detailed example to illustrate how our model can also be used to uncover underlying structures in the data and to highlight the differences and alignments of these structures when evaluating the data based on our Hyperplane model versus the CO model or the CCO model. For our case study, we focus on a subset of $J = 6$ questions from the World Values Survey[18] which we label as Q1 up to Q6:

- Q1 We would like to know your opinion about people from other countries who come to live in [your country]–the immigrants. Do you positively evaluate their impact on the development of [your country]?

- From your point of view, what have been the effects of immigration on the development of [your country]? For each of the following statements about the effect of immigration, please indicate whether you agree or disagree:

  - Q2 Fills important job vacancies
  - Q3 Strengthens cultural diversity
  - Q4 Increases the crime rate.
  - Q5 Increases the risk of terrorism
  - Q6 Leads to social conflict.

The goodness-of-fit index for this set of questions, as well as the metrics for the simulations, are reported in Table 1.

Table 1: Goodness-of-fit across different models and simulations.

Model	v	v_{sim} simulation 1	v_{sim} simulation 2
Hyperplane	0.585	0.312	0.210
CO	0.659	0.422	0.339
CCO	0.796	0.615	0.530

The results indicate that all three models explain a significant share of the data. The hyperplane model has the lowest goodness-of-fit, but this is expected given larger permissiveness of the other

[18] These correspond to questions Q121, Q122, Q123, Q124, Q126, and Q129 in the original questionnaire.

models. However, if we evaluate the models based on how much better they do (on average) compared to the three simulations, we see that the Hyperplane model outperforms the other two.[19] However, it is important to emphasize that this is not a general result as for other questions groupings, the CO (or CCO) model may outperform the Hyperplane model.

Figure 6 illustrates the hyperplane arrangement corresponding to the selected six-question example that explain the largest share of data. Each region in the figure represents a distinct response profile, annotated with the proportion of respondents who selected it. The Figure reveals that the model effectively clusters related questions: Q4, Q5 and Q6 (focused on security concerns related to migration) together, as well as questions Q1, Q2 and Q3 which emphasize the potential positive contributions of migrants to society.

Figure 7 displays the question ordering that maximizes the share of data explained by the CO model. This model also clusters questions Q4, Q5 and Q6, but places Q1 on one end of the sequence and Q2 and Q3 on the opposite end, indicating a different structural interpretation compared to the Hyperplane model.

Finally, Figure 7 shows the circular ordering of questions selected by the CCO model. Like the other models, it groups Q4, Q5 and Q6 as well as Q1, Q2 and Q3. However, the spatial relationships differ: for instance, while the Hyperplane model places Q2 adjacent to Q6, the CCO model positions Q4 as being closer to Q6, reflecting a distinct interpretation of the underlying response patterns.

Figure 6: The hyperplane model

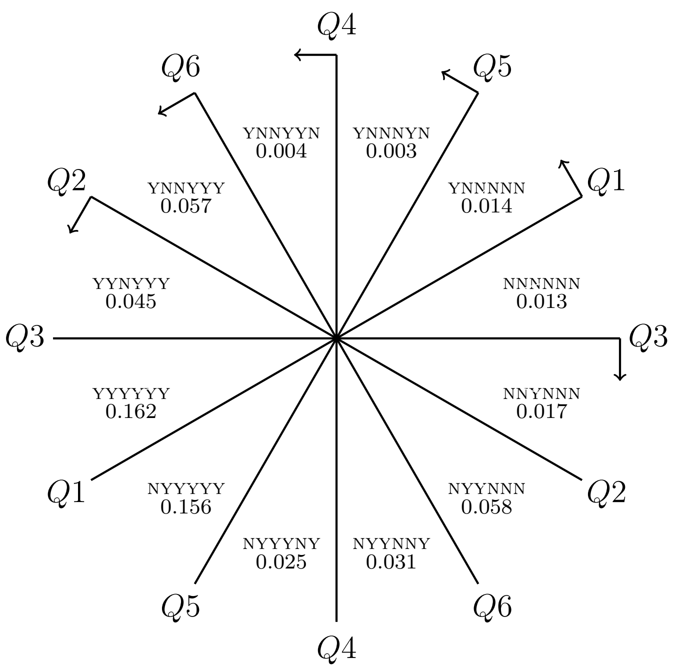


Figure 7: Consecutive ones model

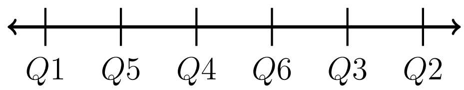
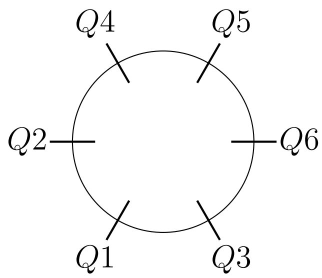


Figure 8: Circular consecutive ones model



7 Conclusion

We introduced the Hyperplane Model of Survey Response, which gives a decision theoretic framework for understanding how individuals respond to survey questions. The model is grounded in two intuitive consistency principles: (1) logical coherence with respect to negation of questions and (2) agreement with convex combinations of more extreme agreed upon

[19] These differences are 0.273 versus 0.237 and 0.181 for Simulation 1, 0.375 versus 0.32 and 0.266 for Simulation 2

questions. The model posits that each respondent can be represented by an ideal point in a multidimensional attribute space and agreement or disagreement with a statement is determined by the position of this ideal point relative to a hyperplane perpendicular to the vector containing the questions attributes.

We characterize the model's testable implications in the two-dimensional case. We show that the Hyperplane model is empirical distinguishable from the CO model. Using data from the World Values Survey, we showed that the Hyperplane model captures meaningful structure in survey responses of thematic domains, exceeding what would be expected by chance.

We see several avenues for future research. First of all, while our analysis focussed on the two-dimensional case for tractability, extending the rationalizability characterization to higher dimensional attribute space would enhance the model's realism and allow it to handle more complex survey data. Second, our current framework is limited to binary (agree/disagree) responses. Future work could explore how to incorporate ordinal or cardinal intensities of agreement. Third, including elements of strategic misreporting, social desirability bias or limited self-knowledge could yield a better and more realistic understanding of real world survey behaviour. Finally, improving the computational efficiency of the rationalizability test and goodness-of-fit measures would increase the model's practical utility and could help in developing more refined goodness-of-fit indices to evaluate the models' performances.

References

Aldrich, J. H., McKelvey, R. D., 1977. A method of scaling with applications to the 1968 and 1972 presidential elections. American Political Science Review 71, 111–130.

Almas, I., Attanasio, O., Pamela, J., 2024. Presidential address: Economics and measurement: New measures to model decision making. Econometrica 92, 947–978.

Ballester, M., Apesteguia, J., 2024. The rationalizability of survey response. Tech. rep.

Beatty, T. K. M., Crawford, I. A., 2011. How demanding is the revealed preference approach to demand. American Economic Review 101, 2782–2795.

Benjamin, D. J., Guzman, J. D., Fleurbaey, M., Heffetz, O., Kimball, M., 2003. What do happiness data mean? theory and survey evidence. Journal of the European Economic Association 21, 2377–2412.

Bertrand, M., Sendhil, M., 2001. Do people mean what they say? Implications for subjective survey data. American Economic Review: Papers and Proceedings of the Hudred Thirteenth Annual Meeting of the American Economics Association.-, 67–72.

Cavaillé, C., Chen, D. L., van Der Straeten, K., 2021. A decision-theoretic approach to understanding survey response: Likert versus quadratic voting for attitudinal research. Tech. rep.

Enelow, J. M., Hinich, M. J., 1984. The Spatial Theory of Voting. Cambridge University Press.

Falk, A., Neuber, T., Strack, P., 2021. Limited self-knowledge and survey response behavior. Tech. rep.

Festa, P., 2024. Encyclopedia of Optimization. Springer Nature Switzerland, Ch. Oriented Matroids.

Guttman, L., 1944. A basis for scaling qualitative data. American Sociology Survey 9, 139–150.

Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., et al., B. P. (Eds.), 2014. World Values Survey: Round Seven. Madrid: JD Systems Institute.

Javaras, K. N., Ripley, B. D., 2007. An ‘ ‘ unfolding” latent variable model for Likert attitude data. Journal of the American Statistical Association 102, 454–463.

Likert, R., 1932. A technique for the measurement of attitudes. Archives of Psychology 140, 1–55.

Liu, S., Netzer, N., 2023. Happy times: measuring happiness using response times. American Economic Review 113, 3289–3322.

Manski, C., 2004. Measuring expectations. Econometrica 72, 1329–1376.

Prelec, D., 2004. A Bayesian truth serum for subjective data. Science 306, 462–466.

Rossi, P. E., Gilula, Z., Allenby, G. M., 2001. Overcoming scale usage heterogeneity: A Bayesian hierarchical approach. Journal of the American Statistical Association 96, 20–31.

Selten, R., 1991. Properties of a measure of predictive success. Mathematical Social Sciences 21, 153–167.

Shor, P. W., 1991. Stretchability of psuedolines is NP-hard. DIMACS Series in Discrete Mathematics 4, 531–554.

Stantcheva, S., 2023. How to run surveys: a guide to creating your own identifying variation and revealing the invisible. Annual Review of Economics 15, 205–234.

Thurstone, L., 1928. Attitudes can be measured. American Journal of Sociology 33, 529–554.

van der Linden, W. J. (Ed.), 1997. Handbook of Item Response Theory. Routledge.

A Proofs

A.1 Proof of theorem 1

Let $A \subseteq \mathbb{R}^n \setminus \{0\}$ denote the questions for which the respondent answers Y, i.e.

$$A = \{q \in \mathbb{R}^n \setminus \{0\} : \psi(q) = \text{Y}\}.$$

Let us first show that A is convex. Suppose $x, y \in A$ and let $z = \alpha x + (1-\alpha)y$ ($\alpha \in [0,1]$). If $z \neq 0$, it follows immediately from Assumption 2 that $z \in A$. Let us now show that the case $z = 0$ cannot occur. Towards a contradiction asssume that $z = 0$ (meaning that $\alpha \notin \{0,1\}$), we consider two cases.

- If $\alpha = 1/2$ then $x = -y$ contradicting Assumption 1.
- If $\alpha \neq 1/2$ (implying that $x \neq -y$), pick $\varepsilon > 0$ such that $\varepsilon < \alpha < 1-\varepsilon$. Let $w = (\alpha-\varepsilon)x + (1-\alpha+\varepsilon)y = \varepsilon(y-x) \neq 0$. Then, by Assumption 2, as w is a convex combination of x and y, it must be in A. Also,

  $$-w = -\varepsilon(y-x) = (\alpha+\varepsilon)x + (1-\alpha-\varepsilon)y$$

  which is also a convex combination of x and y. As such, again by Assumption 2, $-w \in A$, Conclude that $\psi(w) = \psi(-w) = \text{Y}$, contradicting Assumption 1.

Next, from convexity of A and $0 \notin A$, it follows that there is a separating hyperplane $\theta = (\theta_1, \ldots, \theta_n) \neq 0$ through the origin, such that for all $q = (q_1, q_2, \ldots, q_n) \in A$,

$$0 \leq \sum_{i=1}^{n} \theta_i q_i = \langle \theta, q \rangle.$$

Now suppose $0 < \langle \theta, q \rangle$ but $\psi(q) = \text{N}$. Then by Assumption 1, $\psi(-q) = \text{Y}$, so $-q \in A$. But then, $\langle \theta, -q \rangle = -\langle \theta, q \rangle < 0$, a contradiction. Therefore $\psi(q) = \text{Y}$.
Similarly, if $\langle \theta, q \rangle < 0$, then $\langle \theta, -q \rangle > 0$, so $\psi(-q) = \text{Y}$, and by Assumption 1, $\psi(q) = \text{N}$.

A.2 Proof of Lemma 1

Assume, towards a contradiction, that there are two atoms δ_0, δ_1 and a question $j \leq J$, such that $j \in \delta_0 \cap \delta_1$ and $\delta_0 \neq \delta_1$. Let $\delta_0 = \Delta(\overline{r}_0, \overline{r}_1)$ and $\delta_1 = \Delta(\overline{r}_2, \overline{r}_3)$ for some profiles $\overline{r}_0, \overline{r}_1, \overline{r}_2$ and $\overline{r}_3$ in $\overline{\mathcal{D}}$.
First note that all four profiles $\overline{r}_0, \overline{r}_1, \overline{r}_2$ and $\overline{r}_3$ are distinct. To see this, assume, towards a contradiction that $\overline{r}_0 = \overline{r}_2$.

- If also $\overline{r}_1 = \overline{r}_3$ then $\delta_0 = \delta_1$ a contradiction.
- If $\overline{r}_1 \neq \overline{r}_3$. Then given Condition 1, two of the sets $\Delta(\overline{r}_0, \overline{r}_1) = \delta_0, \Delta(\overline{r}_0, \overline{r}_3) = \Delta(\overline{r}_2, \overline{r}_3) = \delta_1$ and $\Delta(\overline{r}_1, \overline{r}_3)$ must be disjoint and the third one should be the union of the other. This means that either $\delta_0 \cap \delta_1 = \emptyset$, which violates the assumption that $j \in \delta_0 \cap \delta_1$, or $\Delta(\overline{r}_0, \overline{r}_1) \cup \Delta(\overline{r}_1, \overline{r}_3) = \delta_1$, which violates the assumption that δ_1 is an atom, or $\Delta(\overline{r}_0, \overline{r}_3) \cup \Delta(\overline{r}_1, \overline{r}_3) = \delta_0$, which violates the assumption that δ_0 is an atom.

Now, consider the question $j \in \delta_0 \cap \delta_1$. Without loss of generality, assume that $\overline{r}_0(j) = \text{Y}$ and, hence, $\overline{r}_1(j) = \text{N}$ (as $j \in \delta_0 = \Delta(\overline{r}_0, \overline{r}_1)$). We consider two cases.

1. $\overline{r}_2(j) = \text{N}$. Then as $j \in \delta_1 = \Delta(\overline{r}_2, \overline{r}_3)$, we have $\overline{r}_3(j) = \text{Y}$.
   It follows that $j \in \Delta(\overline{r}_0, \overline{r}_1)$ and $j \in \Delta(\overline{r}_1, \overline{r}_3)$.
   Applying Condition 1 on the triple $\overline{r}_0, \overline{r}_1, \overline{r}_2$ shows that two of the differences $\delta_0, \Delta(\overline{r}_1, \overline{r}_2)$ and $\Delta(\overline{r}_0, \overline{r}_2)$ are disjoint and the third one is the union of the other. As δ_0 is an atom and $j \in \Delta(\overline{r}_0, \overline{r}_2)$ it follows that:
   $$\delta_0 \subset \Delta(\overline{r}_0, \overline{r}_2)$$
   Next, Applying Condition 1 on the triple $\overline{r}_0, \overline{r}_1, \overline{r}_3$ shows that two of the differences $\delta_0, \Delta(\overline{r}_0, \overline{r}_3)$ and $\Delta(\overline{r}_1, \overline{r}_3)$ are disjoint and the third one is the union of the other. As δ_0 is an atom and $j \in \Delta(\overline{r}_1, \overline{r}_3)$ it follows that:
   $$\delta_0 \subset \Delta(\overline{r}_1, \overline{r}_3).$$
   Now consider an arbitrary question $k \in \delta_0$ and, without loss of generality, assume $\overline{r}_0(k) = \text{Y}$. Then $\overline{r}_1(k) = \text{N}$ and as $k \in \delta_0 \subset \Delta(\overline{r}_0, \overline{r}_2)$, we obtain $\overline{r}_2(k) = \text{N}$. Similarly, $\overline{r}_1(k) = \text{N}$ and $k \in \delta_0 \subset \Delta(\overline{r}_1, \overline{r}_3)$, so $\overline{r}_3(k) = \text{Y}$. As $\overline{r}_2(k) = \text{N}$ and $\overline{r}_3(k) = \text{Y}$ we conclude that $k \in \delta_1$. As $k \in \delta_0$ was arbitrary, we conclude that $\delta_0 \subseteq \delta_1$.
2. $\overline{r}_2(j) = \text{Y}$. Then as $j \in \delta_1 = \Delta(\overline{r}_2, \overline{r}_3)$, we have $\overline{r}_3(j) = \text{N}$. This means that $j \in \Delta(\overline{r}_0, \overline{r}_3)$ and $j \in \Delta(\overline{r}_1, \overline{r}_2)$. Following a similar proof as in case 1 above but exchanging the profiles $\overline{r}_2$ and $\overline{r}_3$ shows that in this case also $\delta_0 \subseteq \delta_1$.

This shows that $\delta_0 \subseteq \delta_1$. Using a similar reasoning as in steps 1 and 2 above but now exchanging $\overline{r}_0$ by $\overline{r}_2$ and $\overline{r}_1$ by $\overline{r}_3$ shows that also $\delta_1 \subseteq \delta_0$. This shows that $\delta_0 = \delta_1$, a contradiction

A.3 Proof of Lemma 2

For part 1, assume, towards a contradiction, that $\overline{r}_0 \in \overline{\mathcal{D}}$ and that $\overline{r}_0$ has three or more edges. This means that there exists $\overline{r}_1, \overline{r}_2, \overline{r}_3 \in \overline{\mathcal{D}}$ such that, $\Delta(\overline{r}_0, \overline{r}_1), \Delta(\overline{r}_0, \overline{r}_2)$ and $\Delta(\overline{r}_0, \overline{r}_3)$ are atoms. Applying Condition 1 on the various triples involving $\overline{r}_0$ gives:

$$\Delta(\overline{r}_0, \overline{r}_1) \cup \Delta(\overline{r}_0, \overline{r}_2) = \Delta(\overline{r}_1, \overline{r}_2),$$
$$\Delta(\overline{r}_0, \overline{r}_2) \cup \Delta(\overline{r}_0, \overline{r}_3) = \Delta(\overline{r}_2, \overline{r}_3),$$
$$\Delta(\overline{r}_0, \overline{r}_1) \cup \Delta(\overline{r}_0, \overline{r}_3) = \Delta(\overline{r}_1, \overline{r}_3)$$

This shows that the three sets $\Delta(\overline{r}_1, \overline{r}_2), \Delta(\overline{r}_1, \overline{r}_3), \Delta(\overline{r}_2, \overline{r}_3)$ have pairwise non-empty intersections, which contradicts Condition 1 when applied to the triple $\overline{r}_1, \overline{r}_2, \overline{r}_3$

For part 2, assume that $\mathcal{G}_{\mathcal{D}}$ is not connected. Then its nodes can be partitioned into two non-empty sets that are not connected to each other. Let A, B be such two subsets of $\overline{\mathcal{D}}$, i.e. $A \cap B = \emptyset$, $A \cup B = \overline{\mathcal{D}}$ and there is no edge between any vertex in A and any vertex in B.

Let $\overline{r}_0 \in A$ and $\overline{r}_1 \in B$ be two vertices that solve:

$$\min |\Delta(\overline{r}_0, \overline{r}_1)| \text{ s.t. } \overline{r}_0 \in A \text{ and } \overline{r}_1 \in B.$$

Let us show that $\Delta(\overline{r}_0, \overline{r}_1)$ is an atom, which will give the desired contradiction, as then $(\overline{r}_0, \overline{r}_1)$ is an edge in $\mathcal{G}_{\mathcal{D}}$.

Towards a contradiction, if the difference $\Delta(\overline{r}_0, \overline{r}_1)$ is not an atom, then it is the union of other elements in $\mathcal{A}_{\mathcal{D}}$. So, it is possible to find $\overline{r}_2, \overline{r}_3 \in \overline{\mathcal{D}}$ such that $\Delta(\overline{r}_2, \overline{r}_3)$ is an atom, $\Delta(\overline{r}_2, \overline{r}_3) \neq \Delta(\overline{r}_0, \overline{r}_1)$ and $\Delta(\overline{r}_2, \overline{r}_3) \subset \Delta(\overline{r}_0, \overline{r}_1)$. Note that it is not possible that one of the nodes $\overline{r}_2$ or $\overline{r}_3$ is in A and the other is in B as this would contradict the definition of $\overline{r}_0, \overline{r}_1$.

Without loss of generality, assume that $\overline{r}_2, \overline{r}_3 \in B$. Note that this implies first and foremost (by definition of $\overline{r}_0$ and $\overline{r}_1$) that $|\Delta(\overline{r}_0, \overline{r}_1)| \leq |\Delta(\overline{r}_0, \overline{r}_2)|$ and $|\Delta(\overline{r}_0, \overline{r}_1)| \leq |\Delta(\overline{r}_0, \overline{r}_3)|$. Let $j \in \Delta(\overline{r}_2, \overline{r}_3) \subset \Delta(\overline{r}_0, \overline{r}_1)$.

Without loss of generality, assume that $\overline{r}_0(j) = \text{Y}$ and hence $\overline{r}_1(j) = \text{N}$. There are two cases to consider.

- If $\overline{r}_2(j) = \text{N}$ and $\overline{r}_3(j) = \text{Y}$ then $j \in \Delta(\overline{r}_0, \overline{r}_2)$ and $j \in \Delta(\overline{r}_1, \overline{r}_3)$. Applying Condition 1 to the triple $\overline{r}_0, \overline{r}_1, \overline{r}_2$, together with $\Delta(\overline{r}_0, \overline{r}_1) \cap \Delta(\overline{r}_0, \overline{r}_2) \neq \emptyset$ and $|\Delta(\overline{r}_0, \overline{r}_1)| \leq |\Delta(\overline{r}_0, \overline{r}_2)|$ gives:

  $$\Delta(\overline{r}_0, \overline{r}_1) \neq \Delta(\overline{r}_0, \overline{r}_2) \text{ and } \Delta(\overline{r}_0, \overline{r}_1) \subset \Delta(\overline{r}_0, \overline{r}_2).$$

  Also, applying Condition 1 to the triple $\overline{r}_0, \overline{r}_1, \overline{r}_3$ together with $\Delta(\overline{r}_0, \overline{r}_1) \cap \Delta(\overline{r}_1, \overline{r}_3) \neq \emptyset$ gives that either $\Delta(\overline{r}_0, \overline{r}_1) \cup \Delta(\overline{r}_0, \overline{r}_3) = \Delta(\overline{r}_1, \overline{r}_3)$ or $\Delta(\overline{r}_0, \overline{r}_3) \cup \Delta(\overline{r}_1, \overline{r}_3) = \Delta(\overline{r}_0, \overline{r}_1)$. The latter implies that $\Delta(\overline{r}_0, \overline{r}_3) \neq \Delta(\overline{r}_0, \overline{r}_1)$ and $\Delta(\overline{r}_0, \overline{r}_3) \subset \Delta(\overline{r}_0, \overline{r}_1)$ which contradicts $|\Delta(\overline{r}_0, \overline{r}_1)| \leq |\Delta(\overline{r}_0, \overline{r}_3)|$.

  Conclude that the first must be the case, so:

  $$\Delta(\overline{r}_0, \overline{r}_1) \neq \Delta(\overline{r}_1, \overline{r}_3) \text{ and } \Delta(\overline{r}_0, \overline{r}_1) \subset \Delta(\overline{r}_1, \overline{r}_3).$$

  Now let $k \in \Delta(\overline{r}_0, \overline{r}_1)$. For example $\overline{r}_0(k) = \text{Y}$ and $\overline{r}_1(k) = \text{N}$. Then $k \in \Delta(\overline{r}_0, \overline{r}_2)$ and $k \in \Delta(\overline{r}_1, \overline{r}_3)$. As such $\overline{r}_2(k) = \text{N}$ and $\overline{r}_3(k) = \text{Y}$, so $k \in \Delta(\overline{r}_2, \overline{r}_3)$. As k was arbitrarily, this shows that $\Delta(\overline{r}_0, \overline{r}_1) \subseteq \Delta(\overline{r}_2, \overline{r}_3)$, a contradiction.

- if $\overline{r}_2(j) = \text{Y}$ and $\overline{r}_3(j) = \text{N}$ the proof is entirely analogues–by exchanging $\overline{r}_2$ and $\overline{r}_3$.

A.4 Proof of Theorem 2

The arguments in the main text shows that the two conditions are necessary. For sufficiency, assume that Conditions 1 and 2 are satisfied. Then Lemma 2 and the argument in the text shows that the network $\mathcal{G}_{\mathcal{D}}$ is a path. Consider the following construction of ℓ.

- Let $\overline{r}_1$ be a vertex in $\mathcal{G}_{\mathcal{D}}$ with a single edge–i.e. an endpoint of the path. Initialize ℓ to be the empty list and set $t = 1$
- While $t < |\overline{\mathcal{D}}|$.
  - Let $\overline{r}_{t+1}$ be the (unique) neighbour of $\overline{r}_t$ in $\overline{\mathcal{D}} \setminus \{\overline{r}_1, \ldots, \overline{r}_t\}$.
  - append to the list ℓ the questions in $\Delta(\overline{r}_t, \overline{r}_{t+1})$ in arbitrary order.
  - set $t := t + 1$.
- If there are remaining questions (different from 1) that are not in $\bigcup_t \Delta(\overline{r}_t, \overline{r}_{t+1})$, add them to the end of the list ℓ in arbitrary order.

As $\mathcal{G}_{\mathcal{D}}$ is a path, $\overline{r}_{t+1}$ is always well defined. Also, as every edge $\Delta(\overline{r}_t, \overline{r}_{t+1})$ is distinct—by Condition 2)—the list ℓ is well defined and contains (from the last step) all questions except question 1. It is also clear that the model $\mathcal{R}(\overline{r}_1, \ell)$ contains all profiles in $\overline{\mathcal{D}}$.

B Goodness-of-fit Indices

B.1 A Dynamic Programming Algorithm for the Hyperplane model

For a profile $\overline{r}$ let,

$$s(\overline{r}) = \frac{|\{t \in T | \overline{r}^t = \overline{r}\}|}{T},$$

be the fraction of respondents in the dataset for which $\overline{r}^t = \overline{r}$. For two profiles $\overline{r}_1, \overline{r}_2$ let $\mathcal{L}(\overline{r}_1, \overline{r}_2)$ denote the set of possible rankings of questions $\{2, \ldots, J\}$ on which $\overline{r}_1$ and $\overline{r}_2$ agree. For example if $\overline{r}_1$ = YYYY and $\overline{r}_2$ = YNYY then $\overline{r}_1$ and $\overline{r}_2$ agree on questions 3 and 4 so $\mathcal{L}(\overline{r}_1, \overline{r}_2)$ contains the rankings $(3, 4)$ and $(4, 3)$. Note that $\mathcal{L}(\overline{r}, \overline{r})$ includes all possible rankings of $\{2, \ldots, m\}$.

For a profile $\overline{r}_1$ and a ranking ℓ of the questions $\{2, \ldots, J\}$ let

$$v(\overline{r}_1, \ell) = \frac{|\{t \leq T | \overline{r}^t \in \mathcal{R}(\overline{r}, \ell)\}|}{T} = \sum_{\overline{r} \in \mathcal{R}(\overline{r}_1, \ell)} s(\overline{r})$$

be the fraction of respondents rationalizable by the response model $(\overline{r}_1, \ell)$. Then define:

$$v(\overline{r}_1) = \max_{\ell \in \mathcal{L}(\overline{r}_1, \overline{r}_1)} v(\overline{r}_1, \ell)$$

as the maximal fraction of respondents that is rationalizable by some response model $(\overline{r}_1, \ell)$ when we allow ℓ to vary over all rankings of $\{2, \ldots, J\}$ Note that:

$$v^H = \max_{\overline{r}_1} v(\overline{r}_1),$$

To compute $v(\overline{r}_1)$ we consider:

$$\pi(\overline{r}_1, \overline{r}_2) = \max_{\ell \in \mathcal{L}(\overline{r}_1, \overline{r}_2)} \sum_{\overline{r}_3 \in \mathcal{R}(\overline{r}_2, \ell)} s(\overline{r}_3).$$

Then:

$$v(\overline{r}_1) = \pi(\overline{r}_1, \overline{r}_1).$$

We compute $\pi(\overline{r}_1, \overline{r}_2)$ recursively:

$$\pi(\overline{r}_1, \overline{r}_2) = s(\overline{r}_2) + \max_k \pi(\overline{r}_1, \overline{r}_{2,-k}) \text{ s.t. } \overline{r}_1(k) = \overline{r}_2(k).$$

where $\overline{r}_{2,-k}$ is the response profile obtained by changing the response of question k in profile $\overline{r}_2$.

$$\overline{r}_{2,-k}(j) = \begin{cases} \overline{r}_2(j) \text{ if } k \neq j \\ \neg\overline{r}_2(j) \text{ if } k = j \end{cases}.$$

Each recursion step reduces the set of questions on which $\overline{r}_1$ and $\overline{r}_2$ agree by one until they only differ on one question as:

$$\{s|\overline{r}_1(s) = \overline{r}_{2,-k}(s)\} = \{s|\overline{r}_1(s) = \overline{r}_2(s)\} \setminus \{k\}.$$

The base case for the recursion (when $\overline{r}_1$ and $\overline{r}_2$ only differ on 1 question) is given by:

$$\pi(\overline{r}_1, \overline{r}_2) = s(\overline{r}_2).$$

Pruning the search space To compute $v^H = \max_{\overline{r}} v(\overline{r})$, one needs to determine, in principle, $v(\overline{r})$ for all possible profiles $\overline{r}$. We can avoid this we pruning the search space using a (known) lower bound on v^H. For example, if we compute $v(\overline{r})$ for one profile $\overline{r}$ then any profile $\overline{r}_1$ with:

$$s(\overline{r}_1) \leq \frac{v(\overline{r})}{J},$$

can be excluded from consideration. This is because any optimal model $(\overline{r}, \ell)$ must include at least one profile $\overline{r}_1 \in \mathcal{R}(r, \ell)$ with $s(\overline{r}_1) \geq \frac{v(\overline{r})}{J}$. In particular, profiles $\overline{r}_1$ with $s(\overline{r}_1) = 0$ can safely be ignored.

To improve efficiency, it might also be beneficial to initialize the search space with a profile $\overline{r}$ for which $s(\overline{r})$ is maximal. This provides a strong lower bound and helps eliminate many unpromising candidates early in the process.

B.2 Integer Programming Implementations for the CO and CCO Models

We denote questions by $i, j, k \in \{1, \ldots, J\}$ and response profiles by r. The goal is to identify the largest subset of profiles consistent with the CO or CCO models. We introduce the following parameters and variables:

Parameters

- $A(r, j) \in \{0, 1\}$: Indicates whether profile $\overline{r}$ answers Y to question j (1 for Y and 0 for N)
- $s(r) \in [0, 1]$: Fraction of respondents with response profile r in the dataset:

$$s(r) = \frac{|\{t \in T : r^t = r\}|}{T}.$$

Variables

- $R(i, j) \in \{0, 1\}$: Indicates whether question i precedes question j in the ranking.

- $Z(r) \in \{0,1\}$: Indicates whether profile r is included in the solution for the CO or CCO model.
- $O(r) \in \{0,1\}$: Indicates whether we test for the consecutive ones or the consecutive zeros property (used only in the CCO model)

The objective function maximizes the total weight of rationalizable profiles

$$\max_{R(i,j),Z(r)} \sum_r s(r)Z(r).$$

Constraints (in total $J(J-1)+J(J-1)(J-2)+J(J-1)(J-2)n$ constraints where n is the number of profiles in $\overline{\mathcal{D}}$):

1 Mutual exclusivity: Either question i precedes question j or vice versa:

$$\forall i,j \leq J : R(i,j)+R(j,i)=1.$$

2 Transitivity: if question i precedes question j and j precedes question k then i precedes k:

$$\forall i,j,k \leq J : R(i,j)+R(j,k) \leq 1+R(i,k).$$

For the CO property we add the following extra condition:

3 Consecutive Ones: if profile r is part of the solution and question i precedes question j and j precedes question k and r replies Y to i and k, then r replies Y to j:

$$\forall r, \forall i,j,k \leq J : Z(r)+R(i,j)+R(j,k)+A(r,i)+A(r,k) \leq 4+A(r,j).$$

For the CCO property, we add instead $O(r) \in \{0,1\}$ to indicate whether we test for the consecutive ones or consecutive zeros property. This gives the following conditions:

4 Circular Consecutive Ones: if profile r is part of the solution and if r must satisfy the consecutive ones property and if question i precedes question j and j precedes question k and r replies Y to both i and k then r replies Y to j.

$$\forall r, \forall i,j,k \leq J : Z(r)+O(r)+R(i,j)+R(j,k)+A(r,i)+A(r,k) \leq 5+A(r,j).$$

5 Circular consecutive zeros: if r is part of the solution and if r must satisfy the consecutive zeros property and if i precedes j and j precedes k and r replies N to both i and j, then $\overline{r}$ replies N to j.

$$\forall r, \forall i,j,k \leq J : Z(r)+(1-O(r))+R(i,j)+R(j,k)+(1-A(r,i))+(1-A(r,k)) \leq 5+(1-A(r,j)).$$

To compute the goodness-of-fit index for the CO model we maximize our objective subject to conditions 1, 2 and 3. For the CCO model we optimize the objective subject to conditions 1, 2, 4 and 5. This gives two Integer Programming problems.

C Additional Tables for Section 6

Table 2: The various topics and the questions

	Topic	Questions
1	Corruption in Public Institutions	Q112, Q113, Q115, Q116, Q118, Q120, Q181
2	Gender Inequality	Q29, Q30, Q31, Q33, Q35, Q189, Q249
3	Sexual Morality	Q25, Q36, Q183, Q184, Q185, Q186, Q193
4	Family Values	Q27, Q28, Q37, Q185, Q186, Q189, Q190
5	Attitudes Towards Neighbors	Q18, Q19, Q20, Q22, Q23, Q24, Q26
6	Attitudes Toward Work I	Q5, Q39, Q41, Q43, Q106, Q108, Q110
7	Attitudes Toward Work II	Q3, Q39, Q40, Q41, Q106, Q108, Q244
8	Religion	Q6, Q15, Q160, Q164, Q169, Q170, Q192
9	Views on Market Institutions	Q106, Q107, Q108, Q109, Q149, Q244, Q247
10	Attitudes Toward Migration	Q121, Q122, Q123, Q124, Q127, Q128, Q129
11	Tolerance	Q12, Q19, Q21, Q26, Q62, Q63, Q135
12	Views on Science and Technology	Q44, Q158, Q159, Q160, Q161, Q162, Q163
13	Views on Authority	Q45, Q150, Q196, Q197, Q198, Q245, Q248
14	Political Participation	Q212, Q213, Q214, Q215, Q216, Q219, Q220
15	Penalization and Security	Q69, Q70, Q150, Q179, Q180, Q181, Q195
16	Trust in Democratic Institutions	Q224, Q225, Q227, Q228, Q229, Q231, Q232
17	Morality and Violence	Q184, Q187, Q188, Q189, Q190, Q191, Q195

Note: the questions refer to the questions as given in the Codebook of the World Values Survey https://www.worldvaluessurvey.org/WVSDocumentationWV7.jsp.

Table 3: Goodness of fit results

Topic	Model		Simulation 1		Simulation 2	
		v	v_{sim}	st.dev	v_{sim}	st.dev
1	Hyperplane	0.464	0.237	(0.001)	0.307	(0.001)
	CO	0.696	0.448	(0.002)	0.432	(0.001)
	CCO	0.728	0.484	(0.002)	0.580	(0.001)
2	Hyperplane	0.521	0.279	(0.001)	0.316	(0.001)
	CO	0.683	0.449	(0.002)	0.493	(0.001)
	CCO	0.708	0.461	(0.002)	0.601	(0.001)
3	Hyperplane	0.586	0.252	(0.002)	0.357	(0.001)
	CO	0.738	0.436	(0.002)	0.454	(0.001)
	CCO	0.758	0.490	(0.002)	0.629	(0.001)
4	Hyperplane	0.521	0.298	(0.001)	0.385	(0.001)
	CO	0.665	0.482	(0.002)	0.456	(0.001)
	CCO	0.773	0.607	(0.002)	0.632	(0.001)
5	Hyperplane	0.661	0.353	(0.002)	0.316	(0.001)
	CO	0.809	0.583	(0.002)	0.488	(0.001)
	CCO	0.840	0.591	(0.002)	0.575	(0.001)

Topic	Model		Simulation 1		Simulation 2	
		v	v_{sim}	st.dev	v_{sim}	st.dev
6	Hyperplane	0.328	0.223	(0.001)	0.253	(0.001)
	CO	0.498	0.409	(0.002)	0.338	(0.001)
	CCO	0.586	0.492	(0.002)	0.527	(0.001)
7	Hyperplane	0.360	0.263	(0.001)	0.250	(0.001)
	CO	0.554	0.459	(0.002)	0.337	(0.001)
	CCO	0.616	0.521	(0.002)	0.526	(0.001)
8	Hyperplane	0.457	0.210	(0.001)	0.219	(0.001)
	CO	0.667	0.404	(0.002)	0.348	(0.001)
	CCO	0.677	0.444	(0.002)	0.498	(0.001)
9	Hyperplane	0.234	0.157	(0.001)	0.247	(0.001)
	CO	0.468	0.342	(0.002)	0.377	(0.001)
	CCO	0.536	0.395	(0.002)	0.518	(0.001)
10	Hyperplane	0.496	0.248	(0.001)	0.392	(0.001)
	CO	0.570	0.334	(0.002)	0.460	(0.001)
	CCO	0.709	0.515	(0.002)	0.641	(0.001)
11	Hyperplane	0.550	0.328	(0.002)	0.335	(0.001)
	CO	0.788	0.597	(0.002)	0.533	(0.001)
	CCO	0.802	0.603	(0.002)	0.597	(0.001)
12	Hyperplane	0.465	0.259	(0.001)	0.263	(0.001)
	CO	0.610	0.487	(0.002)	0.420	(0.001)
	CCO	0.669	0.503	(0.002)	0.528	(0.001)
13	Hyperplane	0.296	0.181	(0.001)	0.235	(0.001)
	CO	0.567	0.382	(0.002)	0.382	(0.001)
	CCO	0.619	0.419	(0.002)	0.508	(0.001)
14	Hyperplane	0.589	0.499	(0.002)	0.647	(0.001)
	CO	0.617	0.517	(0.002)	0.673	(0.001)
	CCO	0.815	0.818	(0.001)	0.807	(0.001)
15	Hyperplane	0.563	0.401	(0.002)	0.343	(0.001)
	CO	0.774	0.605	(0.002)	0.401	(0.001)
	CCO	0.838	0.688	(0.002)	0.611	(0.001)
16	Hyperplane	0.448	0.180	(0.002)	0.285	(0.001)
	CO	0.595	0.367	(0.002)	0.399	(0.001)
	CCO	0.680	0.412	(0.002)	0.532	(0.001)
17	Hyperplane	0.697	0.507	(0.002)	0.611	(0.001)
	CO	0.809	0.584	(0.002)	0.644	(0.001)
	CCO	0.907	0.786	(0.001)	0.784	(0.001)

Note: The first columun specifies the relevant topic. Column 2 refers to the model being tested. Column 3 contains the goodness of fit value v. Columns 4 and 6 contain the mean simulated values v_s for the first and second simulation. Standard deviations of these are between brackets in columns 5 and 7. These standard deviations are very small, rendering most comparisons highly statistically significant.