Quantitative Social Science Methods, I, Lecture Notes: Anchoring Vignettes for Interpersonally Incomparable Survey Responses

Gary King[1]
Institute for Quantitative Social Science
Harvard University

August 17, 2020

[1] GaryKing.org

Readings

GaryKing.org/vign

Readings

- "Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research" Gary King; Christopher J.L. Murray; Joshua A. Salomon; and Ajay Tandon, *American Political Science Review*, 2004

GaryKing.org/vign

Readings

- "Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research" Gary King; Christopher J.L. Murray; Joshua A. Salomon; and Ajay Tandon, *American Political Science Review*, 2004
- "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes" Gary King and Jonathan Wand, *Political Analysis*, 2007

GaryKing.org/vign

Readings

- "Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research" Gary King; Christopher J.L. Murray; Joshua A. Salomon; and Ajay Tandon, *American Political Science Review*, 2004
- "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes" Gary King and Jonathan Wand, *Political Analysis*, 2007
- "Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability" Daniel Hopkins and Gary King, *Public Opinion Quarterly*, 2010

GaryKing.org/vign

Readings

- "Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research" Gary King; Christopher J.L. Murray; Joshua A. Salomon; and Ajay Tandon, *American Political Science Review*, 2004
- "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes" Gary King and Jonathan Wand, *Political Analysis*, 2007
- "Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability" Daniel Hopkins and Gary King, *Public Opinion Quarterly*, 2010
- "Engaging Private and Information Technology in a Digital Age" National Research Council, j.mp/PrivInfoAge, 2007: p.86ff

GaryKing.org/vign

Anchoring Vignettes Address Two Problems

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...
  - The usual advice: You don't have a methodological problem. Get a theory and we'll know what to measure.

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...
  - The usual advice: You don't have a methodological problem. Get a theory and we'll know what to measure.
  - Result of more concreteness: more reliability, no more validity

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...
  - The usual advice: You don't have a methodological problem. Get a theory and we'll know what to measure.
  - Result of more concreteness: more reliability, no more validity
- Incomparability in Survey Responses ("DIF")

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...
  - The usual advice: You don't have a methodological problem. Get a theory and we'll know what to measure.
  - Result of more concreteness: more reliability, no more validity
- Incomparability in Survey Responses ("DIF")
  - Chinese report having more political efficacy than Americans

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...
  - The usual advice: You don't have a methodological problem. Get a theory and we'll know what to measure.
  - Result of more concreteness: more reliability, no more validity
- Incomparability in Survey Responses ("DIF")
  - Chinese report having more political efficacy than Americans
  - The most common measure of health — "How healthy are you? (Excellent, Good, Fair, Poor)"

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...
  - The usual advice: You don't have a methodological problem. Get a theory and we'll know what to measure.
  - Result of more concreteness: more reliability, no more validity
- Incomparability in Survey Responses ("DIF")
  - Chinese report having more political efficacy than Americans
  - The most common measure of health — "How healthy are you? (Excellent, Good, Fair, Poor)" — often correlates negatively with actual health (Amaryta Sen, 2002)

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...
  - The usual advice: You don't have a methodological problem. Get a theory and we'll know what to measure.
  - Result of more concreteness: more reliability, no more validity
- Incomparability in Survey Responses ("DIF")
  - Chinese report having more political efficacy than Americans
  - The most common measure of health — "How healthy are you? (Excellent, Good, Fair, Poor)" — often correlates negatively with actual health (Amaryta Sen, 2002)
  - Ethiopians and Danes report being equally healthy

Anchoring Vignettes Address Two Problems

- Big concepts we can define only by example
  - Illustrations: Freedom, efficacy, pornography, health, compactness, democracy, intelligence, ...
  - The usual advice: You don't have a methodological problem. Get a theory and we'll know what to measure.
  - Result of more concreteness: more reliability, no more validity
- Incomparability in Survey Responses ("DIF")
  - Chinese report having more political efficacy than Americans
  - The most common measure of health — "How healthy are you? (Excellent, Good, Fair, Poor)" — often correlates negatively with actual health (Amaryta Sen, 2002)
  - Ethiopians and Danes report being equally healthy
  - "Individuals understand the 'same' questions in vastly different ways." (Brady, 1985)

Anchoring Vignettes: Political Efficacy Through Voting

How much say [does ‘name’ / do you] have in getting the government to address issues that interest [him / her / you]?

(a) Unlimited say, (b) A lot of say, (c) Some say, (d) Little say, (e) No say at all

Anchoring Vignettes: Political Efficacy Through Voting

- "[Alison] lacks clean drinking water. She and her neighbors are supporting an opposition candidate in the forthcoming elections that has promised to address the issue. It appears that so many people in her area feel the same way that the opposition candidate will defeat the incumbent representative."

How much say [does 'name' / do you] have in getting the government to address issues that interest [him / her / you]?

(a) Unlimited say, (b) A lot of say, (c) Some say, (d) Little say, (e) No say at all

Anchoring Vignettes: Political Efficacy Through Voting

- "[Alison] lacks clean drinking water. She and her neighbors are supporting an opposition candidate in the forthcoming elections that has promised to address the issue. It appears that so many people in her area feel the same way that the opposition candidate will defeat the incumbent representative."
- "[Jane] lacks clean drinking water because the government is pursuing an industrial development plan. In the campaign for an upcoming election, an opposition party has promised to address the issue, but she feels it would be futile to vote for the opposition since the government is certain to win."

How much say [does 'name' / do you] have in getting the government to address issues that interest [him / her / you]?

(a) Unlimited say, (b) A lot of say, (c) Some say, (d) Little say, (e) No say at all

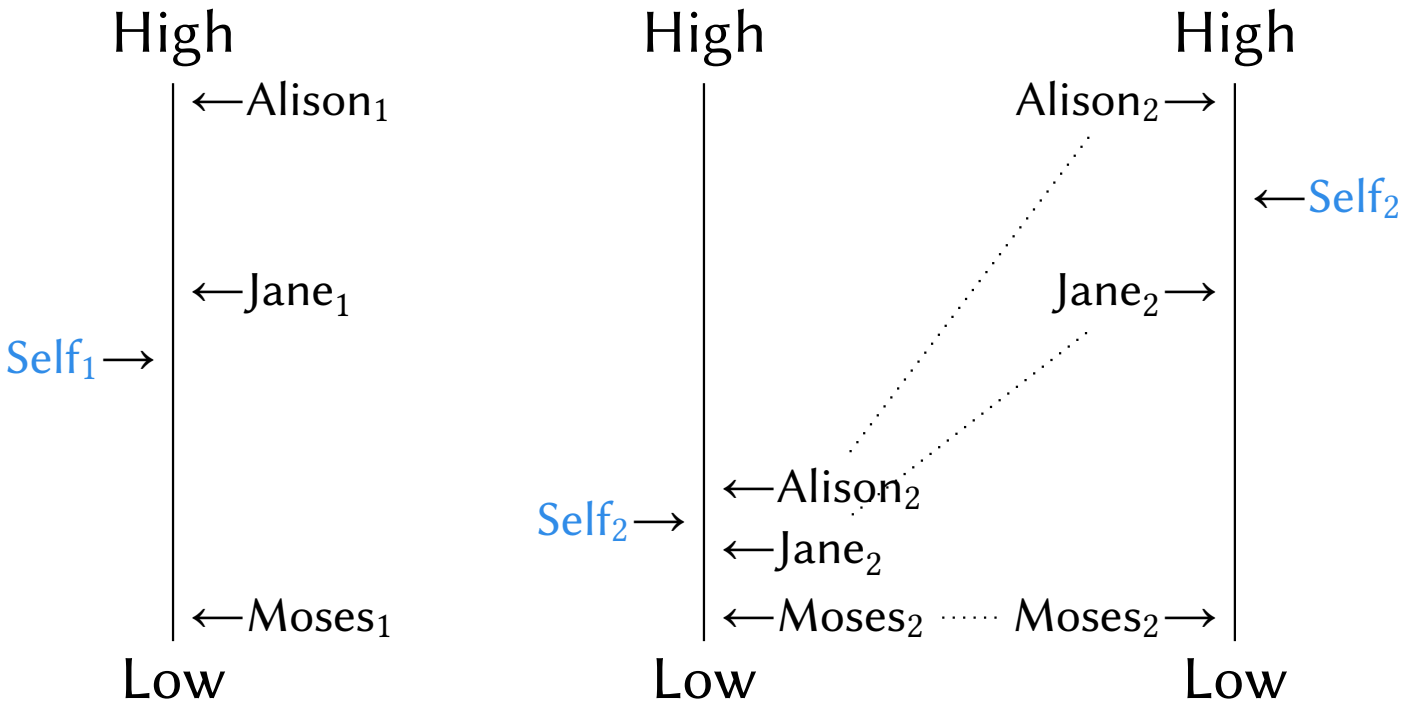
Anchoring Vignettes: Political Efficacy Through Voting

- "[Alison] lacks clean drinking water. She and her neighbors are supporting an opposition candidate in the forthcoming elections that has promised to address the issue. It appears that so many people in her area feel the same way that the opposition candidate will defeat the incumbent representative."
- "[Jane] lacks clean drinking water because the government is pursuing an industrial development plan. In the campaign for an upcoming election, an opposition party has promised to address the issue, but she feels it would be futile to vote for the opposition since the government is certain to win."
- "[Moses] lacks clean drinking water. He would like to change this, but he can't vote, and feels that no one in the government cares about this issue. So he suffers in silence, hoping something will be done in the future."

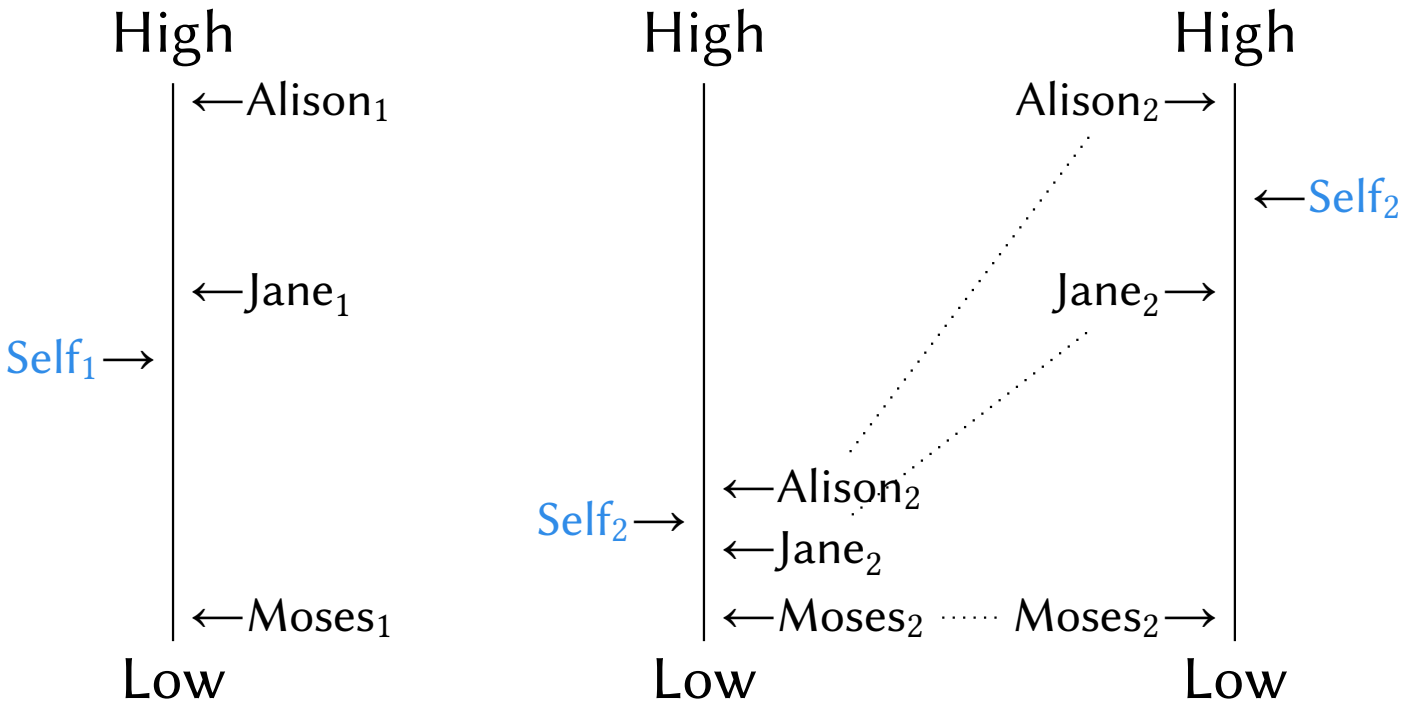
How much say [does 'name' / do you] have in getting the government to address issues that interest [him / her / you]?

(a) Unlimited say, (b) A lot of say, (c) Some say, (d) Little say, (e) No say at all

Who has More Political Efficacy: R_1 or R_2?

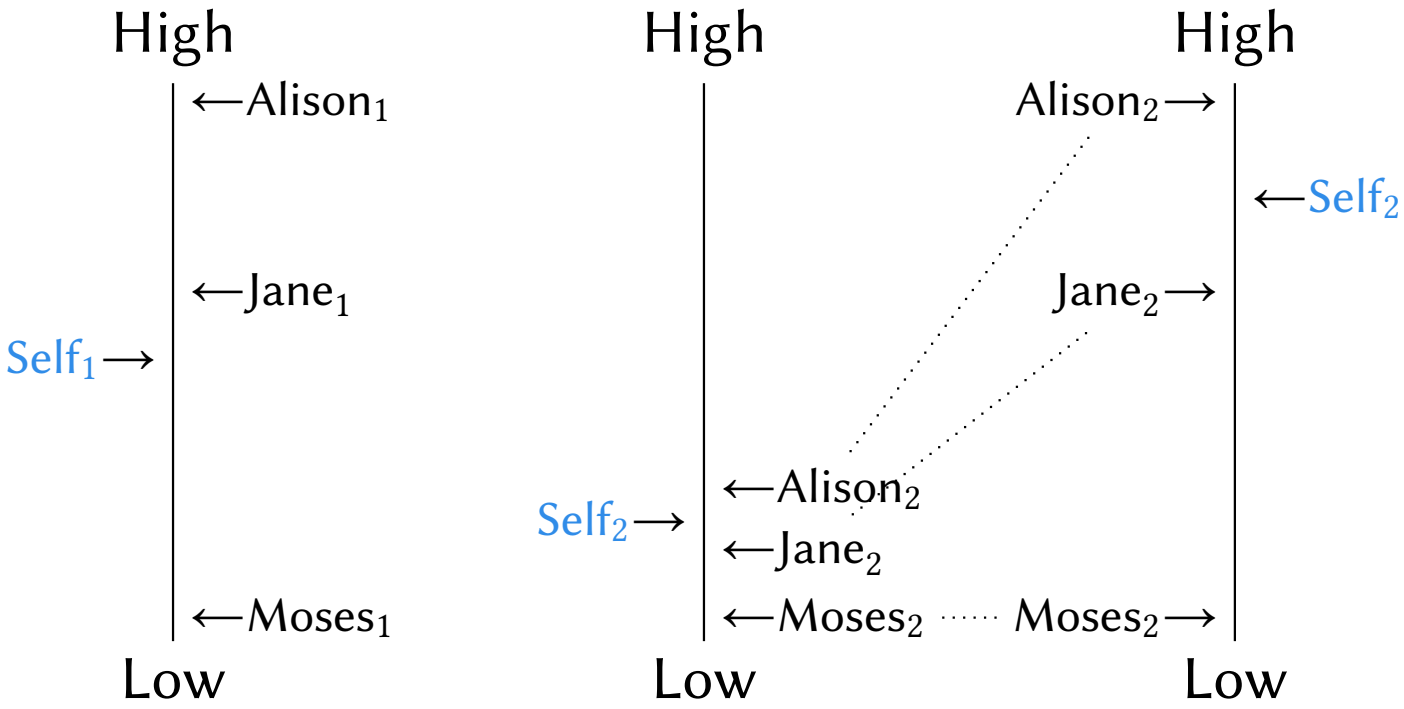


Who has More Political Efficacy: R_1 or R_2?



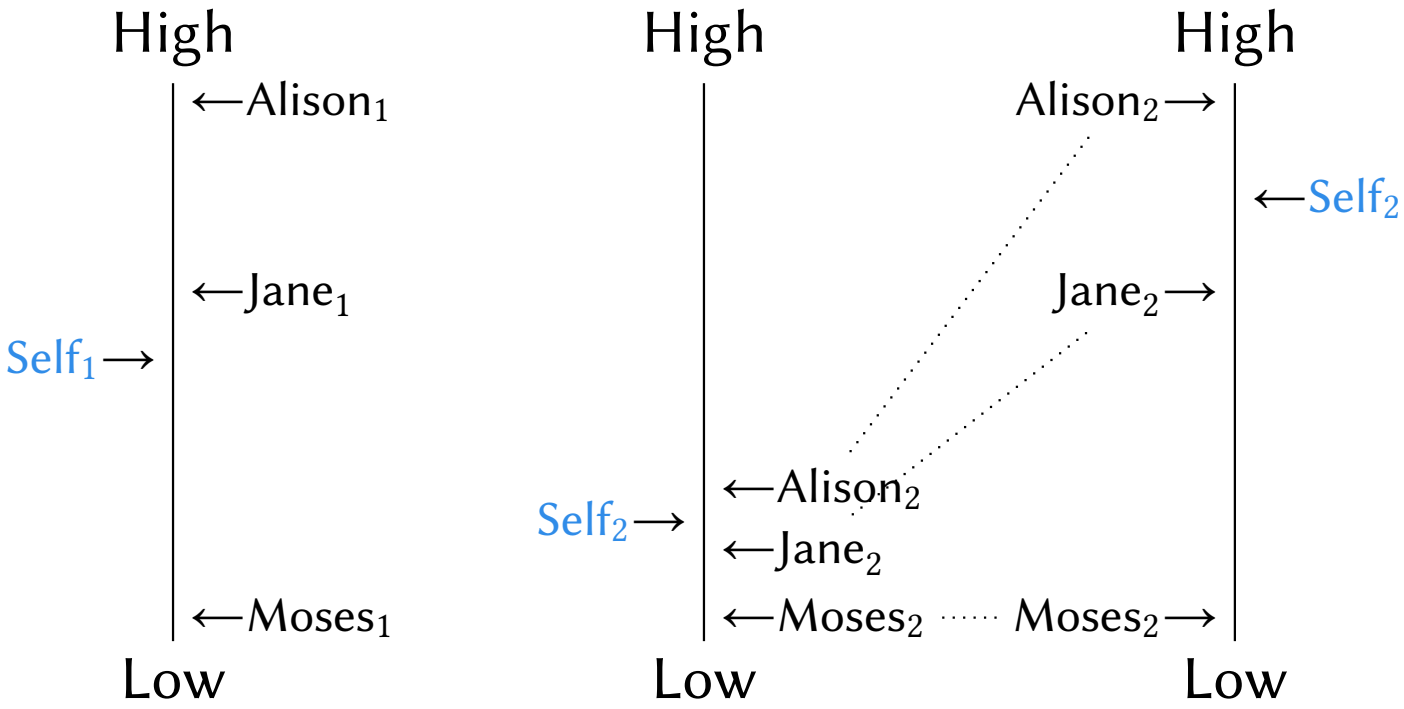
- Quiz: What does *varying vignette assessments* tell us?

Who has More Political Efficacy: R_1 or R_2?



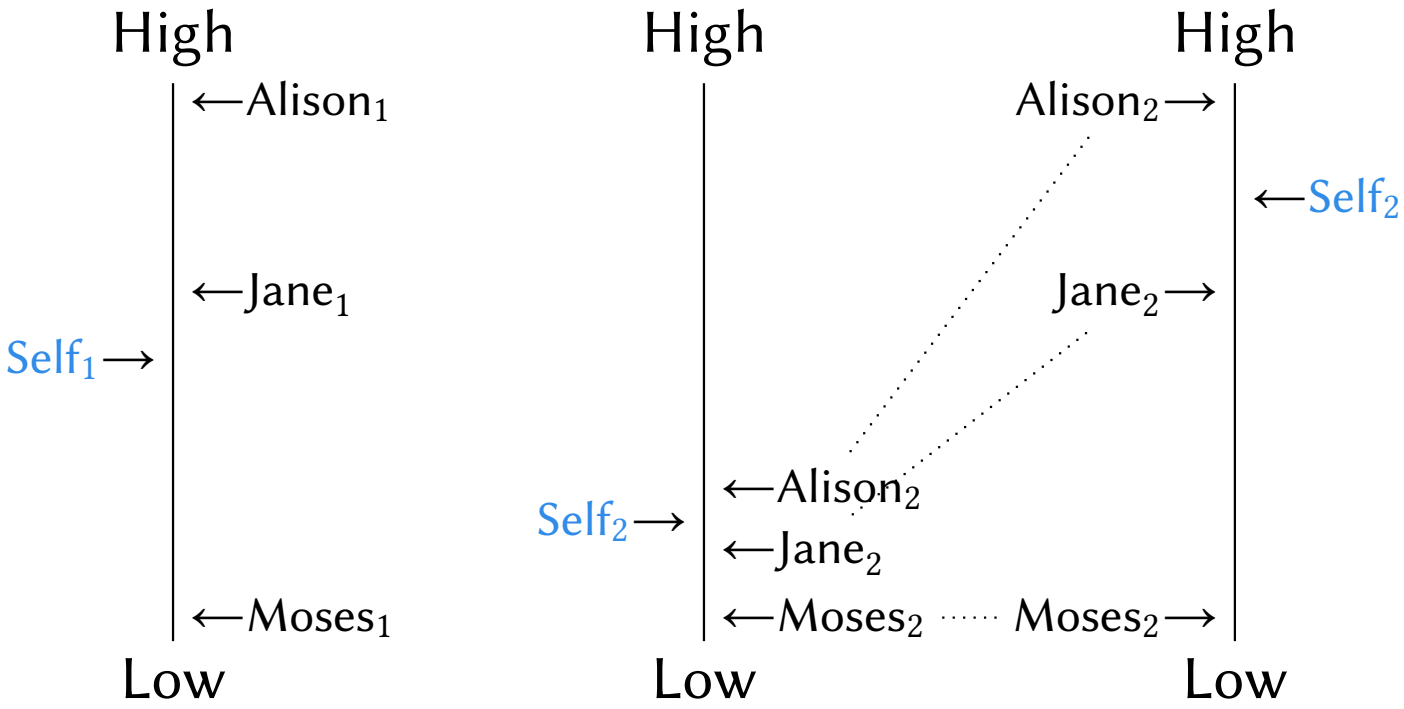
- Quiz: What does *varying vignette assessments* tell us? DIF

Who has More Political Efficacy: R_1 or R_2?



- Quiz: What does *varying vignette assessments* tell us? DIF
- Quiz: How do we know true vignette assessments are fixed over respondents?

Who has More Political Efficacy: R_1 or R_2?



- Quiz: What does *varying vignette assessments* tell us? DIF
- Quiz: How do we know true vignette assessments are fixed over respondents? We create the anchors (Alison, Jane, Moses)!

Formalizing A Nonparametric Methodology

Formalizing A Nonparametric Methodology

- Define self-assessment answers *relative* to vignettes answers.

Formalizing A Nonparametric Methodology

- Define self-assessment answers *relative* to vignettes answers.
- For respondents who rank vignettes, $z_{i1} < z_{i2} < \cdots < z_{iJ}$,

$$C_i = \begin{cases} 1 & \text{if } y_i < z_{i1} \\ 2 & \text{if } y_i = z_{i1} \\ 3 & \text{if } z_{i1} < y_i < z_{i2} \\ \vdots & \vdots \\ 2J+1 & \text{if } y_i > z_{iJ} \end{cases}$$

Formalizing A Nonparametric Methodology

- Define self-assessment answers *relative* to vignettes answers.
- For respondents who rank vignettes, $z_{i1} < z_{i2} < \cdots < z_{iJ}$,

$$C_i = \begin{cases} 1 & \text{if } y_i < z_{i1} \\ 2 & \text{if } y_i = z_{i1} \\ 3 & \text{if } z_{i1} < y_i < z_{i2} \\ \vdots & \vdots \\ 2J+1 & \text{if } y_i > z_{iJ} \end{cases}$$

- Apportion C equally among tied vignette categories

Formalizing A Nonparametric Methodology

- Define self-assessment answers *relative* to vignettes answers.
- For respondents who rank vignettes, $z_{i1} < z_{i2} < \cdots < z_{iJ}$,

$$C_i = \begin{cases} 1 & \text{if } y_i < z_{i1} \\ 2 & \text{if } y_i = z_{i1} \\ 3 & \text{if } z_{i1} < y_i < z_{i2} \\ \vdots & \vdots \\ 2J+1 & \text{if } y_i > z_{iJ} \end{cases}$$

- Apportion C equally among tied vignette categories
- (This is wrong, but simple; we will improve shortly)

Formalizing A Nonparametric Methodology

- Define self-assessment answers *relative* to vignettes answers.
- For respondents who rank vignettes, $z_{i1} < z_{i2} < \cdots < z_{iJ}$,

$$C_i = \begin{cases} 1 & \text{if } y_i < z_{i1} \\ 2 & \text{if } y_i = z_{i1} \\ 3 & \text{if } z_{i1} < y_i < z_{i2} \\ \vdots & \vdots \\ 2J+1 & \text{if } y_i > z_{iJ} \end{cases}$$

- Apportion C equally among tied vignette categories
- (This is wrong, but simple; we will improve shortly)
- Treat vignette ranking inconsistencies as ties

Formalizing A Nonparametric Methodology

- Define self-assessment answers *relative* to vignettes answers.
- For respondents who rank vignettes, $z_{i1} < z_{i2} < \cdots < z_{iJ}$,

$$C_i = \begin{cases} 1 & \text{if } y_i < z_{i1} \\ 2 & \text{if } y_i = z_{i1} \\ 3 & \text{if } z_{i1} < y_i < z_{i2} \\ \vdots & \vdots \\ 2J+1 & \text{if } y_i > z_{iJ} \end{cases}$$

- Apportion C equally among tied vignette categories
- (This is wrong, but simple; we will improve shortly)
- Treat vignette ranking inconsistencies as ties
- Requires vignettes and self-assessments asked of all

Formalizing A Nonparametric Methodology

- Define self-assessment answers *relative* to vignettes answers.
- For respondents who rank vignettes, $z_{i1} < z_{i2} < \cdots < z_{iJ}$,

$$C_i = \begin{cases} 1 & \text{if } y_i < z_{i1} \\ 2 & \text{if } y_i = z_{i1} \\ 3 & \text{if } z_{i1} < y_i < z_{i2} \\ \vdots & \vdots \\ 2J+1 & \text{if } y_i > z_{iJ} \end{cases}$$

- Apportion C equally among tied vignette categories
- (This is wrong, but simple; we will improve shortly)
- Treat vignette ranking inconsistencies as ties
- Requires vignettes and self-assessments asked of all
- (A parametric method to come doesn't)

Measurement Assumptions

Measurement Assumptions

- Response Consistency

Measurement Assumptions

- Response Consistency
  - Each respondent uses the self-assessment and vignette categories in approximately the same way across questions

Measurement Assumptions

- Response Consistency
  - Each respondent uses the self-assessment and vignette categories in approximately the same way across questions
  - I.e., DIF occurs across respondents, not across questions for any one respondent

Measurement Assumptions

- Response Consistency
  - Each respondent uses the self-assessment and vignette categories in approximately the same way across questions
  - I.e., DIF occurs across respondents, not across questions for any one respondent
- Vignette Equivalence

Measurement Assumptions

- Response Consistency
  - Each respondent uses the self-assessment and vignette categories in approximately the same way across questions
  - I.e., DIF occurs across respondents, not across questions for any one respondent
- Vignette Equivalence
  - The actual level for any vignette is the same for all respondents

Measurement Assumptions

- Response Consistency
  - Each respondent uses the self-assessment and vignette categories in approximately the same way across questions
  - I.e., DIF occurs across respondents, not across questions for any one respondent
- Vignette Equivalence
  - The actual level for any vignette is the same for all respondents
  - The quantity being estimated exists

Measurement Assumptions

- Response Consistency
  - Each respondent uses the self-assessment and vignette categories in approximately the same way across questions
  - I.e., DIF occurs across respondents, not across questions for any one respondent
- Vignette Equivalence
  - The actual level for any vignette is the same for all respondents
  - The quantity being estimated exists
  - The scale being tapped is perceived as unidimensional

Measurement Assumptions

- Response Consistency
  - Each respondent uses the self-assessment and vignette categories in approximately the same way across questions
  - I.e., DIF occurs across respondents, not across questions for any one respondent
- Vignette Equivalence
  - The actual level for any vignette is the same for all respondents
  - The quantity being estimated exists
  - The scale being tapped is perceived as unidimensional
- In other words: we allow response-category DIF but assume stem question equivalence

Comparing Political Efficacy in Mexico and China

Comparing Political Efficacy in Mexico and China

Mexico

Opposition leader Vicente Fox elected president. 71-year rule of PRI Party ends. Peaceful transition of power begins.

Comparing Political Efficacy in Mexico and China

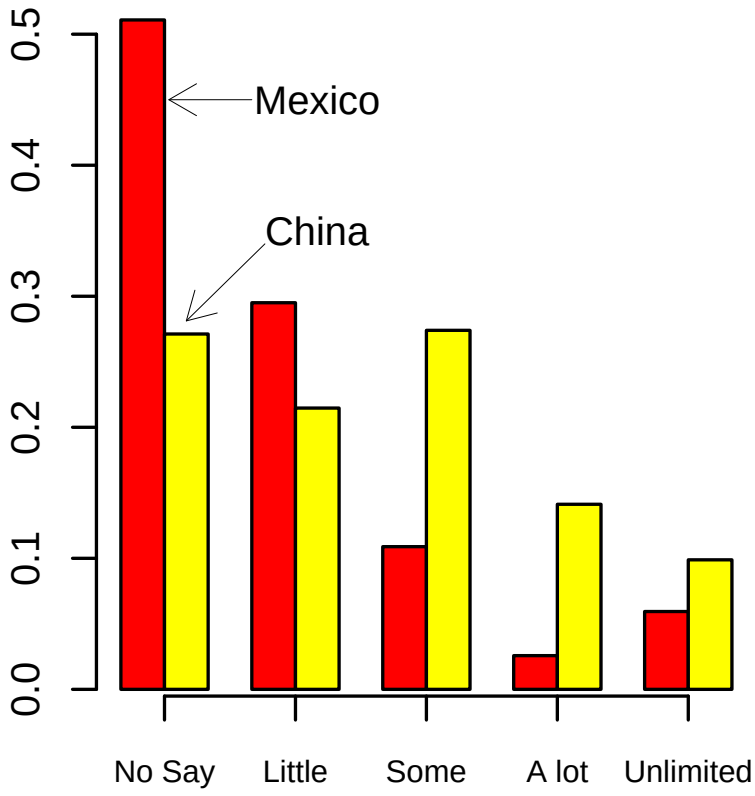
Mexico

Opposition leader Vicente Fox elected president. 71-year rule of PRI Party ends. Peaceful transition of power begins.

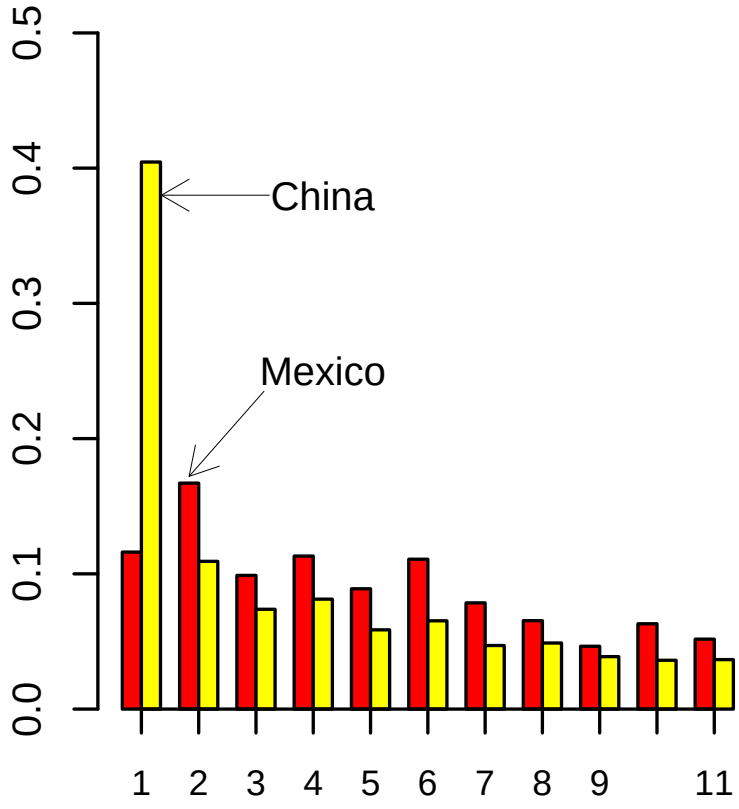
China

How much say do you have getting government to address issues that interest you?

Nonparametric Estimates of Political Efficacy



observed categorical self-assessments



C, our nonparametric DIF-corrected estimate of the same distribution

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C

Scalars:

$y < z_1 < z_2$	T	$\{1\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$
Ties:						

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$
Ties:						
$y < z_1 = z_2$	T					$\{1\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$
Ties:						
$y < z_1 = z_2$	T					$\{1\}$
$y = z_1 = z_2$		T		T		$\{2,3,4\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$
Ties:						
$y < z_1 = z_2$	T					$\{1\}$
$y = z_1 = z_2$		T		T		$\{2,3,4\}$
$z_1 = z_2 < y$					T	$\{5\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					{1}
$y = z_1 < z_2$		T				{2}
$z_1 < y < z_2$			T			{3}
$z_1 < y = z_2$				T		{4}
$z_1 < z_2 < y$					T	{5}
Ties:						
$y < z_1 = z_2$	T					{1}
$y = z_1 = z_2$		T		T		{2,3,4}
$z_1 = z_2 < y$					T	{5}
Inconsistencies:						

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$
Ties:						
$y < z_1 = z_2$	T					$\{1\}$
$y = z_1 = z_2$		T		T		$\{2,3,4\}$
$z_1 = z_2 < y$					T	$\{5\}$
Inconsistencies:						
$y < z_2 < z_1$	T					$\{1\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$
Ties:						
$y < z_1 = z_2$	T					$\{1\}$
$y = z_1 = z_2$		T		T		$\{2,3,4\}$
$z_1 = z_2 < y$					T	$\{5\}$
Inconsistencies:						
$y < z_2 < z_1$	T					$\{1\}$
$y = z_2 < z_1$	T			T		$\{1,2,3,4\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$
Ties:						
$y < z_1 = z_2$	T					$\{1\}$
$y = z_1 = z_2$		T		T		$\{2,3,4\}$
$z_1 = z_2 < y$					T	$\{5\}$
Inconsistencies:						
$y < z_2 < z_1$	T					$\{1\}$
$y = z_2 < z_1$	T			T		$\{1,2,3,4\}$
$z_2 < y < z_1$	T				T	$\{1,2,3,4,5\}$

Ties and Inconsistencies Produce *Ranges*

Survey	1:	2:	3:	4:	5:	
Responses	$y < z_1$	$y = z_1$	$z_1 < y < z_2$	$y = z_2$	$y > z_2$	C
Scalars:						
$y < z_1 < z_2$	T					$\{1\}$
$y = z_1 < z_2$		T				$\{2\}$
$z_1 < y < z_2$			T			$\{3\}$
$z_1 < y = z_2$				T		$\{4\}$
$z_1 < z_2 < y$					T	$\{5\}$
Ties:						
$y < z_1 = z_2$	T					$\{1\}$
$y = z_1 = z_2$		T		T		$\{2,3,4\}$
$z_1 = z_2 < y$					T	$\{5\}$
Inconsistencies:						
$y < z_2 < z_1$	T					$\{1\}$
$y = z_2 < z_1$	T			T		$\{1,2,3,4\}$
$z_2 < y < z_1$	T				T	$\{1,2,3,4,5\}$
$z_2 < y = z_1$		T			T	$\{2,3,4,5\}$
$z_2 < z_1 < y$					T	$\{5\}$

A Likelihood for Scalar or Vector Responses

A Likelihood for Scalar or Vector Responses

Parametric likelihood used in a nonparametric methodology

A Likelihood for Scalar or Vector Responses

Parametric likelihood used in a nonparametric methodology

- Define an unobserved variable: $Y_i \sim \text{Normal}(x_i\beta, 1)$

A Likelihood for Scalar or Vector Responses

Parametric likelihood used in a nonparametric methodology

- Define an unobserved variable: $Y_i \sim \text{Normal}(x_i\beta, 1)$
- With observation mechanism, for scalar C, same as ordered probit: $C_i = c \quad \text{if } \tau_{c-1} \leq Y_i < \tau_c$

A Likelihood for Scalar or Vector Responses

Parametric likelihood used in a nonparametric methodology

- Define an unobserved variable: $Y_i \sim \text{Normal}(x_i\beta, 1)$
- With observation mechanism, for scalar C, same as ordered probit: $C_i = c \quad \text{if } \tau_{c-1} \leq Y_i < \tau_c$
- Probability of observing category c, for $X = x_0$:

$$\Pr(C = c|x_0) = \int_{\tau_{c-1}}^{\tau_c} \text{Normal}(y|x_0\beta, 1)\, dy$$

A Likelihood for Scalar or Vector Responses

Parametric likelihood used in a nonparametric methodology

- Define an unobserved variable: $Y_i \sim \text{Normal}(x_i\beta, 1)$
- With observation mechanism, for scalar C, same as ordered probit: $C_i = c \quad \text{if } \tau_{c-1} \leq Y_i < \tau_c$
- Probability of observing category c, for $X = x_0$:

$$\Pr(C = c|x_0) = \int_{\tau_{c-1}}^{\tau_c} \text{Normal}(y|x_0\beta, 1)\, dy$$

- Observation mechanism for vector valued C:

$$C_i = c \qquad \text{if } \tau_{\min(c)-1} \leq Y_i < \tau_{\max(c)}$$

A Likelihood for Scalar or Vector Responses

Parametric likelihood used in a nonparametric methodology

- Define an unobserved variable: $Y_i \sim \text{Normal}(x_i\beta, 1)$
- With observation mechanism, for scalar C, same as ordered probit: $C_i = c \quad \text{if } \tau_{c-1} \leq Y_i < \tau_c$
- Probability of observing category c, for $X = x_0$:

$$\Pr(C = c|x_0) = \int_{\tau_{c-1}}^{\tau_c} \text{Normal}(y|x_0\beta, 1)\, dy$$

- Observation mechanism for vector valued C:

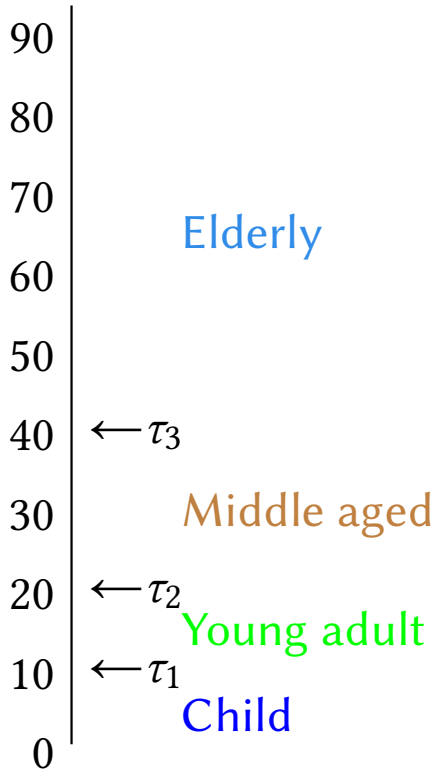
$$C_i = c \qquad \text{if } \tau_{\min(c)-1} \leq Y_i < \tau_{\max(c)}$$

- Quiz: Do you know how to form the likelihood function?

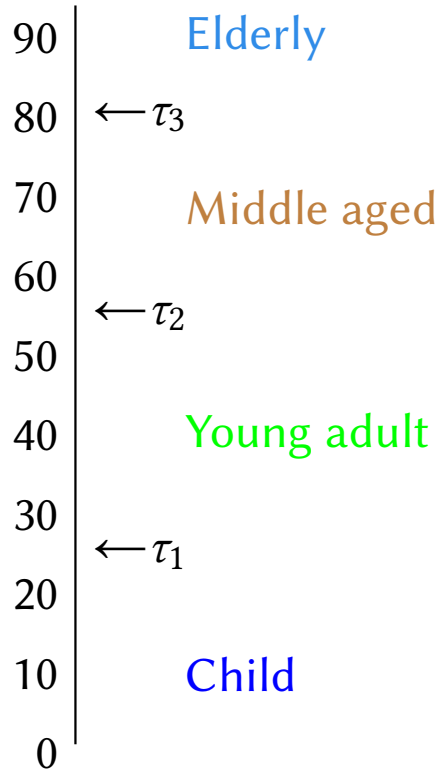
Categorizing Years of Age

Categorizing Years of Age

Respondent 1

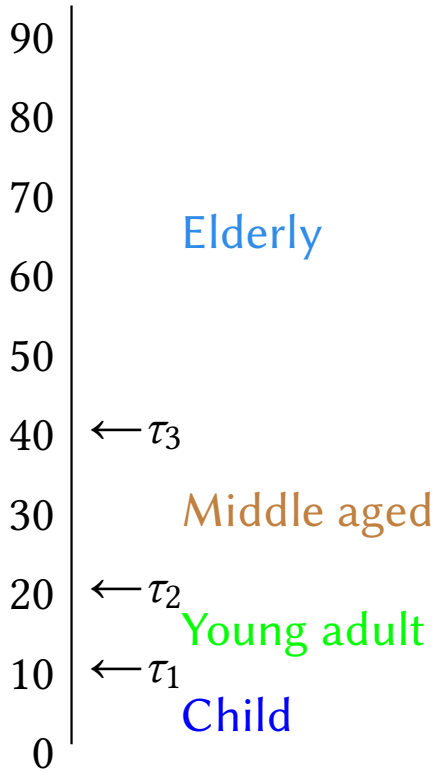


Respondent 2

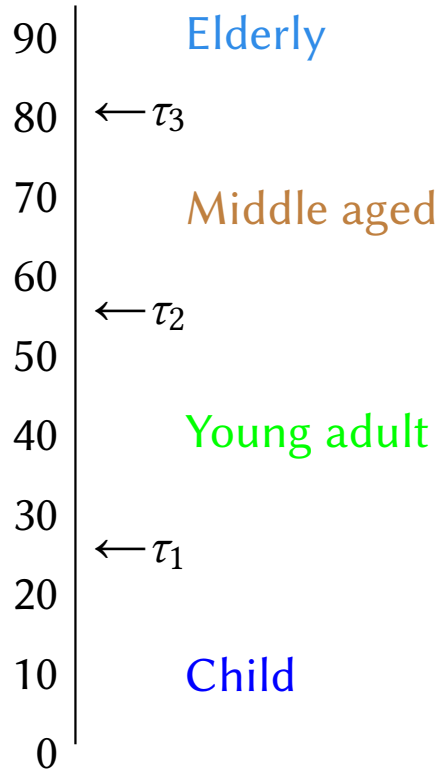


Categorizing Years of Age

Respondent 1

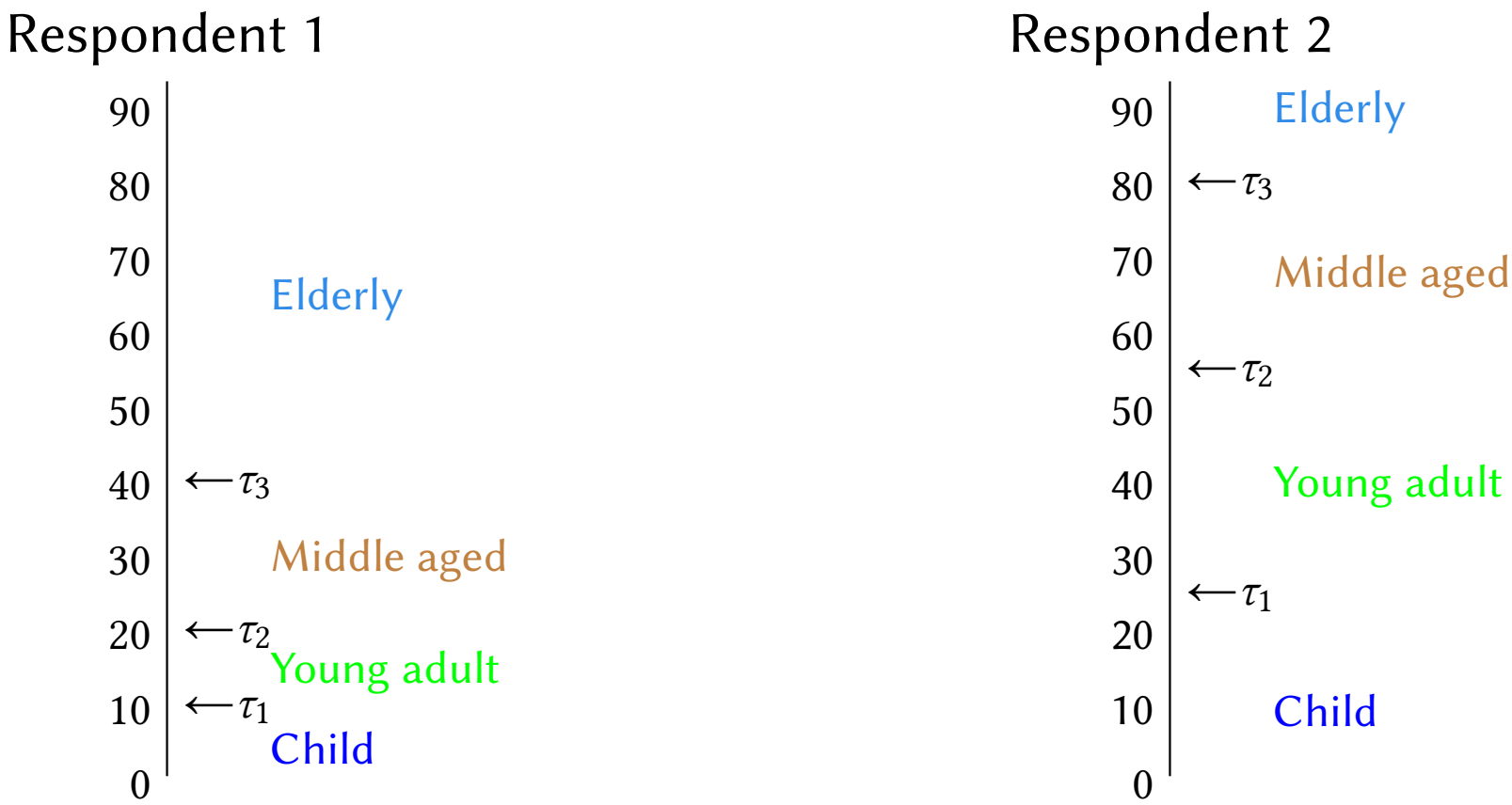


Respondent 2



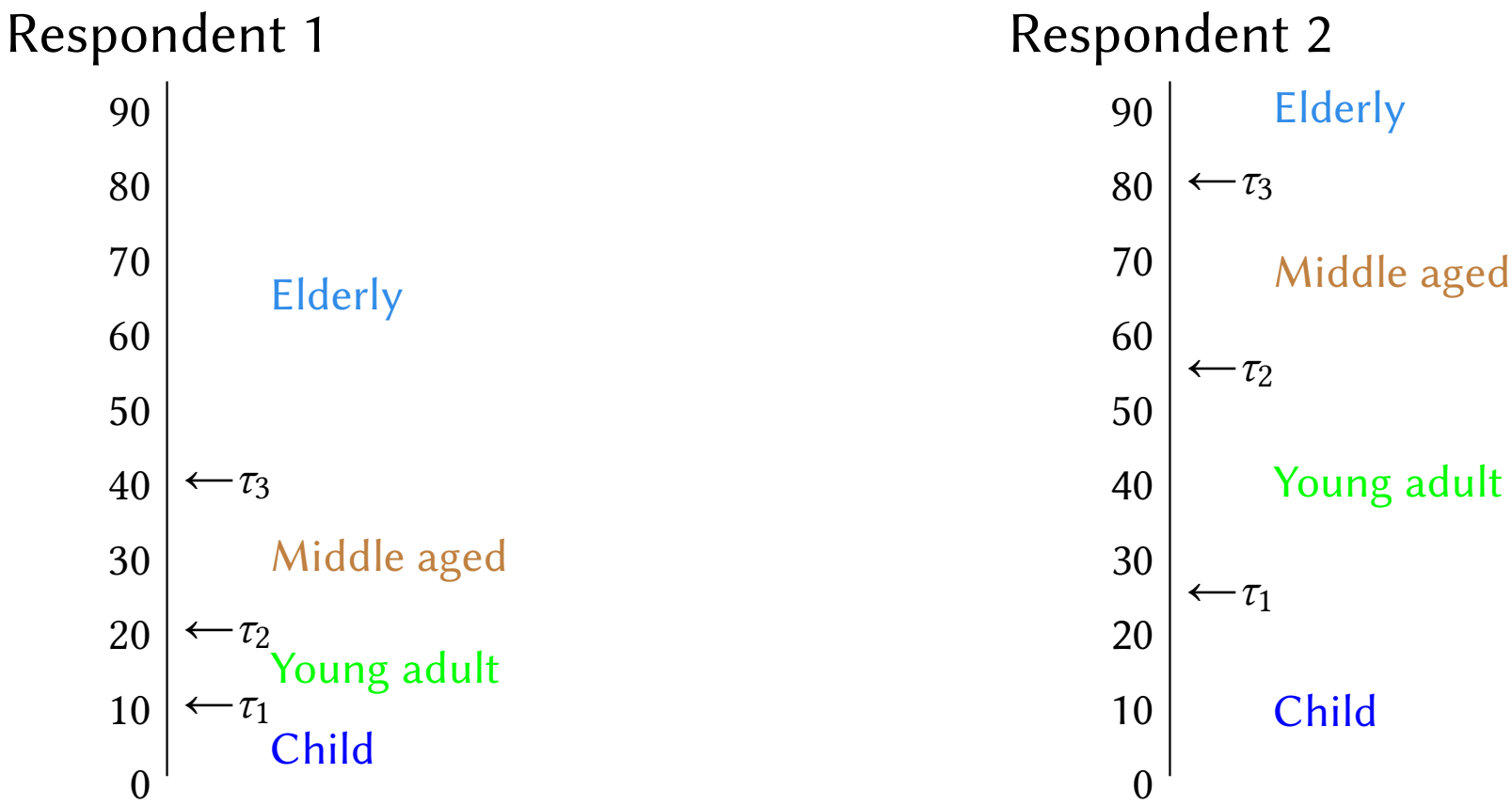
- If thresholds vary: categorical answers are meaningless

Categorizing Years of Age



- If thresholds vary: categorical answers are meaningless
- Quiz: In Ordered Logit, can we subscript τ_i by person?

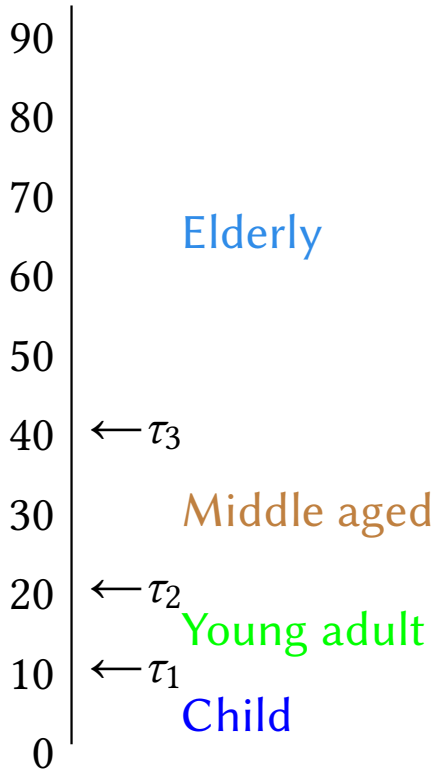
Categorizing Years of Age



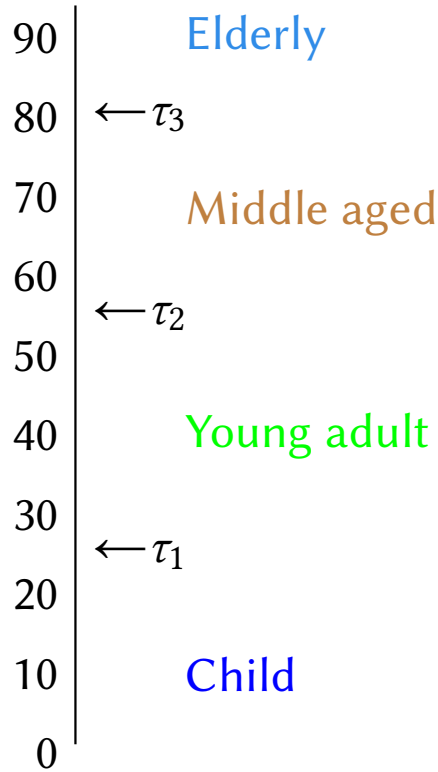
- If thresholds vary: categorical answers are meaningless
- Quiz: In Ordered Logit, can we subscript τ_i by person?
- Quiz: What information would identify the τ_i's?

Categorizing Years of Age

Respondent 1



Respondent 2



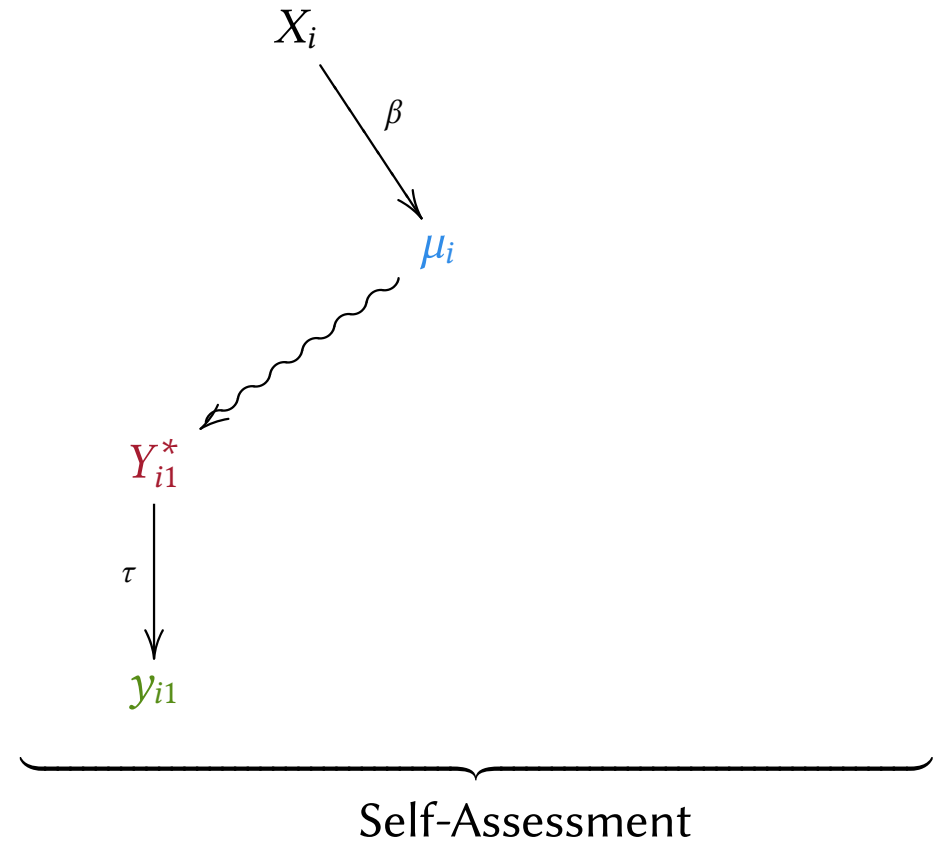
- If thresholds vary: categorical answers are meaningless
- Quiz: In Ordered Logit, can we subscript τ_i by person?
- Quiz: What information would identify the τ_i's?
- A parametric model works by estimating the thresholds

Statistical Model: Visual Summary

Actual:

Perceived:

Reported:



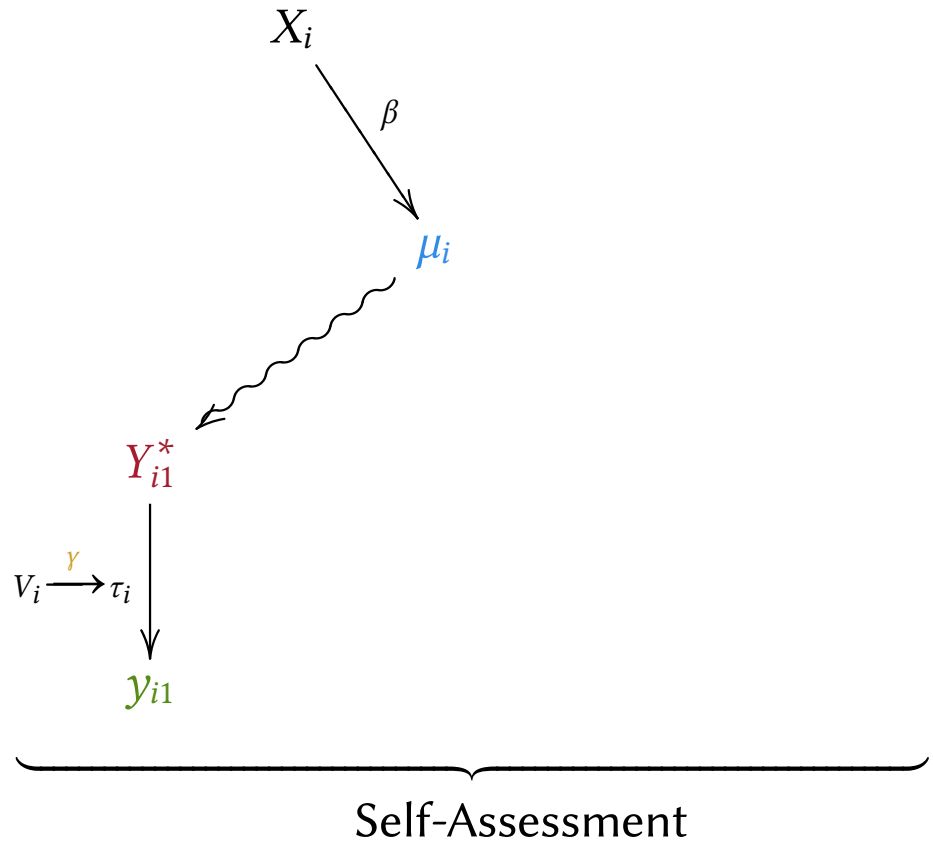
- An ordinal probit model,

Statistical Model: Visual Summary

Actual:

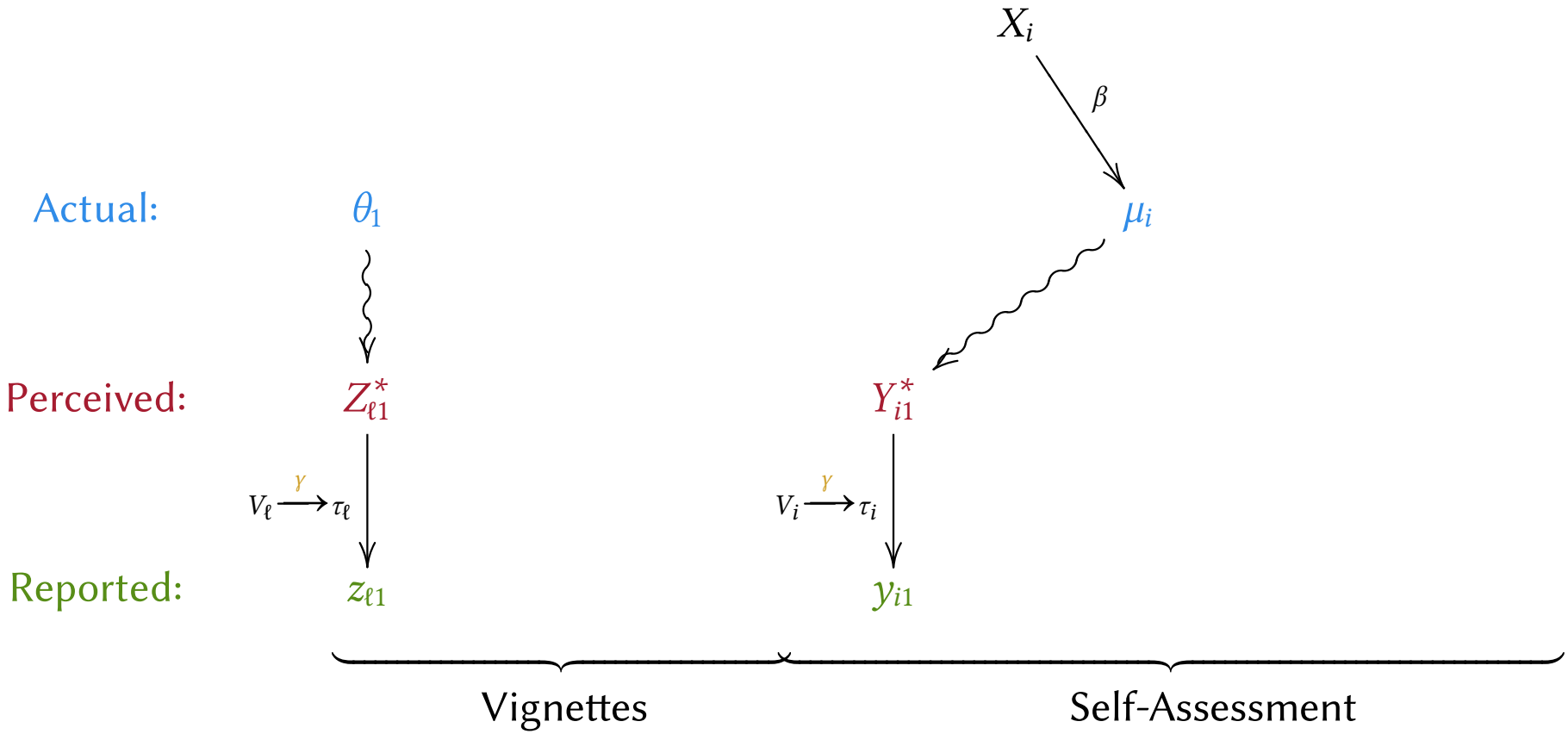
Perceived:

Reported:



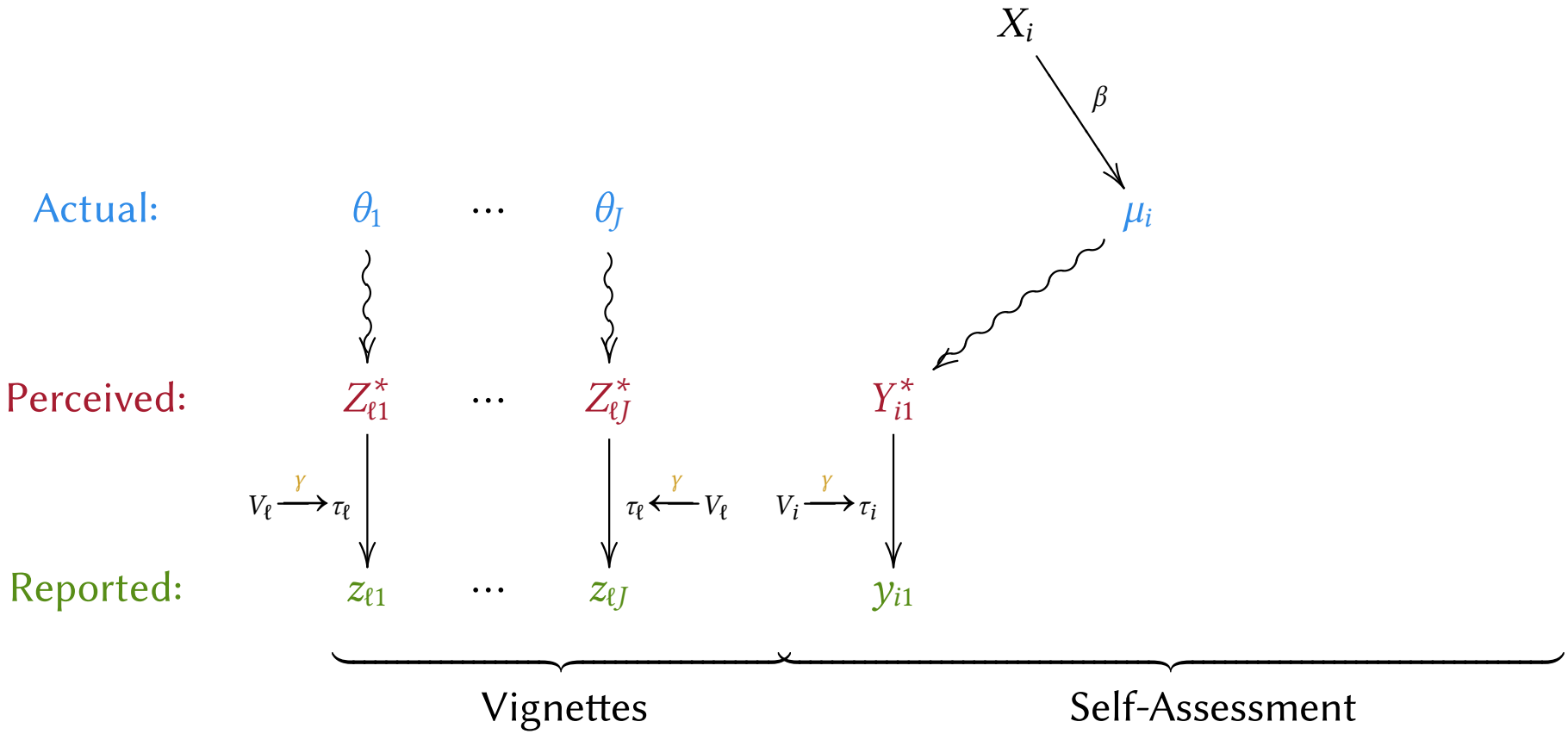
- An ordinal probit model,
- with varying thresholds,

Statistical Model: Visual Summary



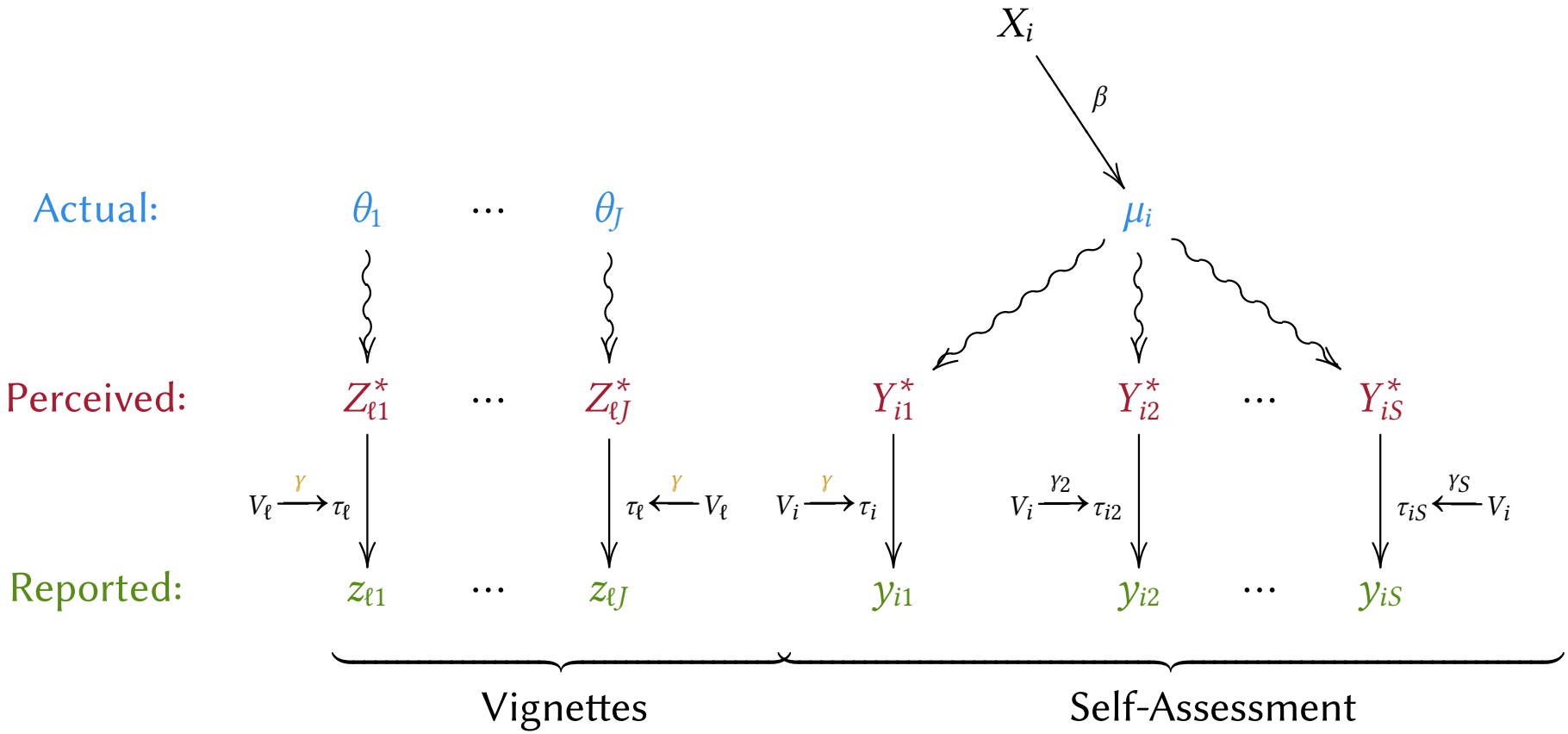
- An ordinal probit model,
- with varying thresholds,
- a vignette for identification (with the same γ),

Statistical Model: Visual Summary



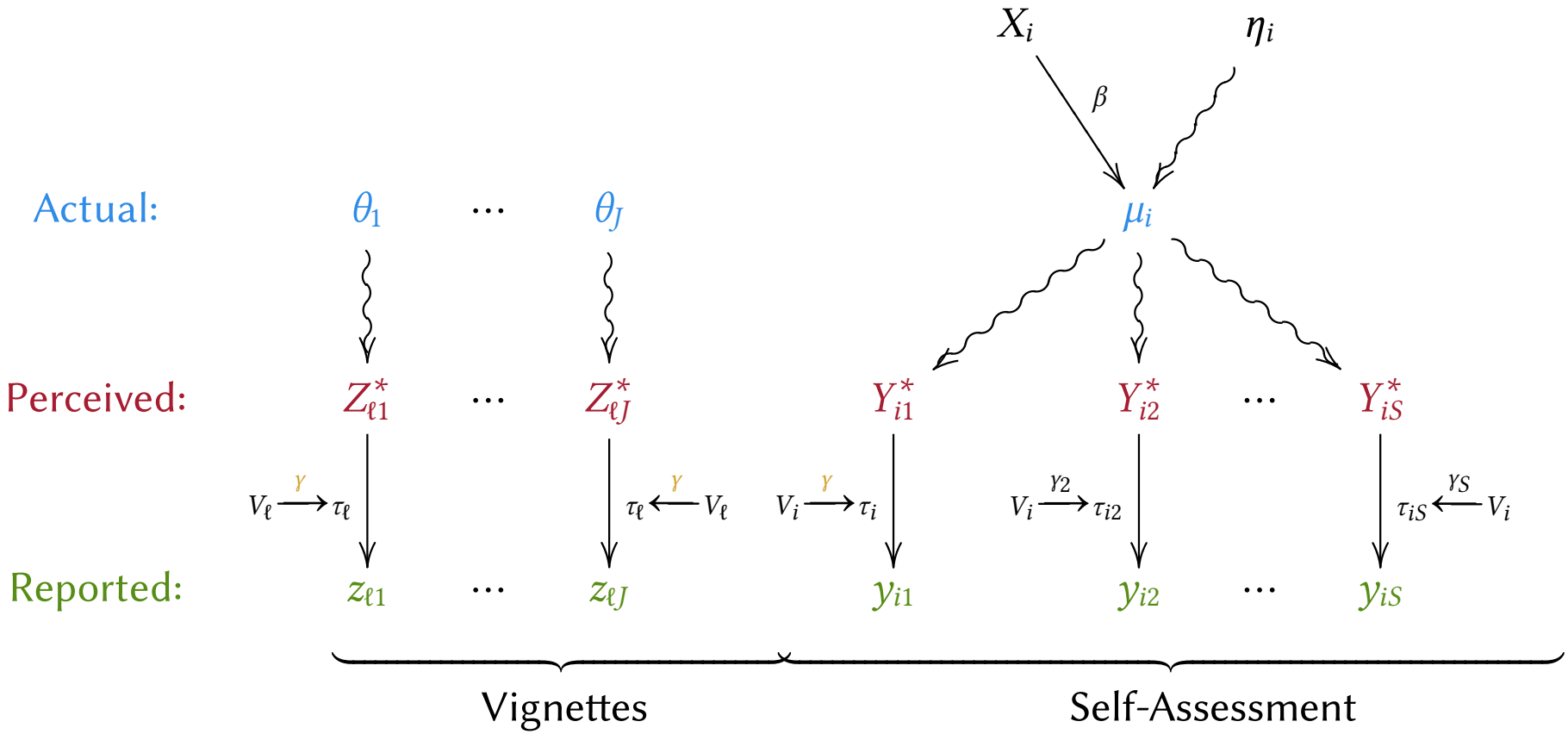
- An ordinal probit model,
- with varying thresholds,
- a vignette for identification (with the same γ),
- more vignettes for better discrimination,

Statistical Model: Visual Summary



- An ordinal probit model,
- with varying thresholds,
- a vignette for identification (with the same γ),
- more vignettes for better discrimination,
- optional multiple self-assessment questions for more efficiency,

Statistical Model: Visual Summary



- An ordinal probit model,
- with varying thresholds,
- a vignette for identification (with the same γ),
- more vignettes for better discrimination,
- optional multiple self-assessment questions for more efficiency,
- and an optional random effect.

Self-Assessment Model Component ($i = 1, \ldots, n$)

Self-Assessment Model Component ($i = 1, \ldots, n$)

Vignette Model Component ($\ell = 1, \ldots, N$)

Self-Assessment Model Component ($i = 1, \ldots, n$)

- Actual: $\mu_i = X_i\beta + \eta_i,$ with random effect $\eta_i \sim N(0, \omega^2)$

Vignette Model Component ($\ell = 1, \ldots, N$)

Self-Assessment Model Component ($i = 1, \ldots, n$)

- Actual: $\mu_i = X_i\beta + \eta_i$, with random effect $\eta_i \sim N(0, \omega^2)$
- Perceived: $Y_{i1}^* \sim N(\mu_i, 1) \quad \ldots \quad Y_{iS}^* \sim N(\mu_i, 1)$

Vignette Model Component ($\ell = 1, \ldots, N$)

Self-Assessment Model Component ($i = 1, \ldots, n$)

- Actual: $\mu_i = X_i\beta + \eta_i$, with random effect $\eta_i \sim N(0, \omega^2)$
- Perceived: $Y^*_{i1} \sim N(\mu_i, 1) \quad \ldots \quad Y^*_{iS} \sim N(\mu_i, 1)$
- Reported (observation mechanism):

$$y_{i1} = k \quad \text{if } \tau^{k-1}_{i1} \leq Y^*_{i1} < \tau^k_{i1}$$
$$\vdots$$
$$y_{iS} = k \quad \text{if } \tau^{k-1}_{is} \leq Y^*_{is} < \tau^k_{is}$$

Vignette Model Component ($\ell = 1, \ldots, N$)

Self-Assessment Model Component ($i = 1, \ldots, n$)

- Actual: $\mu_i = X_i\beta + \eta_i$, with random effect $\eta_i \sim N(0, \omega^2)$
- Perceived: $Y^*_{i1} \sim N(\mu_i, 1) \quad \ldots \quad Y^*_{iS} \sim N(\mu_i, 1)$
- Reported (observation mechanism):

$$y_{i1} = k \quad \text{if } \tau^{k-1}_{i1} \leq Y^*_{i1} < \tau^k_{i1}$$
$$\vdots$$
$$y_{iS} = k \quad \text{if } \tau^{k-1}_{is} \leq Y^*_{is} < \tau^k_{is}$$

- where $\tau^1_{is} = \gamma_1 V_i, \quad \tau^k_{is} = \tau^{k-1}_{is} + e^{\gamma_k V_i}, \quad (k = 2, \ldots, K_s)$

Vignette Model Component ($\ell = 1, \ldots, N$)

Self-Assessment Model Component ($i = 1, \ldots, n$)

- Actual: $\mu_i = X_i\beta + \eta_i$, with random effect $\eta_i \sim N(0, \omega^2)$
- Perceived: $Y^*_{i1} \sim N(\mu_i, 1) \quad \ldots \quad Y^*_{iS} \sim N(\mu_i, 1)$
- Reported (observation mechanism):

$$y_{i1} = k \quad \text{if } \tau^{k-1}_{i1} \leq Y^*_{i1} < \tau^k_{i1}$$
$$\vdots$$
$$y_{iS} = k \quad \text{if } \tau^{k-1}_{is} \leq Y^*_{is} < \tau^k_{is}$$

- where $\tau^1_{is} = \gamma_1 V_i, \quad \tau^k_{is} = \tau^{k-1}_{is} + e^{\gamma_k V_i}, \quad (k = 2, \ldots, K_s)$

Vignette Model Component ($\ell = 1, \ldots, N$)

- Actual: $\theta_1, \ldots, \theta_J$

Self-Assessment Model Component ($i = 1, \ldots, n$)

- Actual: $\mu_i = X_i\beta + \eta_i$, with random effect $\eta_i \sim N(0, \omega^2)$
- Perceived: $Y_{i1}^* \sim N(\mu_i, 1) \quad \ldots \quad Y_{iS}^* \sim N(\mu_i, 1)$
- Reported (observation mechanism):

$$y_{i1} = k \quad \text{if } \tau_{i1}^{k-1} \leq Y_{i1}^* < \tau_{i1}^k$$
$$\vdots$$
$$y_{iS} = k \quad \text{if } \tau_{is}^{k-1} \leq Y_{is}^* < \tau_{is}^k$$

- where $\tau_{is}^1 = \gamma_1 V_i, \quad \tau_{is}^k = \tau_{is}^{k-1} + e^{\gamma_k} V_i, \quad (k = 2, \ldots, K_s)$

Vignette Model Component ($\ell = 1, \ldots, N$)

- Actual: $\theta_1, \ldots, \theta_J$
- Perceived: $Z_{\ell 1}^* \sim N(\theta_1, \sigma^2) \quad \ldots \quad Z_{\ell J}^* \sim N(\theta_J, \sigma^2)$

Self-Assessment Model Component ($i = 1, \ldots, n$)

- Actual: $\mu_i = X_i\beta + \eta_i$, with random effect $\eta_i \sim N(0, \omega^2)$
- Perceived: $Y^*_{i1} \sim N(\mu_i, 1) \quad \ldots \quad Y^*_{iS} \sim N(\mu_i, 1)$
- Reported (observation mechanism):

$$y_{i1} = k \quad \text{if } \tau^{k-1}_{i1} \leq Y^*_{i1} < \tau^k_{i1}$$
$$\vdots$$
$$y_{iS} = k \quad \text{if } \tau^{k-1}_{is} \leq Y^*_{is} < \tau^k_{is}$$

- where $\tau^1_{is} = \gamma_1 V_i, \quad \tau^k_{is} = \tau^{k-1}_{is} + e^{\gamma_k V_i}, \quad (k = 2, \ldots, K_s)$

Vignette Model Component ($\ell = 1, \ldots, N$)

- Actual: $\theta_1, \ldots, \theta_J$
- Perceived: $Z^*_{\ell 1} \sim N(\theta_1, \sigma^2) \quad \ldots \quad Z^*_{\ell J} \sim N(\theta_J, \sigma^2)$
- Reported: $z_{\ell j} = k \quad \text{if } \tau^{k-1}_{\ell 1} \leq Z^*_{\ell j} < \tau^k_{\ell 1},$

Self-Assessment Model Component ($i = 1, \ldots, n$)

- Actual: $\mu_i = X_i\beta + \eta_i$, with random effect $\eta_i \sim N(0, \omega^2)$
- Perceived: $Y^*_{i1} \sim N(\mu_i, 1) \quad \ldots \quad Y^*_{iS} \sim N(\mu_i, 1)$
- Reported (observation mechanism):

$$y_{i1} = k \quad \text{if } \tau^{k-1}_{i1} \le Y^*_{i1} < \tau^k_{i1}$$
$$\vdots$$
$$y_{iS} = k \quad \text{if } \tau^{k-1}_{is} \le Y^*_{is} < \tau^k_{is}$$

- where $\tau^1_{is} = \gamma_1 V_i, \quad \tau^k_{is} = \tau^{k-1}_{is} + e^{\gamma_k V_i}, \quad (k = 2, \ldots, K_s)$

Vignette Model Component ($\ell = 1, \ldots, N$)

- Actual: $\theta_1, \ldots, \theta_J$
- Perceived: $Z^*_{\ell 1} \sim N(\theta_1, \sigma^2) \quad \ldots \quad Z^*_{\ell J} \sim N(\theta_J, \sigma^2)$
- Reported: $z_{\ell j} = k \quad \text{if } \tau^{k-1}_{\ell 1} \le Z^*_{\ell j} < \tau^k_{\ell 1},$
- where $\tau^1_{\ell s} = \gamma_1 V_\ell, \quad \tau^k_{\ell s} = \tau^{k-1}_{\ell s} + e^{\gamma_k V_\ell} \quad (k = 2, \ldots, K_s)$

Self-Assessments v. Medical Tests

Self-Assessment: In the last 30 days, how much difficulty did [you/name] have in seeing and recognizing a person you know across the road (i.e. from a distance of about 20 meters)?

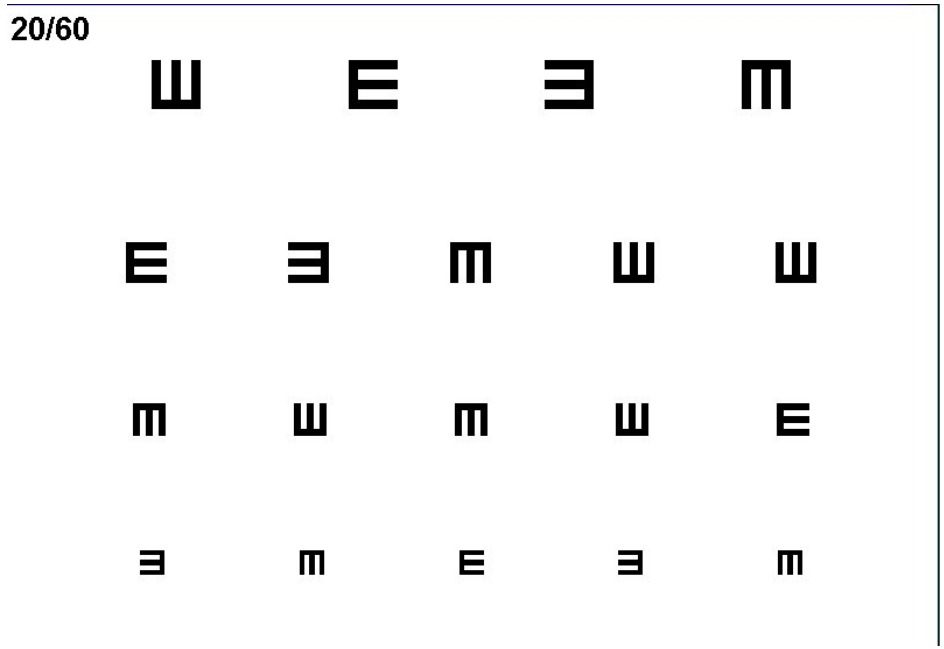
Self-Assessments v. Medical Tests

Self-Assessment: In the last 30 days, how much difficulty did [you/name] have in seeing and recognizing a person you know across the road (i.e. from a distance of about 20 meters)? (A) none, (B) mild, (C) moderate, (D) severe, (E) extreme/cannot do

Self-Assessments v. Medical Tests

Self-Assessment: In the last 30 days, how much difficulty did [you/name] have in seeing and recognizing a person you know across the road (i.e. from a distance of about 20 meters)? (A) none, (B) mild, (C) moderate, (D) severe, (E) extreme/cannot do

The Snellen Eye Chart Test:



Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

- Medical test: Slovakians see much better than the Chinese

Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

- Medical test: Slovakians see much better than the Chinese
- Ordinal probit: no difference

Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

- Medical test: Slovakians see much better than the Chinese
- Ordinal probit: no difference
- Chopit: same result as the medical test (on a different scale)

Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

- Medical test: Slovakians see much better than the Chinese
- Ordinal probit: no difference
- Chopit: same result as the medical test (on a different scale)
- (Better to calculate QOIs! The articles explain how.)

Fixing DIF in China and Mexico

Fixing DIF in China and Mexico

		Ordinal Probit		Chopit	
Eqn.	Variable	Coeff.	(s.e.)	Coeff.	(s.e.)
μ	China	.670	(.081)	−.362	(.090)
	age	.004	(.003)	.006	(.003)
	male	.087	(.076)	.113	(.081)
	education	.020	(.008)	.019	(.008)
Vignettes	θ_1			1.393	(.190)
	θ_2			1.304	(.190)
	θ_3			.953	(.189)
	θ_4			.902	(.188)
	θ_5			.729	(.188)
	$\ln \sigma$			−.238	(.042)

Fixing DIF in China and Mexico

		Ordinal Probit		Chopit	
Eqn.	Variable	Coeff.	(s.e.)	Coeff.	(s.e.)
μ	China	.670	(.081)	−.362	(.090)
	age	.004	(.003)	.006	(.003)
	male	.087	(.076)	.113	(.081)
	education	.020	(.008)	.019	(.008)
Vignettes	θ_1			1.393	(.190)
	θ_2			1.304	(.190)
	θ_3			.953	(.189)
	θ_4			.902	(.188)
	θ_5			.729	(.188)
	$\ln \sigma$			−.238	(.042)

- Ordered Probit: Chinese people have *more* political efficacy than Mexicans

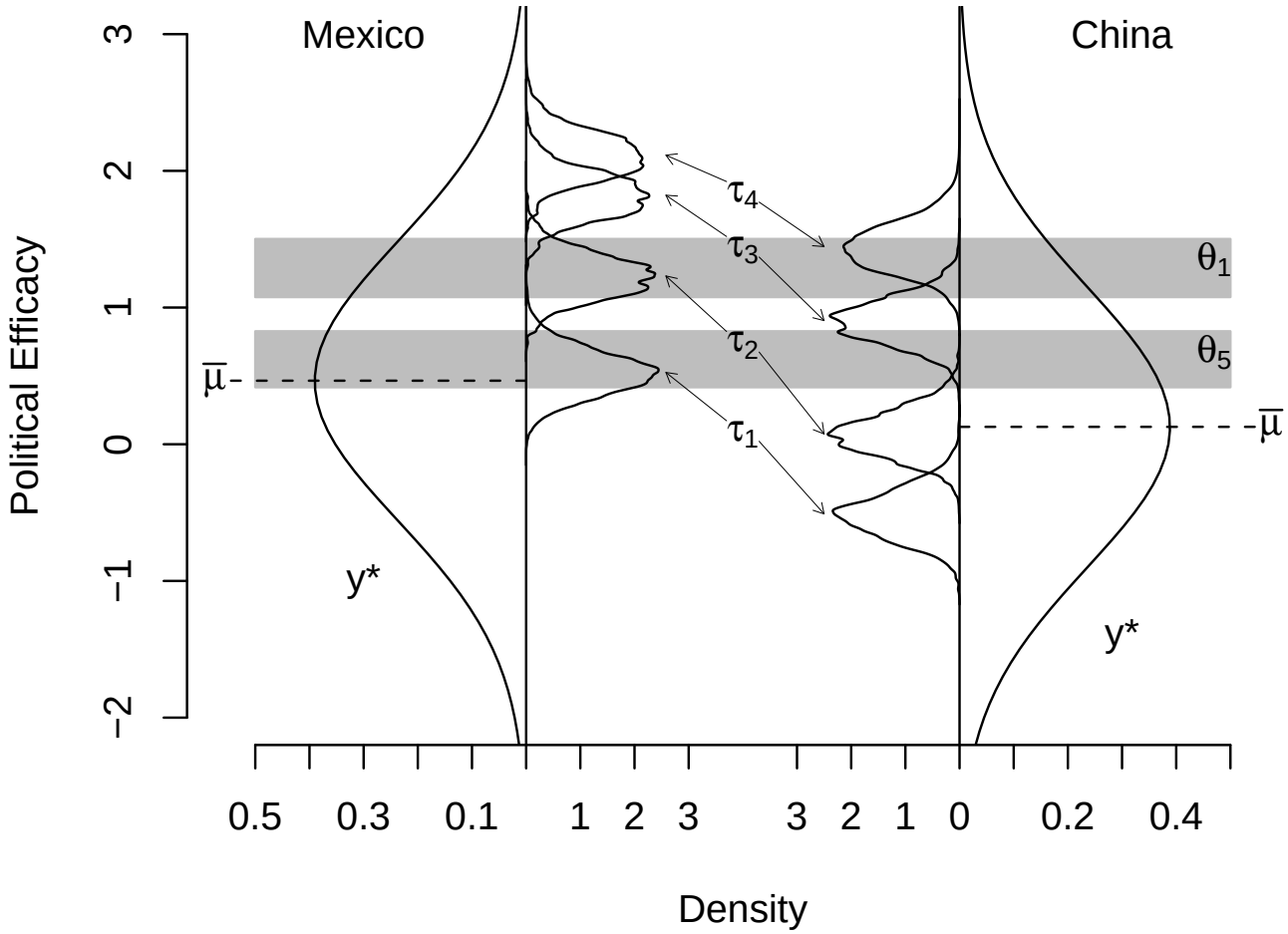
Fixing DIF in China and Mexico

		Ordinal Probit		Chopit	
Eqn.	Variable	Coeff.	(s.e.)	Coeff.	(s.e.)
μ	China	.670	(.081)	−.362	(.090)
	age	.004	(.003)	.006	(.003)
	male	.087	(.076)	.113	(.081)
	education	.020	(.008)	.019	(.008)
Vignettes	θ_1			1.393	(.190)
	θ_2			1.304	(.190)
	θ_3			.953	(.189)
	θ_4			.902	(.188)
	θ_5			.729	(.188)
	$\ln \sigma$			−.238	(.042)

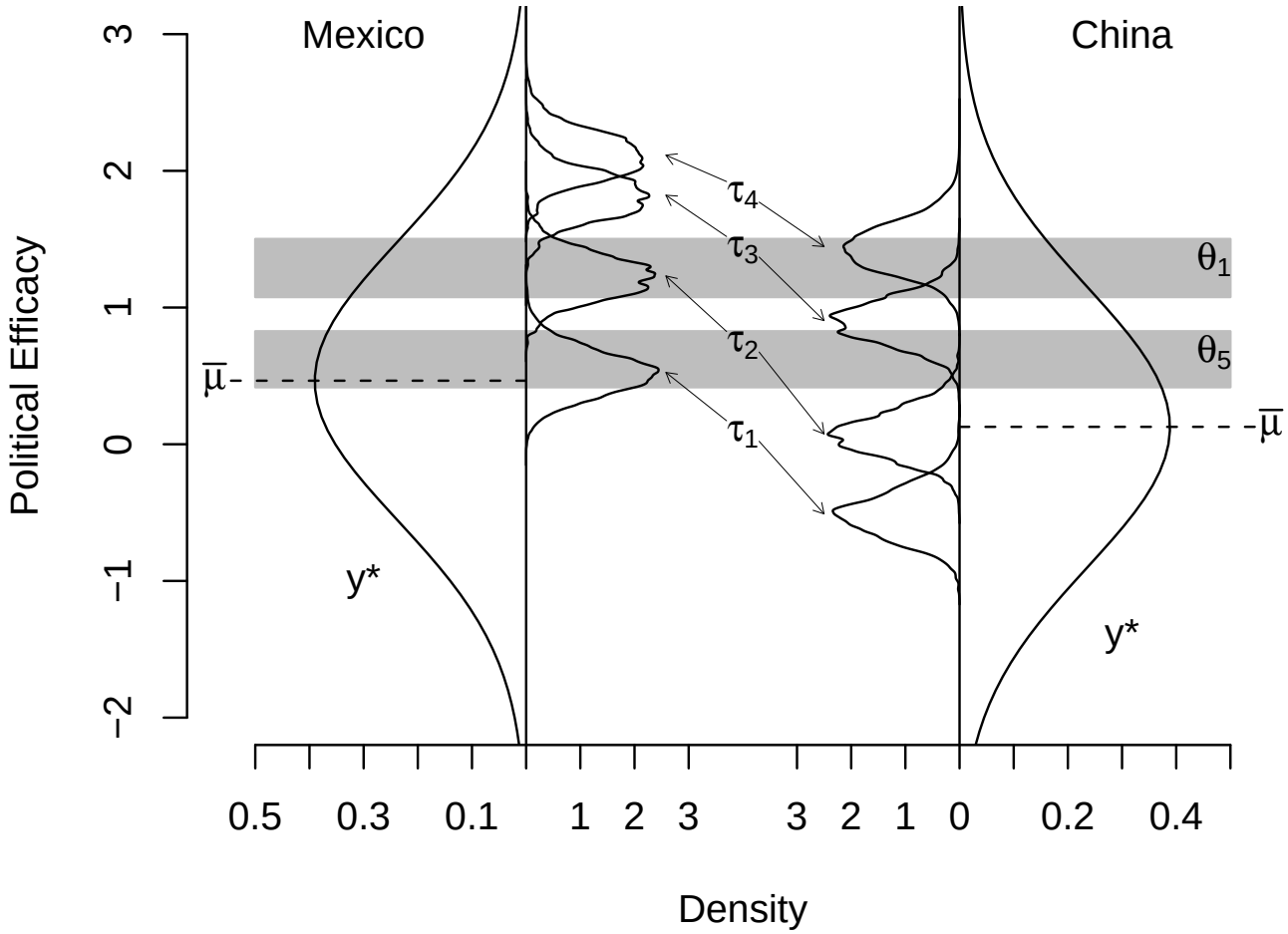
- Ordered Probit: Chinese people have *more* political efficacy than Mexicans
- Chopit: Chinese people have *less* political efficacy than Mexicans

DIF as Threshold Variation: Mexico v. China

DIF as Threshold Variation: Mexico v. China

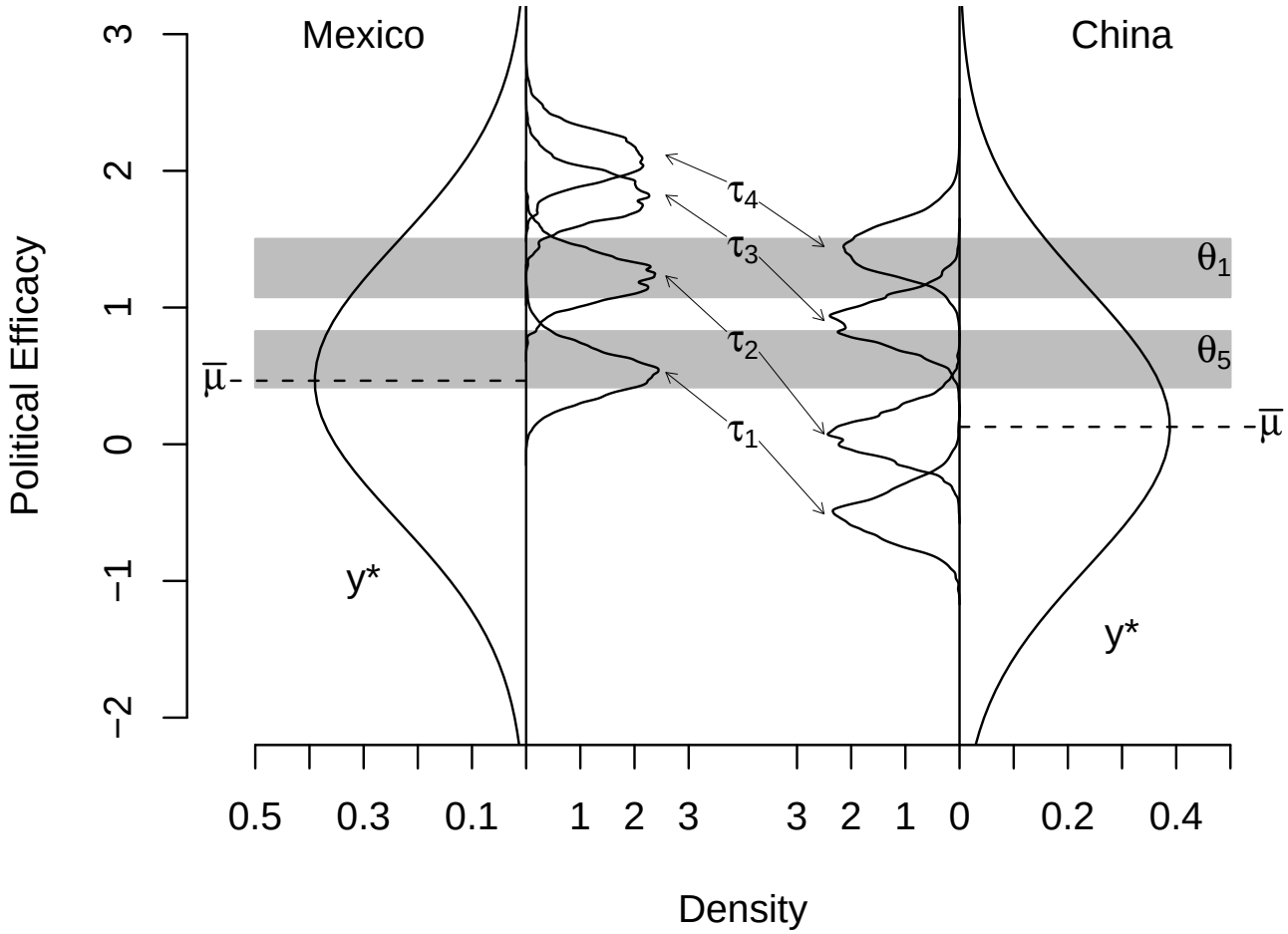


DIF as Threshold Variation: Mexico v. China



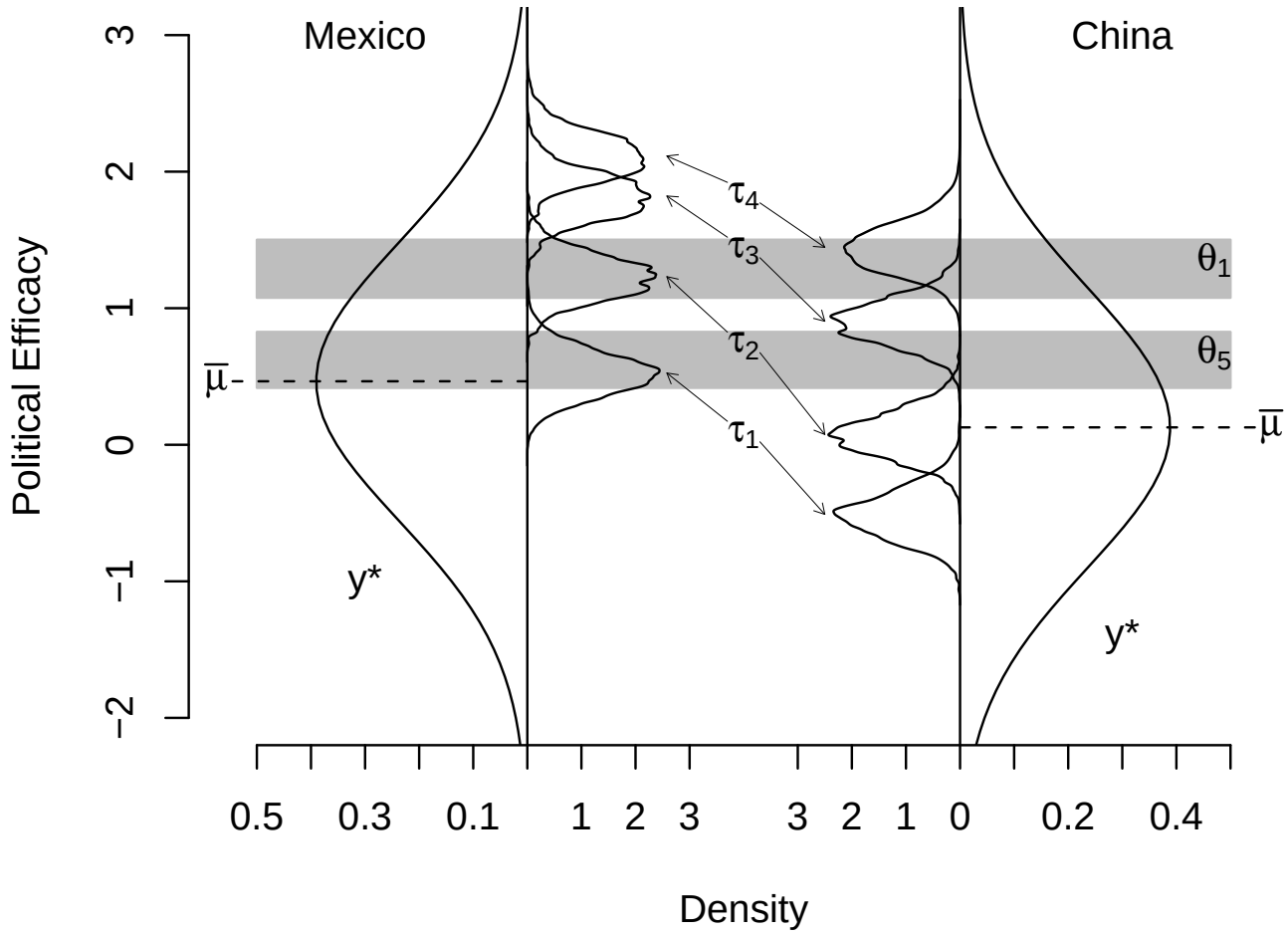
- Self-assessments: Actual (mean), μ; Perceived y^*

DIF as Threshold Variation: Mexico v. China



- Self-assessments: Actual (mean), μ; Perceived y^*
- Vignette Assessments: θ_1, θ_2

DIF as Threshold Variation: Mexico v. China



- Self-assessments: Actual (mean), μ; Perceived y^*
- Vignette Assessments: θ_1, θ_2
- Thresholds: Chinese have lower standards for satisfying τ_j

Summary: Anchoring Vignettes to fix DIF

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete
- Writing vignettes: aids in clarification and discovery of domains of concept of interest — even if you do not do a survey.

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete
- Writing vignettes: aids in clarification and discovery of domains of concept of interest — even if you do not do a survey.
- Other common survey problems: Question wording,

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete
- Writing vignettes: aids in clarification and discovery of domains of concept of interest — even if you do not do a survey.
- Other common survey problems: Question wording, Accurate translation,

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete
- Writing vignettes: aids in clarification and discovery of domains of concept of interest — even if you do not do a survey.
- Other common survey problems: Question wording, Accurate translation, Question order,

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete
- Writing vignettes: aids in clarification and discovery of domains of concept of interest — even if you do not do a survey.
- Other common survey problems: Question wording, Accurate translation, Question order, Sampling design,

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete
- Writing vignettes: aids in clarification and discovery of domains of concept of interest — even if you do not do a survey.
- Other common survey problems: Question wording, Accurate translation, Question order, Sampling design, Interview length,

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete
- Writing vignettes: aids in clarification and discovery of domains of concept of interest — even if you do not do a survey.
- Other common survey problems: Question wording, Accurate translation, Question order, Sampling design, Interview length, Social backgrounds of interviewer and respondent

Summary: Anchoring Vignettes to fix DIF

- Assumptions: response consistency, vignette equivalence
- Probably won't eliminate all DIF: but problems would have to be unrealistically big to make unadjusted measures better than the adjusted ones
- Reduce expense: assign each vignette to subsample. E.g., 4 vignettes asked for 1/4 of the sample
- If you think you have DIF-free questions: you now have the first real opportunity to test that hypothesis.
- Whether or not you have DIF: vignettes can help us follow standard advice of making questions concrete
- Writing vignettes: aids in clarification and discovery of domains of concept of interest — even if you do not do a survey.
- Other common survey problems: Question wording, Accurate translation, Question order, Sampling design, Interview length, Social backgrounds of interviewer and respondent
- Other uses: Defining concepts you only know when you see