Comparing latent inequality with ordinal health data

David M. Kaplan
University of Missouri

Longhao Zhuo
University of Missouri

Midwest Econometrics Group
October 2018

Outline

Outline

Motivation: health inequality

From Deaton and Paxson (1998, pp. 248–9):

- "Our interest in health inequality stems from a more general interest in the distribution of welfare."

Motivation: health inequality

From Deaton and Paxson (1998, pp. 248–9):

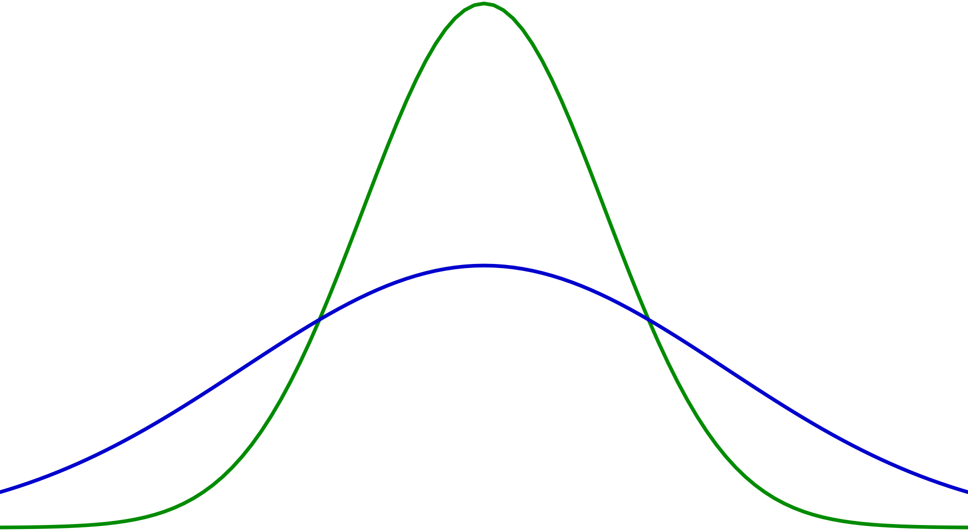
- "Our interest in health inequality stems from a more general interest in the distribution of welfare."
- SRHS: "self-reported health status"
- SRHS scale: excellent, very good, good, fair, poor.

Motivation: health inequality

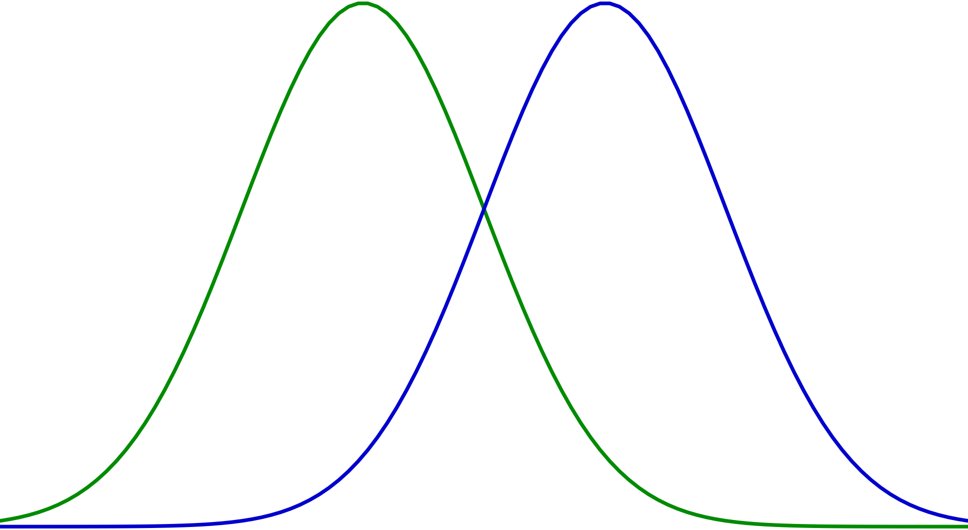
From Deaton and Paxson (1998, pp. 248–9):

- "Our interest in health inequality stems from a more general interest in the distribution of welfare."
- SRHS: "self-reported health status"
- SRHS scale: excellent, very good, good, fair, poor.
- SRHS benefits:
  1. "Useful over the complete adult life cycle"
  2. Strongly correlated with objective measures
  3. Widely available (PSID, NHIS, etc.)
  4. Synthesizes all dimensions of health

Inequality #1: within-group (dispersion)



Inequality #2: between-group (better/worse)



SRHS: latent model

- Why isn't SRHS distribution of ultimate interest?

SRHS: latent model

- Why isn't SRHS distribution of ultimate interest?
  1. No cardinal meaning. (No mean/variance.)

SRHS: latent model

- Why isn't SRHS distribution of ultimate interest?
  1. No cardinal meaning. (No mean/variance.)
  2. Ignores within-category variation. (All "good" identical.)

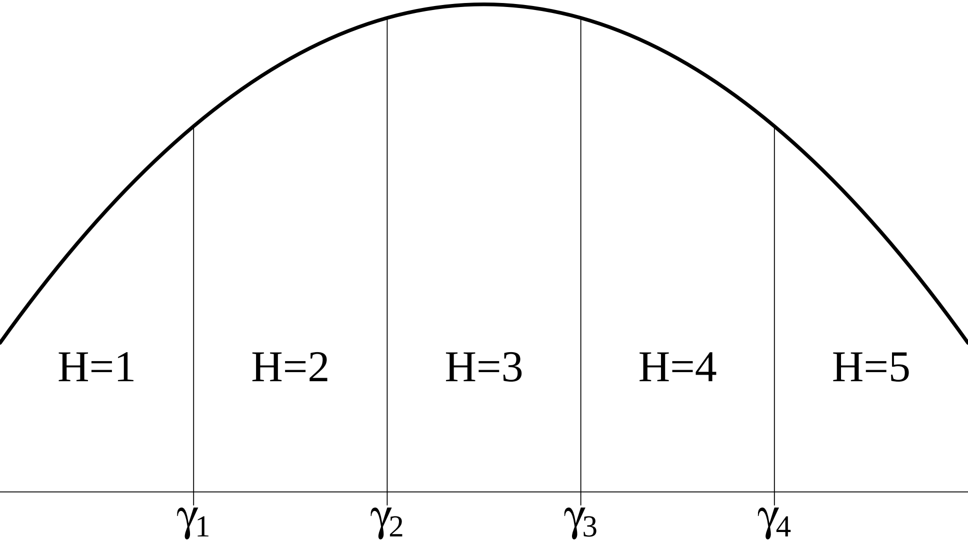
SRHS: latent model

- Why isn't SRHS distribution of ultimate interest?
  1. No cardinal meaning. (No mean/variance.)
  2. Ignores within-category variation. (All "good" identical.)
- Latent health H^*: cardinal, continuous, of interest, but censored.

SRHS: latent model

- Why isn't SRHS distribution of ultimate interest?
  1. No cardinal meaning. (No mean/variance.)
  2. Ignores within-category variation. (All "good" identical.)
- Latent health H^*: cardinal, continuous, of interest, but censored.
- SRHS H, fixed thresholds γ_j.

Latent model



Literature: SRHS inequality methodology

- Goal: compare latent health distributions using ordinal data.

Literature: SRHS inequality methodology

- Goal: compare latent health distributions using ordinal data.
- Literature: either parametric/MLE or discrete latent distribution.
- Literature: statistical inference rare.

Literature: SRHS inequality methodology

- Goal: compare latent health distributions using ordinal data.
- Literature: either parametric/MLE or discrete latent distribution.
- Literature: statistical inference rare.
- Allison and Foster (2004): "median-preserving spread," called "the breakthrough in analyzing inequality with [SRHS] data" by Madden (2014, p. 206).

Literature: SRHS inequality methodology

- Goal: compare latent health distributions using ordinal data.
- Literature: either parametric/MLE or discrete latent distribution.
- Literature: statistical inference rare.
- Allison and Foster (2004): "median-preserving spread," called "the breakthrough in analyzing inequality with [SRHS] data" by Madden (2014, p. 206).
- SRHS-based inequality indexes:
  - Good: "complete ordering" of distributions.
  - Bad: many possible indexes/weighting parameters/functions; implicit assumptions.

Our contribution

- Semi/nonparametric restrictions (not parametric).

Our contribution

- Semi/nonparametric restrictions (not parametric).
- Continuous latent distribution (not discrete).

Our contribution

- Semi/nonparametric restrictions (not parametric).
- Continuous latent distribution (not discrete).
- Results: characterize ordinal conditions informative about latent inequality.

Our contribution

- Semi/nonparametric restrictions (not parametric).
- Continuous latent distribution (not discrete).
- Results: characterize ordinal conditions informative about latent inequality.
- Statistical inference: frequentist and Bayesian. (Code examples.)

Outline

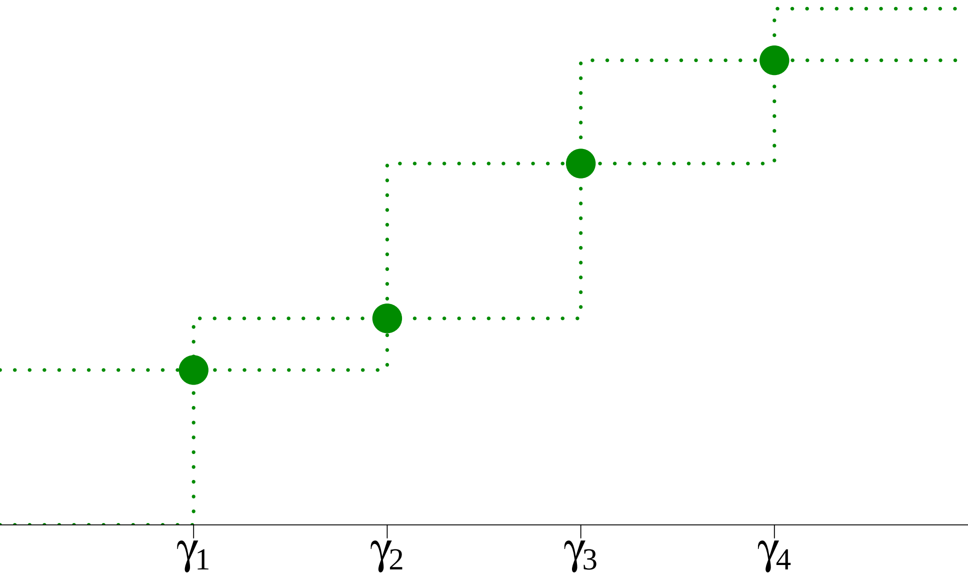
(Partial) Identification

- Given H distributions, what can we learn about H^* distributions?

(Partial) Identification

- Given H distributions, what can we learn about H^* distributions?
- One view: latent CDFs are partially identified given thresholds γ_j.
- Related: Stoye (2010).

(Partial) Identification



Thresholds: assumptions

- Do all populations have the same γ_j thresholds?

Thresholds: assumptions

- Do all populations have the same γ_j thresholds?
- Evidence of "yes" or "constant shift": Lindeboom and van Doorslaer (2004), Hernández-Quevedo, Jones, and Rice (2005).

SD1: Proposition 2

- Between-group inequality.

SD1: Proposition 2

- Between-group inequality.
- "Healthier" if latent first-order stochastic dominance (SD1).
  - $X \; \mathrm{SD}_1 \; Y \iff F_X(\cdot) \le F_Y(\cdot)$

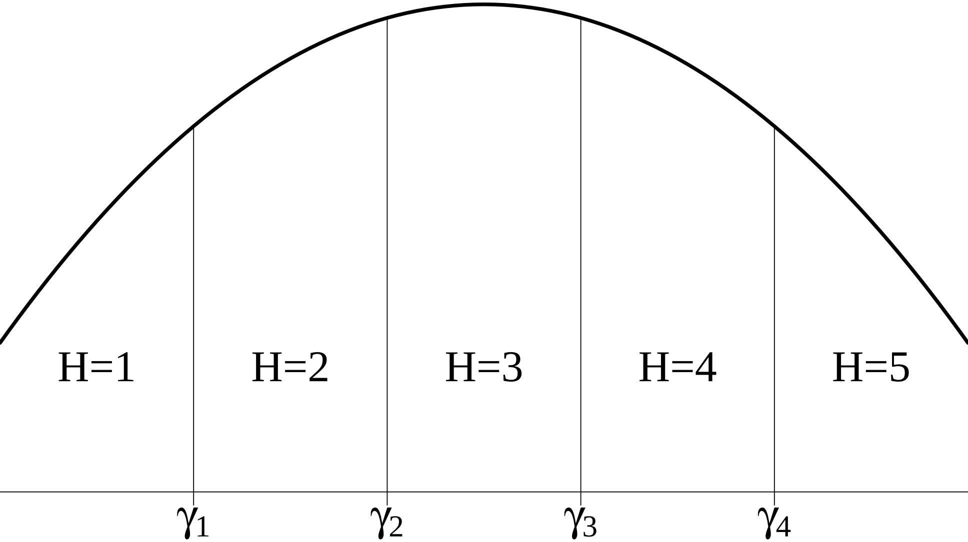
SD1: Proposition 2

- Between-group inequality.
- "Healthier" if latent first-order stochastic dominance (SD1).
  - $X \text{ SD}_1 \; Y \iff F_X(\cdot) \le F_Y(\cdot)$
- Ordinal SD1 $\implies$ latent SD1?
- Latent SD1 $\implies$ ordinal SD1?

SD1: Proposition 2

- Between-group inequality.
- "Healthier" if latent first-order stochastic dominance (SD1).
  - $X \ \mathrm{SD}_1 \ Y \iff F_X(\cdot) \leq F_Y(\cdot)$
- Ordinal SD1 $\implies$ latent SD1?
- Latent SD1 $\implies$ ordinal SD1?
- Proposition 2(i–vi) in paper.

Latent model (again)



SD1: Proposition 2

- Same γ_j.

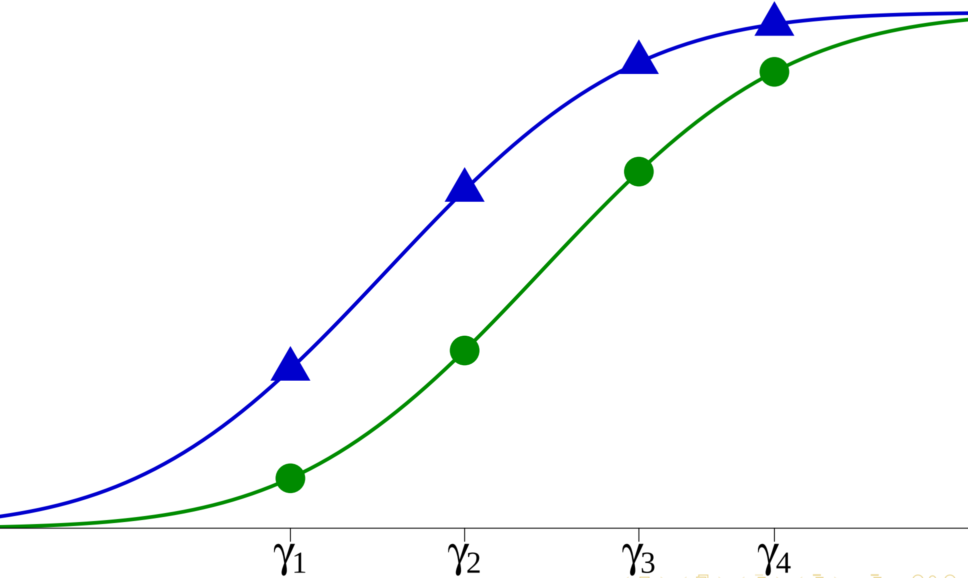
SD1: Proposition 2

- Same γ_j.
- Prop 2(i): latent SD1 $\implies$ ordinal SD1, but not $\impliedby$.
  - Reject ordinal SD1 $\implies$ reject latent SD1.

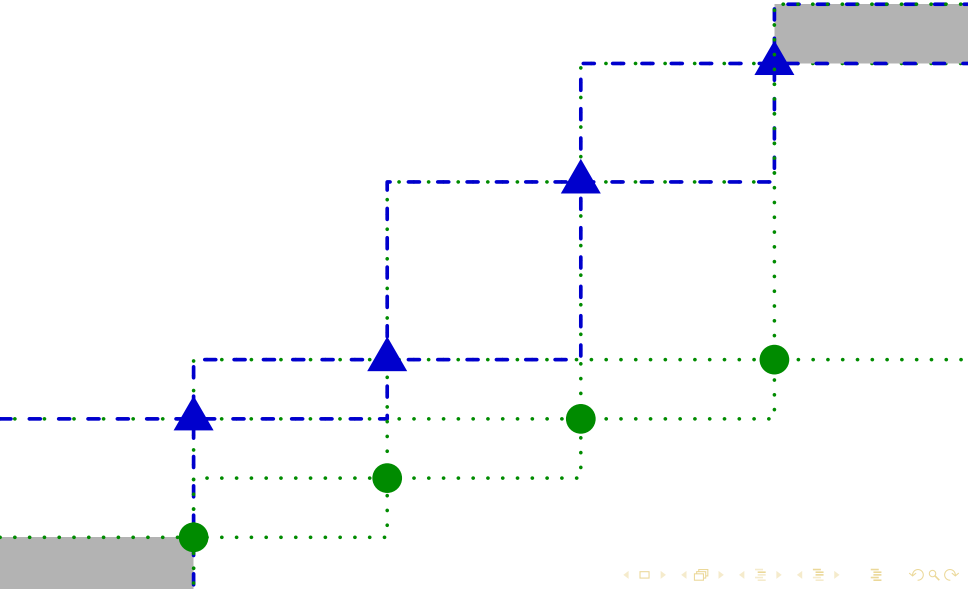
SD1: Proposition 2

- Same γ_j.
- Prop 2(i): latent SD1 $\implies$ ordinal SD1, but not $\impliedby$.
  - Reject ordinal SD1 $\implies$ reject latent SD1.
- Prop 2(v,vi): ordinal SD1 $\implies$ latent "restricted SD1" (Atkinson, 1987).

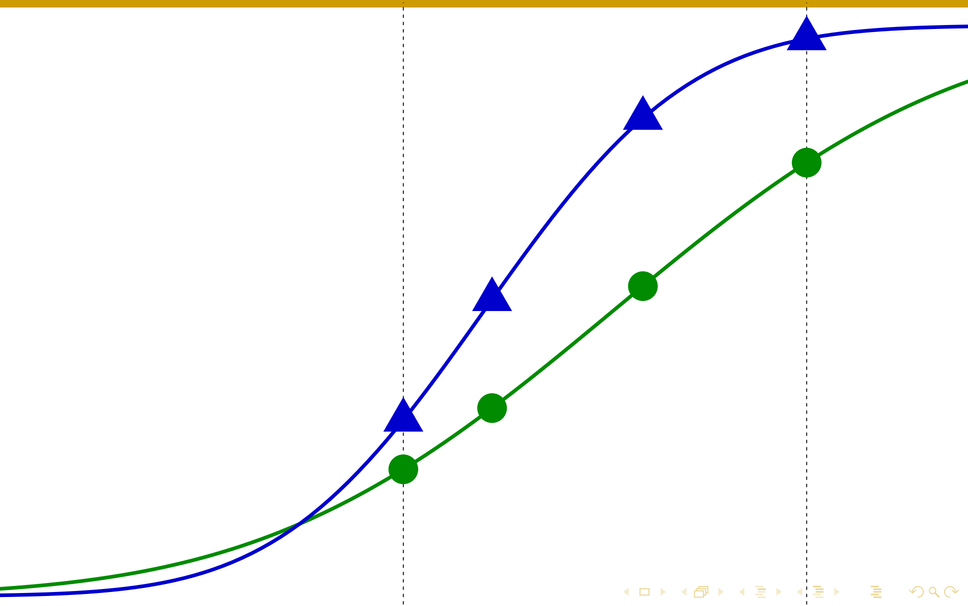
SD1: Prop 2(i)



SD1: Prop 2(v)



SD1: Prop 2(vi)



Dispersion: Proposition 3 (CDF crossing)

- Within-group inequality.

Dispersion: Proposition 3 (CDF crossing)

- Within-group inequality.
- Pure location shift has zero effect on "dispersion."
- Similarly, γ_j can all shift by a constant (unlike for SD1).

Dispersion: Proposition 3 (CDF crossing)

- Within-group inequality.
- Pure location shift has zero effect on "dispersion."
- Similarly, γ_j can all shift by a constant (unlike for SD1).
- Prop 3(i): can learn about latent interquantile range (IQR) differences.

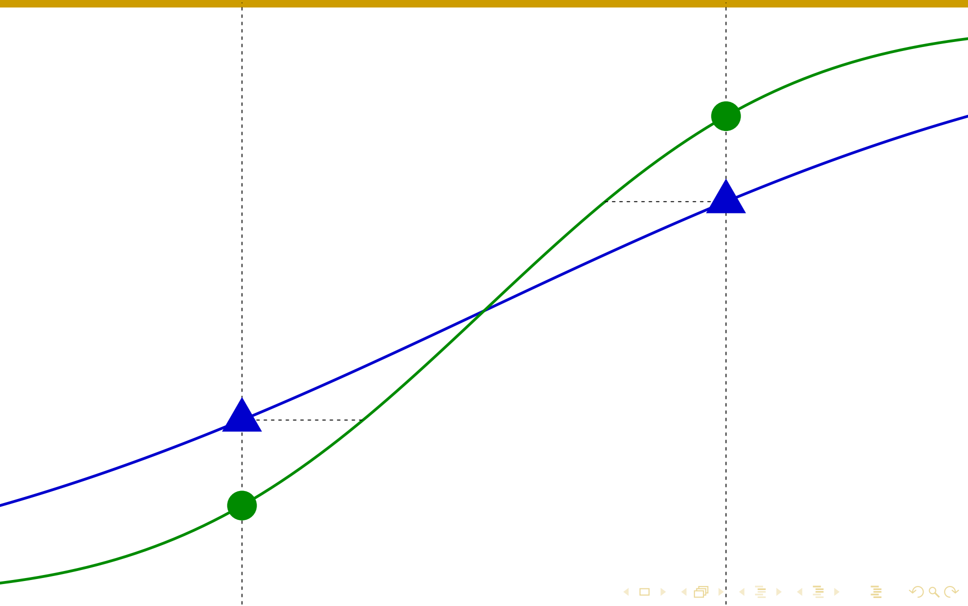
Dispersion: Proposition 3 (CDF crossing)

- Within-group inequality.
- Pure location shift has zero effect on "dispersion."
- Similarly, γ_j can all shift by a constant (unlike for SD1).
- Prop 3(i): can learn about latent interquantile range (IQR) differences.
- Prop 3(ii): location–scale $\implies$ extrapolate to all IQR differences.

Dispersion: Proposition 3 (CDF crossing)

- Within-group inequality.
- Pure location shift has zero effect on "dispersion."
- Similarly, γ_j can all shift by a constant (unlike for SD1).
- Prop 3(i): can learn about latent interquantile range (IQR) differences.
- Prop 3(ii): location–scale $\implies$ extrapolate to all IQR differences.
- Prop 3(iii): even stronger assumptions $\implies$ latent SD2.

Dispersion: Prop 3(i,ii)



Dispersion: Proposition 4 (fanning out)

- Can we ever infer dispersion changes without a CDF crossing?

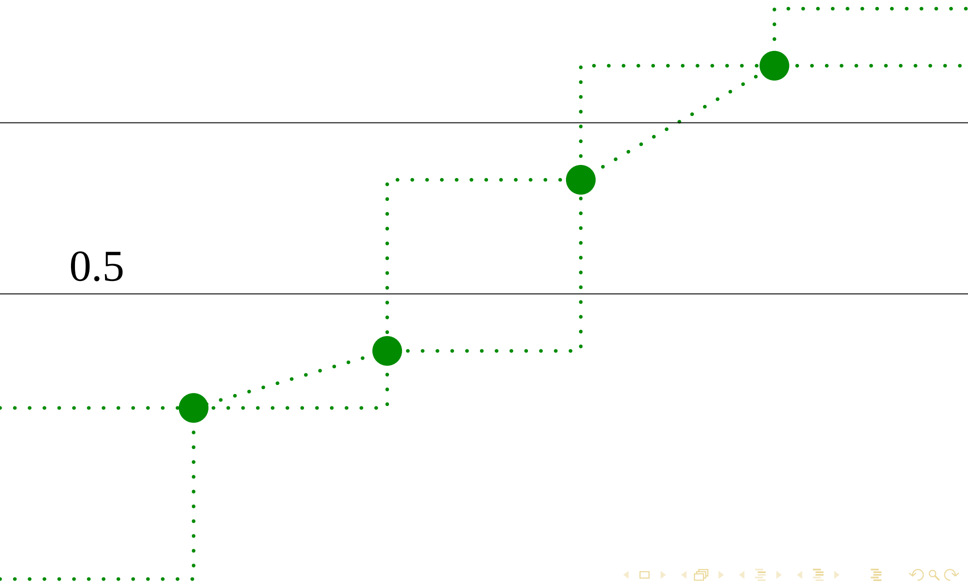
Dispersion: Proposition 4 (fanning out)

- Can we ever infer dispersion changes without a CDF crossing?
- Yes, with stronger assumptions: symmetric, unimodal latent distributions.

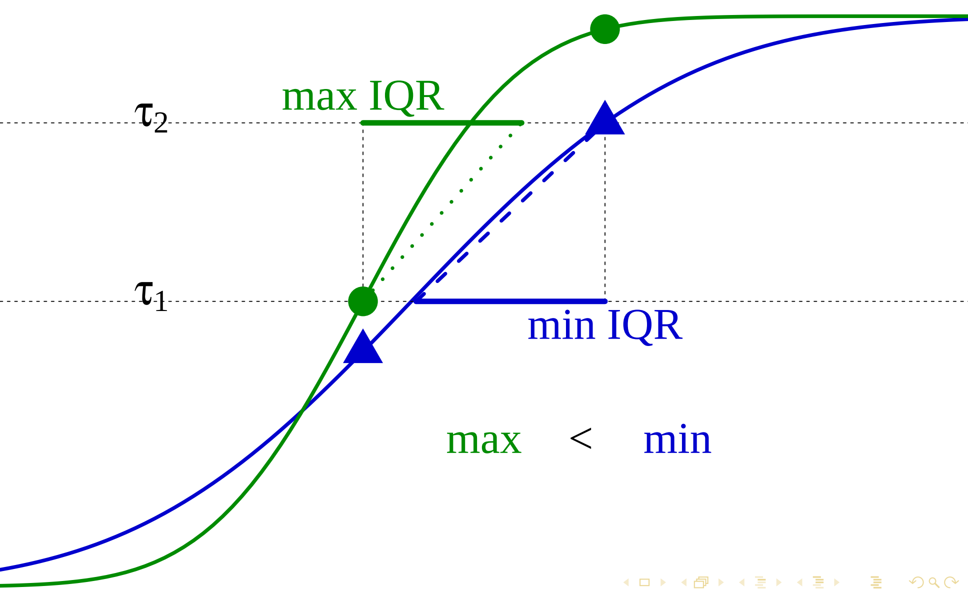
Dispersion: Proposition 4 (fanning out)

- Can we ever infer dispersion changes without a CDF crossing?
- Yes, with stronger assumptions: symmetric, unimodal latent distributions.
- Again, location–scale assumption allows extrapolation.

Dispersion: Prop 4(i)



Dispersion: Prop 4(i)



Outline

Ordinal SD1 inference

- Let $\theta_j \equiv F_2(j) - F_1(j) = \mathbb{E}[\mathbb{1}\{H_2 \leq j\} - \mathbb{1}\{H_1 \leq j\}]$.

Ordinal SD1 inference

- Let $\theta_j \equiv F_2(j) - F_1(j) = \mathbb{E}[\mathbb{1}\{H_2 \le j\} - \mathbb{1}\{H_1 \le j\}]$.
- Ordinal SD1: $H_2 \text{ SD}_1 H_1 \iff \theta_j \le 0,\ j = 1, 2, 3, 4.$

Ordinal SD1 inference

- Let $\theta_j \equiv F_2(j) - F_1(j) = \mathbb{E}[\mathbb{1}\{H_2 \le j\} - \mathbb{1}\{H_1 \le j\}]$.
- Ordinal SD1: $H_2 \text{ SD}_1 H_1 \iff \theta_j \le 0,\ j = 1, 2, 3, 4$.
- Can use recent moment inequality tests: Andrews and Barwick (2012), Romano, Shaikh, and Wolf (2014), McCloskey (2015), et al.

Ordinal SD1 inference

- Let $\theta_j \equiv F_2(j) - F_1(j) = \mathbb{E}[\mathbb{1}\{H_2 \le j\} - \mathbb{1}\{H_1 \le j\}]$.
- Ordinal SD1: $H_2 \ \mathrm{SD}_1 \ H_1 \iff \theta_j \le 0, \ j = 1, 2, 3, 4$.
- Can use recent moment inequality tests: Andrews and Barwick (2012), Romano, Shaikh, and Wolf (2014), McCloskey (2015), et al.
- Bayesian: "nonparametric" posterior for category probabilities $\Longrightarrow$ posterior probabilities for all relationships.

Other relationships

- Other relationships are unions and/or intersections of inequalities.
- Bayes: just compute posteriors.
- Frequentist: intersection–union test (sometimes).

Outline

Treatment effects

- Jones, Molitor, and Reif (2018): large, randomized workplace wellness program.
- Measure SRHS.

Treatment effects

- Jones, Molitor, and Reif (2018): large, randomized workplace wellness program.
- Measure SRHS.
- Can compare treated/untreated latent health distributions using SRHS: does treatment improve health (SD1)? increase inequality? etc.

Treatment effects

- Jones, Molitor, and Reif (2018): large, randomized workplace wellness program.
- Measure SRHS.
- Can compare treated/untreated latent health distributions using SRHS: does treatment improve health (SD1)? increase inequality? etc.
- Can examine selection effects, too.
- Stay tuned. . .

Comparisons of U.S. states

- Goal: compare latent health across different states in U.S.
- Data: 2011 PSID. Around 50–300 observations per state.

Comparisons of U.S. states

- Goal: compare latent health across different states in U.S.
- Data: 2011 PSID. Around 50–300 observations per state.
- Posteriors: Bayesian bootstrap of Dong, Elliott, and Raghunathan (2014) to incorporate stratification, clustering, weights.

PSID 2011 posterior probabilities (%)

		SD_1		SC		fans out	
X	Y	$\succ$	$\prec$	$\succ$	$\prec$	$\succ$	$\prec$
AZ	MO	0	90	4	2	11	34
NY	UT	0	3	0	94	1	70
IL	NY	20	0	66	1	92	3
MN	NY	24	0	57	0	96	3
IA	MO	0	10	2	16	42	98

PSID 2011 posterior probabilities (%)

	X SD$_1$ Y; Y is:					X SC Y; Y is:				
X	MO	KS	NE	IA	IL	MO	KS	NE	IA	IL
MO	—	0	10	10	0	—	7	67*	16	30
KS	34*	—	20*	10	3	6	—	24	6	16
NE	3	0	—	6	0	2	0	—	3	0
IA	0	0	6	—	0	2	0	66*	—	6
IL	40*	4	18*	43*	—	4	14	48	15	—

Asterisk (*): satisfied in-sample.

Outline

Purpose and setup

- Goal: compare type I error rates of different frequentist and Bayesian tests.

Purpose and setup

- Goal: compare type I error rates of different frequentist and Bayesian tests.
- DGP 1: $\mathrm{P}(X = j) = \mathrm{P}(Y = j) = 1/5$, $j = 1, \ldots, 5$; all SD1 inequalities binding.
- DGP 2: change to $\mathrm{P}(X = j) = 1/10$ for $j = 1, 2, 3$ and $\mathrm{P}(X = 4) = 1/2$; only one binding inequality.

Methods

- KS: Kolmogorov–Smirnov.
- RMS: "refined moment selection" of Andrews and Barwick (2012).
- Bayes: Dirichlet–multinomial, uninformative prior on parameters. Reject if posterior below α.
- adj: adjust prior to $\mathrm{P}(H_0) = 1/2$. (Goutis, Casella, and Wells, 1996)

Results: $\alpha = 0.1$

		$H_0: X \text{ SD}_1\ Y$				$H_0: X \text{ SC } Y$		
DGP	n	KS	RMS	Bayes	adj	RMS	Bayes	adj
1	50	0.038	0.089	0.436	0.204	0.032	0.439	0.175
1	100	0.022	0.084	0.430	0.205	0.029	0.359	0.142
1	500	0.027	0.092	0.447	0.199	0.034	0.428	0.171
1	1000	0.032	0.079	0.454	0.228	0.032	0.408	0.155
2	50	0.004	0.057	0.127	0.032	0.031	0.125	0.034
2	100	0.002	0.068	0.105	0.031	0.085	0.133	0.041
2	500	0.006	0.087	0.098	0.029	0.095	0.114	0.032
2	1000	0.003	0.074	0.084	0.025	0.060	0.084	0.018

Results: $\alpha = 0.1$

		$H_0: X \text{ SD}_1 Y$				$H_0: X \text{ SC } Y$		
DGP	n	KS	RMS	Bayes	adj	RMS	Bayes	adj
1	50	0.038	0.089	0.436	0.204	0.032	0.439	0.175
1	100	0.022	0.084	0.430	0.205	0.029	0.359	0.142
1	500	0.027	0.092	0.447	0.199	0.034	0.428	0.171
1	1000	0.032	0.079	0.454	0.228	0.032	0.408	0.155
2	50	0.004	0.057	0.127	0.032	0.031	0.125	0.034
2	100	0.002	0.068	0.105	0.031	0.085	0.133	0.041
2	500	0.006	0.087	0.098	0.029	0.095	0.114	0.032
2	1000	0.003	0.074	0.084	0.025	0.060	0.084	0.018

RMS "better" than KS

Results: $\alpha = 0.1$

		H_0: X SD$_1$ Y				H_0: X SC Y		
DGP	n	KS	RMS	Bayes	adj	RMS	Bayes	adj
1	50	0.038	0.089	0.436	0.204	0.032	0.439	0.175
1	100	0.022	0.084	0.430	0.205	0.029	0.359	0.142
1	500	0.027	0.092	0.447	0.199	0.034	0.428	0.171
1	1000	0.032	0.079	0.454	0.228	0.032	0.408	0.155
2	50	0.004	0.057	0.127	0.032	0.031	0.125	0.034
2	100	0.002	0.068	0.105	0.031	0.085	0.133	0.041
2	500	0.006	0.087	0.098	0.029	0.095	0.114	0.032
2	1000	0.003	0.074	0.084	0.025	0.060	0.084	0.018

RMS "better" than KS

Even with Bayes (adj), cannot treat posterior as p-value, or vice-versa

Outline

Conclusion

- Summary: link latent and ordinal inequality, without parametric or discrete assumptions.
- Summary: relationships defined by inequalities; frequentist and Bayesian inference (code on website).

Conclusion

- Summary: link latent and ordinal inequality, without parametric or discrete assumptions.
- Summary: relationships defined by inequalities; frequentist and Bayesian inference (code on website).
- Future work: other shape restrictions? multivariate? regression? health, happiness, . . . ?

Conclusion

- Summary: link latent and ordinal inequality, without parametric or discrete assumptions.
- Summary: relationships defined by inequalities; frequentist and Bayesian inference (code on website).
- Future work: other shape restrictions? multivariate? regression? health, happiness, . . . ?
- Thank you!
- (And further questions or comments are welcome)

References I

Allison, R. A., Foster, J. E., 2004. Measuring health inequality using qualitative data. Journal of Health Economics 23 (3), 505–524.
URL https://doi.org/10.1016/j.jhealeco.2003.10.006

Andrews, D. W. K., Barwick, P. J., 2012. Inference for parameters defined by moment inequalities: A recommended moment selection procedure. Econometrica 80 (6), 2805–2826.
URL https://www.jstor.org/stable/23357242

Atkinson, A. B., 1987. On the measurement of poverty. Econometrica 55 (4), 749–764.
URL https://doi.org/10.2307/1911028

References II

Deaton, A., Paxson, C., 1998. Aging and inequality in income and health. American Economic Review (Papers and Proceedings) 88 (2), 248–253.
URL https://www.jstor.org/stable/116928

Dong, Q., Elliott, M. R., Raghunathan, T. E., 2014. A nonparametric method to generate synthetic populations to adjust for complex sampling design features. Survey Methodology 40 (1), 29.
URL https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X201400114003

Goutis, C., Casella, G., Wells, M. T., 1996. Assessing evidence in multiple hypotheses. Journal of the American Statistical Association 91 (435), 1268–1277.
URL https://www.jstor.org/stable/2291745

References III

Hernández-Quevedo, C., Jones, A. M., Rice, N., 2005. Reporting bias and heterogeneity in self-assessed health. evidence from the British Household Panel Survey. HEDG Working Paper 05/04, Health, Econometrics and Data Group, The University of York. URL https://ideas.repec.org/p/yor/hectdg/05-04.html

Jones, D., Molitor, D., Reif, J., 2018. What do workplace wellness programs do? evidence from the Illinois Workplace Wellness Study. NBER Working Paper 24229, National Bureau of Economic Research. URL http://www.nber.org/papers/w24229

Lindeboom, M., van Doorslaer, E., 2004. Cut-point shift and index shift in self-reported health. Journal of Health Economics 23 (6), 1083–1099. URL https://doi.org/10.1016/j.jhealeco.2004.01.002

References IV

Madden, D., 2014. Dominance and the measurement of inequality. In: Culyer, A. J. (Ed.), Encyclopedia of Health Economics. Vol. 1. Elsevier, pp. 204–208.
URL https://doi.org/10.1016/B978-0-12-375678-7.00725-2

McCloskey, A., 2015. On the computation of size-correct power-directed tests with null hypotheses characterized by inequalities, working paper, available at http://www.brown.edu/Departments/Economics/Faculty/Adam_McCloskey/Research.html.

Romano, J. P., Shaikh, A. M., Wolf, M., 2014. A practical two-step method for testing moment inequalities. Econometrica 82 (5), 1979–2002.
URL https://www.jstor.org/stable/24029299

References V

Stoye, J., 2010. Partial identification of spread parameters. Quantitative Economics 1 (2), 323–357.
URL https://doi.org/10.3982/QE24