The Mismeasure of Association: The Unsoundness of the Rate Ratio and Other Measures That Are Affected by the Prevalence of an Outcome

Inequality and Methods Workshop,
Minnesota Population Center and Division of Epidemiology and Community Health of the School of Public Health,
University of Minnesota,
September 5, 2014

James P. Scanlan
Attorney at Law
Washington, DC
jpscanlan.com

- Correct or Incorrect?

- If correct, how many appraisals of differences between outcome rates of two groups have been sound?

Personal background

- Lawyer in Washington, DC

- EEOC v. Sears, Roebuck and Co., 839 F.2d 302 (7th Cir. 1988)
  - 10-month trial in 1984-85; almost entirely statistical
  - Milkman "Women's History and the Sears Case," *Feminist Studies* (1986)
  - Sears Case page of jpscanlan.com
  - Sears Case Illustration subpage of Scanlan's Rule page
  - See discussion of Table 15 *infra*

Digression re Sears Case

- *Washington Post* quotation of EEOC Commissioner's Decision underlying the Sears case (Feb. 25, 1979):

  > "in retail stores women hold 77 percent of the lower paying, noncommission sales jobs and only 23 percent of the desirable commission sales positions …"

- "Illusions of Job Segregation," *Public Interest* (1988)

- "The Mismeasure of Discrimination," Univ Kansas School of Law Faculty Workshop (2013) (Section F)

- "The Perverse Enforcement of Fair Lending Laws," *Mortgage Banking* (2014)

- Employment Discrimination page of jpscanlan.com (Section A)

Key Points (1)

One: Standard measures of differences between outcome rates (proportions) cannot effectively quantify differences in the circumstances of advantaged and disadvantaged groups because – for reasons inherent in the underlying risk distributions – each measure tends to be systematically affected by the prevalence of an outcome.

- The two relative (percentage) differences
- Absolute (percentage point) differences
- Odds Ratios

Two: Efforts to appraise differences in the circumstances of two groups reflected by a pair of outcome rates in the law and the social and medical sciences have been almost universally undermined by failure to recognize the way chosen measures tend to be affected by the prevalence of an outcome.

Key Points (2)

Three: Even when broadly correct, research employing standard measures of differences between outcome rates is misleading by implying that the measures employed effectively quantify the difference in circumstances of two groups reflected by their differing rates.

Four: There exists only one answer to the question of whether differences in the circumstances of advantaged and disadvantaged groups reflected by their outcome rates have increased or decreased over time or are larger in one setting than another.

Five: That answer can be divined, albeit imperfectly, by deriving from pairs of outcome rates the difference between means of the underlying risk distributions.

Caveat One

- Do not be distracted by the fact that one commonly finds departures from the patterns described here. Observed patterns are invariably functions of
  - (a) the strength of the forces causing rates to differ (differences in the circumstances of the groups being compared) and
  - (b) the prevalence-related/distributionally-driven forces described here.
- Society's interest is in (a).
- Only with a mastery of (b) can one understand (a).

Caveat Two

- Do not think that presenting relative and absolute differences (or even both of the two relative differences and the absolute difference) addresses the issues raised here.
- The fundamental problem is that none of the measures is statistically sound.

Clinical Settings

- Discussion here will largely focus in advantaged and disadvantaged groups.
- Points made apply equally in clinical setting where treated subjects are the advantaged group and control subjects are the disadvantaged group.
- See Subgroup Effects subpage of the Scanlan's Rule page.

Key references

(1) "Race and Mortality Revisited," Society (2014)*

(2) "Can We Actually Measure Health Disparities?," Chance (2006)

(3) "Race and Mortality," Society (2000)

(4) "Divining Difference," Chance (1994)

(5) "The Perils of Provocative Statistics," Public Interest (1991)

(6) "'Feminization of Poverty' is Misunderstood," Plain Dealer (1987)

Some other extended treatments

- "Measuring Health and Healthcare Disparities," Federal Committee on Statistical Methodology 2013 Research Conference. PowerPoint presentation is succinct and well annotated.

- "The Mismeasure of Discrimination," Faculty Workshop, Univ of Kansas School of Law (2013)

- Letter to Harvard University, Mass. General Hospital et al. re Commissioned Paper (2012) [see Society 2014 at 16-18] *

- Letter to Harvard University re Measurement at Harvard (2012)*

- "The Mismeasure of Group Differences in the Law and the Social and Medical Sciences," Applied Stat Workshop, Harvard Inst for Quantitative Social Science (2012)

- "The Misinterpretation of Health Inequalities in the United Kingdom," Brit. Soc. for Pop. Studies (BSPS) 2006 Conference

Measurement pages of jpscanlan.com

- Measuring Health Disparities (MHD)
  - Journal Comments (144)*
  - Whitehall Studies
- Scanlan's Rule (SR)
  - Subgroup Effects
  - Illogical Premises
  - Collected Illustrations
  - Consensus
- Immunization Disparities
- Mortality and Survival
- Immunization Disparities
- Educational Disparities
- Disparate Impact
- Discipline Disparities
- Lending Disparities
- Employment Discrimination
- Feminization of Poverty
- Vignettes

Institutional Correspondence

- Education Law Center (Aug. 14, 2014)
- IDEA Data Center (Aug. 11, 2014) [see Tables 19 and 20 *infra*]
- Institute of Medicine II (May 28, 2014)
- Annie E. Casey Foundation (May 13, 2014) [see Table 24 *infra*]
- Education Trust (April 30, 2014)
- Investig and Oversight Subcomm of House Finance Comm (Dec. 4, 2013)
- Mailman School of Public Health of Columbia University (May 24, 2013)
- Senate Committee on Health, Education, Labor and Pensions (Apr. 1, 2013)
- Federal Reserve Board (March 4, 2013)
- Harvard University , Mass General Hospital, et al. (Oct. 26, 2012)
- Harvard University (Oct. 9, 2012)
- United States Department of Justice (Apr. 23, 2012)
- United States Department of Education (Apr. 18, 2012)
- The Commonwealth Fund (June 1, 2010)
- Institute of Medicine (June 1, 2010)
- National Quality Forum (Oct. 22, 2009)
- Robert Wood Johnson Foundation (Apr. 8, 2009)

Interpretive Rule 1 (IR1): The Two Relative Differences (aka Heuristic Rule X (HRX), Scanlan's Rule)

The rarer an outcome

(a) the greater tends to be the relative difference in experiencing it and

(b) the smaller tends to be the relative difference in avoiding it.

IR1 Implications – General

- As mortality and poverty decline, relative differences in experiencing those outcomes tend to increase while relative differences in avoiding them tend to decrease.

- As procedures like immunization and cancer screening become more common, relative differences in receipt of those procedures tend to decrease while relative differences in failing to receive them tend to increase.

- More survivable cancers tend to show larger relative differences in mortality, but smaller relative differences in survival than less survivable cancers. Mortality and Survival page and Table 13 *infra.*

- Generally reducing blood pressure (or improving folate levels) tends to increase relative differences in hypertension (or low folate) while reducing relative differences in normal blood pressure (or adequate folate). NHANES Illustrations subpage of SR.

- Relaxing mortgage lending , employment, or public school discipline standards tends to increase relative differences in failing to meet the standards while reducing relative differences in meeting the standards.

IR1 Implications – Comparatively Advantaged Populations/Subpopulations (1)

- Star Tribune Commentary (Feb. 8, 2014): "It's easy to misunderstand gaps and mistake good fortune for a crisis")

- Relative racial, gender, socioeconomic differences in adverse outcomes tend to be larger, while relative differences in the corresponding favorable outcomes tend to be smaller, among comparatively advantaged populations/subpopulations (where the outcomes are less common) than among less advantaged populations/subpopulations.

IR1 Implications – Comparatively Advantaged Populations/Subpopulations (2)

- Racial diff in infant health outcomes among highly-educated or low risk groups ("Race and Mortality")
- Occupational diff in mortality among British Civil Servants (Whitehall Studies)
- Racial, gender, and SES diff in mortality among young (Life Tables Illustrations)
- Racial diff in loan rejection among high-income applicants (Disp – High Income)
- Racial diff in completion/non-completion rates at elite universities ("Race and Mortality")
- Suburban discipline disparities (Suburban Disparities)
- Racial and SES diff in mortality in Norway and Sweden (or Minnesota and Massachusetts)

IR1 Implications – Comparatively Advantaged Populations/Subpopulations (3)

- Racial diff in adverse outcomes among high SES groups; SES differences in adverse outcomes among whites
- Racial diff in healthcare among the insured compare with the uninsured.
- Racial and gender diff in selection among highly qualified applicants.
- Racial diff in suspensions in pre-school versus K12. Table 7 or Society 2014 and Table 8 *infra*.
- Effect of records on employment prospect of whites versus blacks (or effect of being black on employment prospects of those with out or with without criminal. Table 8 of Society 2014 and Table 17 *infra.*

IR1 Implications – Comparatively Advantaged Populations/Subpopulations (4)

- Scholars describe patterns of large racial differences in adverse outcomes among advantaged subpopulations as “poorly understood.”
- It is fairer to say that they are not understood at all.
- Drawing of inferences based on perceptions about either (a) the large relative differences in adverse outcomes or (b) the small relative differences in favorable outcome within advantaged subpopulations has never sound.

Corollary 1 to IR1

As an outcome changes in overall prevalence,

(a) the group with the lower baseline rate outcome will tend to undergo a larger proportionate change in its rate for the outcome, while

(b) other group will tend to undergo a larger proportionate change in its rate for the opposite outcome.

Implications of Corollary 1 to IR 1

- Effects of reductions/increases in poverty
- Effects of lowering/raising cutoffs (improving performance)
- Effects of improving health outcomes
- Explanatory theories: “diffusion of innovation,” “inverse equity hypothesis” (Explanatory Theories)*
- Effects of chronic conditions on self-rated health* (Reporting Heterogeneity, Comment on Delpierre BMC Pub Hlth 2012)
- Subgroup Effects subpage of SR
- Subgroup Effects – Nonclinical subpage (Minneapolis housing study)*

Corollary 2 to IR1

When an outcome declines in overall prevalence, there will tend to be an increases in the proportion the most susceptible group comprises of both

(a) those experiencing the outcome; and

(b) those failing to experience the outcome.

(Feminization of Poverty, Table 1 of Chance 2006)

Implications of Corollary 2 to IR1

- Feminization of Poverty
- Racial impact of Proposition 48
- ISAIAH study of housing crisis in St. Paul (Kansas paper)
- Any discussion of the proportion a group comprise of persons experiencing some adverse outcome (addressed *infra*)

Absolute Differences/Odds Ratios

- Absolute differences and differences measured by odds ratios are unaffected by whether one examines the favorable or the adverse outcome.
- But an effective indicator must remain constant when there occurs a change in overall prevalence akin to that effected by lowering a test cutoff.
- Absolute differences and odds ratios tend also to be affected by the prevalence of an outcome, but in a more complicate way than the two relative differences.

Interpretive Rule 2(IR 2): Absolute Differences/Odds Ratios

- As an outcome goes from being rare to being universal, absolute differences between rates tend to:

  (a) increase to the point where the first group's rate reaches 50%;

  (b) behave inconsistently until the second group's rate reaches 50%;

  (c) then decline.

- As the prevalence of an outcome changes, differences measured by odds ratios tend to change in the opposite direction of absolute differences.

Relationship of the Absolute Difference to the Two Relative Differences (1)

- As the prevalence of an outcome changes, the absolute difference tends to change in the same direction as the smaller relative difference.
- Since observers commonly choose to report the larger relative difference, there is a systematic tendency for the absolute difference and the reported relative difference to change in opposite directions.

Relationship of the Absolute Difference to the Two Relative Differences (2)

- All measures may change in the same direction, in which case we can assume there has been a change the strength of the forces causing the outcome rates to differ.
- But anytime an observer reports that a relative difference and the absolute difference have changed in opposite direction, the unmentioned relative difference will necessarily have changed in the opposite direction of the mentioned relative difference and the same direction as the absolute difference.

Table 1. Explanation of Terms

(a) AG Fav Rt	(b) DG Fav Rt	(c) AG Adv Rt	(d) DG Adv Rt	(1) AG/DG Ratio Fav	(2) DG/AG Ratio Adv	(3) Abs Df (pp)	(4) Odds Ratio
90%	80%	10%	20%	1.125	2.00	10	2.25

In this presentation, the larger figure is always used as the numerator in the rate ratio (RR); hence the relative difference is always RR -1

(1) AG/DG Ratio Fav = a/b (1.125; relative difference is 12.5%) - BLUE

(2) DG/AG Ratio Adv = d/c (2.00; relative difference is 10%) - RED

(3) Abs Df (pp)= a-b (10 percentage points) - GREEN
[see Percentage Point subpage of Vignettes page]

(4) Odd Ratio = (a/c)/(d/b) (2.25)

Table 2: Simplified Illustration of Effects of Lowering Test Cutoff on Relative Difference Between Pass Rates and Relative Difference Between Failure Rates

Cutoff	AG Pass	DG Pass	AG Fail	DG Pass	DG/AG Ratio Pass	DG/AG Ratio Fail
High	80%	63%	20%	37%	1.27	1.85
Low	95%	87%	95%	87%	1.09	2.60

As a result of lowering the cutoff:

(a) Rate ratio for passing decreased from 1.27 to 1.09 (i.e., relative difference between pass rates decreased from 27% to 9%);

(b) Rate ratio for failure increased from 1.85 to 2.60 (i.e., relative difference between pass rates increased from 85 percent to 160%).

Table 2a: Simplified Illustration of Effects of Patterns of the Two Relative Differences in Advantaged and Disadvantaged Setting

Setting	AG Pass	DG Pass	AG Fail	DG Pass	DG/AG Ratio Pass	DG/AG Ratio Fail
Disadvantaged (e.g., inner city)	80%	63%	20%	37%	1.27	1.85
Advantaged (e.g., suburbs)	95%	87%	5%	87%	1.09	2.60

Advantaged setting has larger difference in failure rates but smaller difference in pass rates.

Table 3: Simplified Illustration of Effects of Lowering Test Cutoff on Relative Difference Between Pass Rates and Relative Difference Between Failure Rates (with IR1 corollaries 1 and 2)

Cutoff	AG Pass	DG Pass	AG Fail	DG Pass	DG/AG Ratio Pass	DG/AG Ratio Fail
High	80%	63%	20%	37%	1.27	1.85
Low	95%	87%	5%	13%	1.09	2.60

Corollary 1: Lowering the cutoff caused:
(a) pass rates to increase by 38% for DG but only 19% for AG;
(b) failure rates to decrease by 75% for AG but only 65% for DG.

Corollary 2: Lowering the cutoff t(assuming equal-sized groups) caused:
(a) prop DG comprised of passes to increase from to 44% to 48%;
(b) prop DG comprised of fails to increase from 65% to 72%.

Table 4: Simplified Illustration of Effects of Lowering Test Cutoff on Relative Difference Between Pass Rates and Relative Difference Between Failure Rates (with absolute differences and odds ratios)

Cutoff	AG Pass	DG Pass	DG/AG Ratio Pass	DG/AG Ratio Fail	Abs Df (pp)	Odds Ratio
High	80%	63%	1.27	1.85	17	2.35
Low	95%	87%	1.09	2.60	8	2.84

Fig. 1. Ratios of (1) DG Fail Rate to AG Fail Rate and (2) AG Pass Rate to DG Pass Rate at Various Cutoff Points Defined by AG Fail Rate

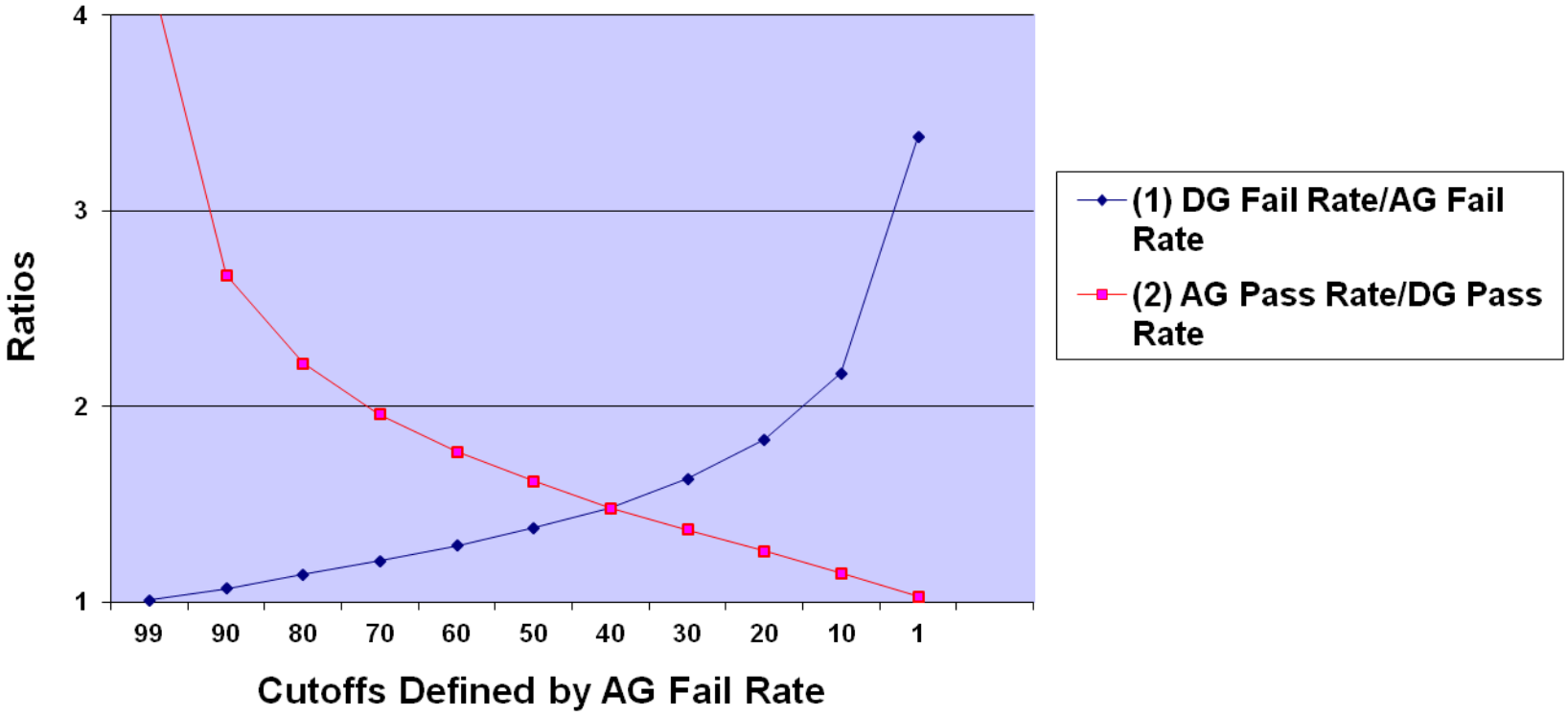


Fig. 2: Absolute Difference Between Rates at various Cutoffs Defined by AG Fail Rate

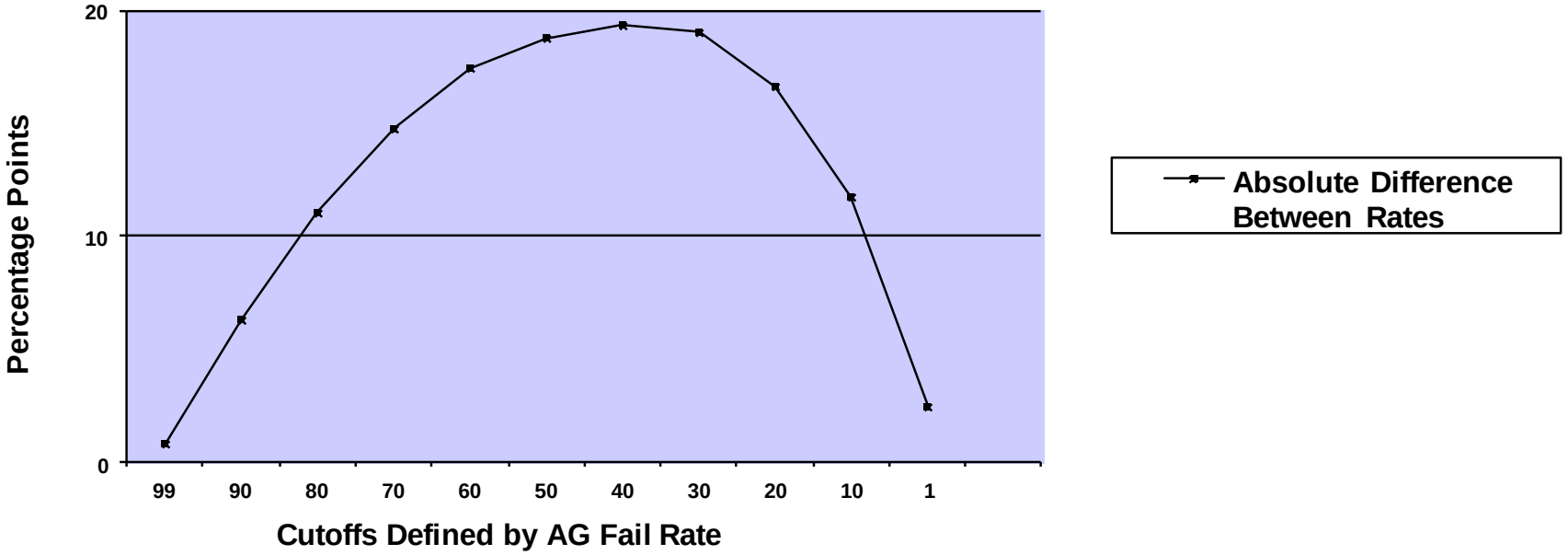
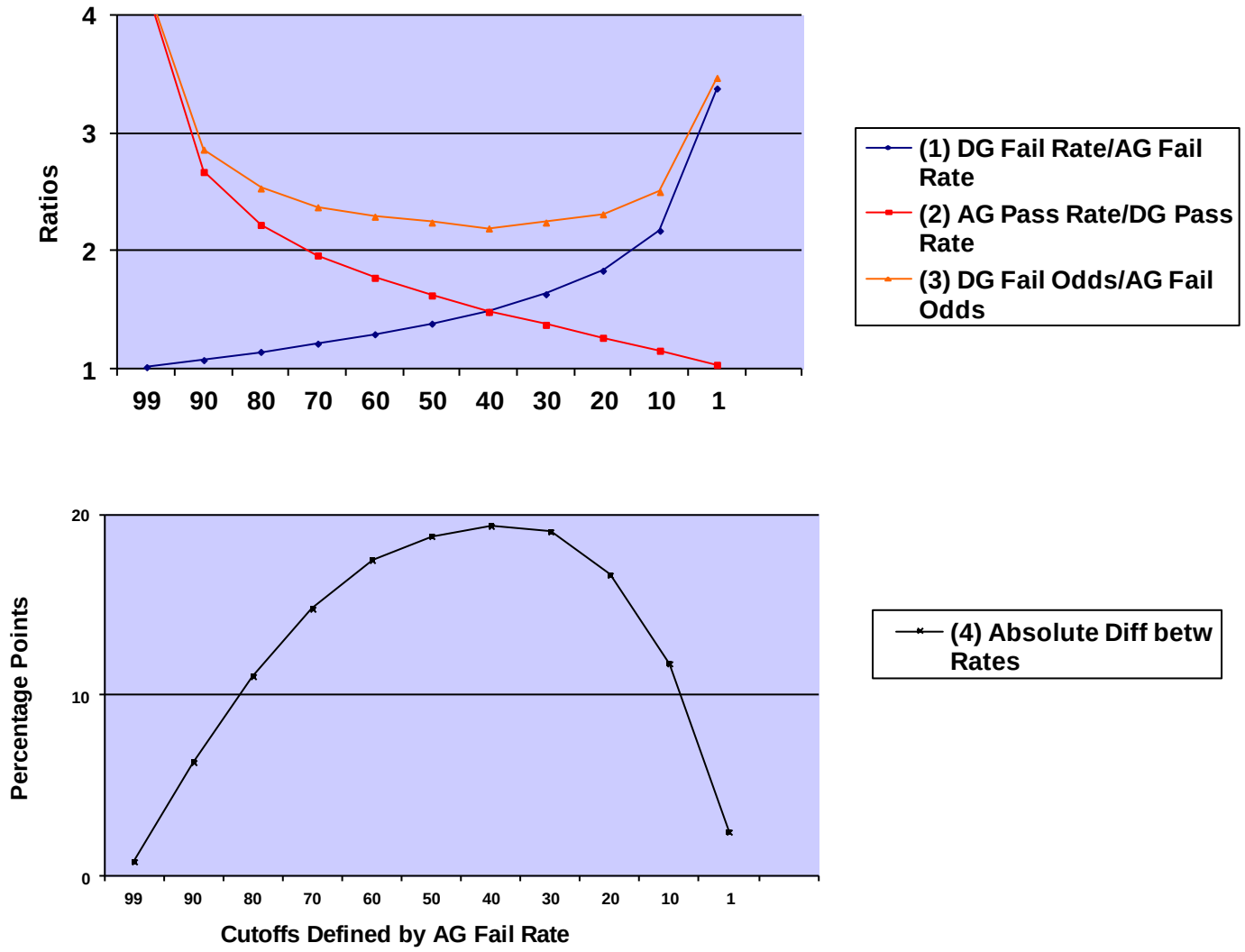


Fig. 3 Ratios of (1) DG Fail Rate to AG Fail Rate, (2) AG Pass Rate to DG Pass Rate, (3) DG Failure Odds to AG Failure Odds; and (4) Absolute Difference Between Rates



Notes on Lowering Standards

- For years, federal agencies have been encouraging mortgage lenders and public schools to relax lending and discipline criteria under the mistaken belief that doing so will reduce relative (racial/ethnic) differences in adverse borrower/discipline outcomes.
- Federal agencies continue to monitor fairness of practices on the basis of relative differences in adverse outcomes.
- By responding to federal encouragements to relax standards, lenders and public schools increase the chances that the federal government will accuse them of discrimination.
- No agency of government is aware, in any institutional sense, that lowering a test cutoff tends to increase relative differences in failure rates (save, to a degree, NCHS, as will be discussed).

References re lowering standards

- Society 2014 at 14-16
- "The Perverse Enforcement of Fair Lending Laws," *Mortgage Banking* (2014)
- "Things government doesn't know about racial disparities," The Hill (2014).
- "The Paradox of Lowering Standards," *Baltimore Sun* (2013)
- "Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies," *Amstat News* (Dec. 2012)
- **"Getting it Straight When Statistics Can Lie," *Legal Times* (1993) (Fisher v. Transco Services, IRS, Postal Service) ***
- **"An Issue of Numbers," *National Law Journal* (1990) (NCAA Proposition 42) ***
- Lending Disparities page and subpages
- **Discipline Disparities page and subpages, especially Minneapolis Disparities and St. Paul Disparities subpages**

Table 5. Illustration of Effect on Standard Measures of Reducing Poverty Such as to Enable Everyone with and Income Above 75 Percent of Poverty Line to Escape Poverty

(2004 data from "Can We Actually Measure Health Disparities," *Chance* (2006))

Row #	Perc of Pov Line	Prop of Wh Above	Prop of Bl Above	Prop of Wh Below	Prop of Bl Below	B/W Ratio Below	W/B Ratio Above	Abs Df (PP)	Odds Ratio
1 (bef)	100	89.2%	75.3%	10.8%	24.7%	2.29	1.18	13.9	2.71
2 (aft)	75	92.8%	82.2%	7.2%	17.8%	2.47	1.13	10.6	2.79

Fig. 4. Ratios of (1) Black to White Rates of Falling Below Percentages of Poverty Line, (2) White to Black Rates of Falling Above the Percentage, (3) Black to White Odds of Falling Below the Percentage, and (4) Absolute Differences Between Rates

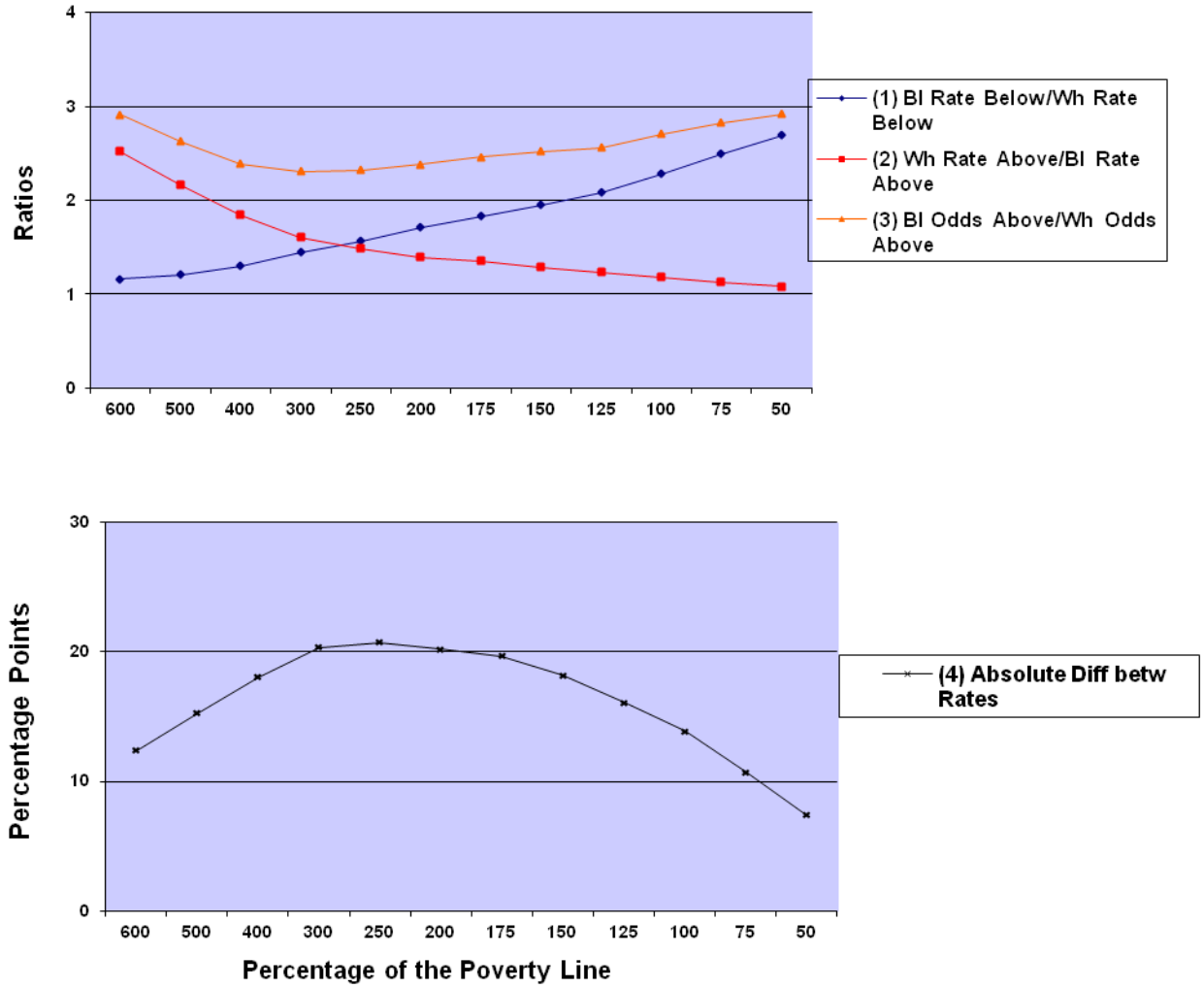


Table 6. Illustration of Effect on Standard Measures of Reducing Poverty Such as to Enable Everyone with and Income Above 75 Percent of Poverty Line to Escape Poverty (with EES)

Perc of Pov Line	Prop Wh Below	Prop Bl Below	B/W Ratio Below	W/B Ratio Above	Abs Df (PP)	Odds Ratio	EES
100	10.8%	24.7%	2.29	1.18	13.9	2.71	.55
75	7.2%	17.8%	2.47	1.13	10.6	2.79	.54

Question: Could one justify exploring the reasons for changes in any of the standard measures measure – say, to evaluate the role of a particular administration's civil rights enforcement policy in the differences – without consideration of the patterns described here?

EES (Estimate Effect Size)Explained

- Derive from any pair of outcome rates the differences between means of the (hypothesized) underlying distributions in terms of standard deviations. In test score hypotheticals EES was .50.
- Probit coefficient
- See Solutions subpage of Measuring Health Disparities page regarding limitations, nuances.

Table 7. Illustrations of EES Values

RR Adverse	DG Adverse Rt	AG Adverse Rt	EES	Percent of DG Above AG Mean
1.2	60.0%	50.0%	0.25	40.3%
1.2	18.4%	15.4%	0.12	45.4%
1.5	75.0%	50.0%	0.67	25.3%
1.5	45.0%	30.0%	0.39	35.0%
2	60.0%	30.0%	0.78	22.0%
2	40.0%	20.0%	0.58	28.3%
2	20.0%	10.0%	0.43	33.7%
2	1.0%	0.5%	0.24	40.9%
2.5	24.2%	9.7%	0.6	27.6%
2.5	7.2%	2.9%	0.43	33.7%
3	14.4%	4.8%	0.59	27.9%
3	2.7%	0.9%	0.43	33.7%

Table 8. White and black rates of multiple suspensions in preschool and K-12, with measures of difference

Level	White Mult Susp Rate	Black Mult Susp Rate	B/W Ratio Susp	W/B Ratio No Susp	Abs Df (pp)	EES
Preschool	0.15%	0.67%	1.01	4.41	0.52	.49
K-12	2.23%	6.72%	1.05	3.01	4.49	.51

See Society 2014 at 15 re its Table 8 and Preschool Disparities subpage of Discipline Disparities page.

Table 9. Illustration on the EES From an Alternative Perspective (comparison of group's rate changes rather than comparisons of difference between groups before and after change)

Table	Subject	Group	Initial Adverse Rate	Final Adverse Rate	Fav Perc Increase	Adv Perc Decrease	EES
2	Test	AG	20.0%	5.0%	18.8%	75.0%	0.80
2	Test	DG	37.0%	13.0%	38.1%	64.9%	0.80
3	Poverty	White	10.8%	7.2%	4.0%	33.3%	0.22
3	Poverty	Black	24.7%	17.8%	9.2%	27.9%	0.24

BLUE and RED Columns show Corollary 1 previously mentioned (in terms of relative change rather than rate ratio).
EES is based on before and after rates. Since no meaningful change, EES should remain constant. That is, EES for 20% versus 37%, and for 5% versus 13%, is .50.
Compare with Tables on the Educational Disparities page and Table 7 of Society 2014 (Table 17 infra).

Interjection re subgroups

- Subgroup Effects subpage of SR explains why assuming that an intervention that reduces a baseline adverse outcome rate from 10% to 5% will cause a like 50% reduction in a baseline rate of 20% (i.e., to 10%) is not only incorrect but illogical. See also See also Comment on Hingorani BMJ 2013 and Illogical Premises, Illogical Premises II, and Inevitability of Interaction subpages SR.

- But one can, on the basis of the .36 EES difference reflected by the change from 10% to 5%, estimate that the intervention will reduce a 20% rate to approximately 11.5%.

NCHS Recognition of IR1

- In five official and unofficial documents between 2004 and 2009 (responding to Society 2000 and Chance 1994), NCHS recognized that determinations of whether health and healthcare disparities were increasing or decreasing would commonly turn on whether one examined relative differences in favorable outcome or relative differences in adverse outcomes.

- Key document: 2005 NCHS monograph "Methodological Issues in Measuring Health Disparities"

- Agency merely recommended that all disparities be analyzed in terms of relative differences in adverse outcomes. Has never addressed the implications of the fact that measures change as the prevalence of an outcome changes with respect to the utility of the measures.

- See Society 2014 at 4 to 9.

Healthy People 2010 Technical Appendix at A-8

"Those dichotomous objectives that are expressed in terms of favorable events or conditions are re-expressed using the adverse event or condition for the purpose of computing disparity [12 [sic],18,19], but they are not otherwise restated or changed."

13. Keppel KG, Pearcy JN, Klein RJ. Measuring progress in Healthy People 2010. Statistical Notes, no. 25. Hyattsville, MD: National Center for Health Statistics. September 2004.

18. Keppel KG, Pamuk E, Lynch J, et al. Methodological issues in measuring health disparities. National Center for Health Statistics. Vital Health Stat 2(141). 2005.

19. Keppel KG, Pearcy JN. Measuring relative disparities in terms of adverse outcomes. J Public Health Manag Pract 11(6). 2005.

Note: Few readers of the Technical Appendix would imagine that by measuring things like immunization disparities in terms of relative differences in no immunization one commonly reverses the direction of change over time, at times causing dramatic decreases to be dramatic increases (as in the Morita study in Table 9 infra).

CDC and AHRQ

- CDC usually measures disparities in terms of absolute differences between rates.
- AHRQ seeks to measure disparities in terms of relative differences in adverse outcomes, but does not invariably do so (see Table 4 of FCSM 2013 presentation).
- Neither CDC nor AHRQ has shown any awareness that measure change because prevalence changes or even that NCHS (an arm of CDC) has found that determinations of directions of change will commonly turn on which relative difference observer examines.
- Has the disparities research of NCHS, CDC, AHRQ (including the yearly National Healthcare Disparities Report) been of value?

Other treatments of IR1

- See generally Consensus subpage of Scanlan's Rule page.

- Mackenbach JP. The persistence of health inequalities in modern welfare states: The explanation of a paradox. Social Science and Medicine 2012;75:761-769.

  See Comment II of Marmot BMJ 2013; see also comment nos. 113, 79, 72, 70, 50.

- Lambert, P.J. and S. Subramanian. Disparities in socio-economic outcomes: some positive propositions and their normative implications. *Social Choice and Welfare* 2014;43(3):565-576.

- Lambert P, Subramanian S. Group inequalities and "Scanlan's Rule": Two apparent conundrums and how we might address them. Working Paper 84/2014, Madras School of Economics.

Table 10: Illustration Based on Morita (*Pediatrics* 2008) Data on Black and White Hepatitis-B Vaccination Rates Before and After School-Entry Vaccination Requirement

Grade	Year	Period	Wh Vac Rate	Bl Vac Rate	W/B Ratio Vac (Morita)	BW Ratio No Vac (NCHS)	Abs Df (PP) (CDC)	OR	EES
5	1996	Pre	8%	3%	2.67	1.05	5	2.81	47
5	1997	Post	46%	33%	1.39	1.24	13	1.73	34
9	1996	Pre	46%	32%	1.44	1.26	14	1.81	37
9	1997	Post	89%	84%	1.06	1.45	5	1.54	24

Table 11. Illustration from Harper et al. (CEBP 2009) Data on Differences in Mammography by Income (see Comment on Harper)

Year	High Inc Mam Rt	Low Inc Mam Rt	H/L Ratio Mam	L/H Ratio No Mam	Abs Df (pp)	OR	EES
1987	36.3%	17.2%	2.11	1.30	19	2.74	0.60
2004	77.4%	55.2%	1.40	1.98	22	2.78	0.62

Abstract: "In contrast, relative area-socioeconomic disparities in mammography use increased by 161%."

Text: "Whether a health outcome is defined in favorable or adverse terms (e.g., survival versus death) can affect the magnitude of measures of health disparity based on ratios (11, 12). Consistent with the Healthy People 2010 framework for comparing across outcomes (13), we measured all breast cancer outcomes in adverse terms."

Relative difference for mammography decreased 64% (111% to 4%); relative difference for no mammography increased by 227%. (3% to 98%)

.

Table 12. Illustration from Baker and Middleton (JECH 2003) Data on Differences in Mammography of Least and Most Deprived (see Mortality and Survival page)

Year	Lst Dpr Mam Rt	Mst Dpr Mam Rt	LD/MD Ratio Mam	MD/LD Ratio No Mam	Abs Df (pp)	Odds Ratio	EES
1991	84.1%	39.0%	2.15	3.83	45	8.26	1.27
1999	98.6%	76.0%	1.30	17.14	23	22.24	1.49

Authors would relied on relative differences in mammography rates to find a decreased disparity.

Harper et al. would find a **570% increase** in the disparity (from 283% to 1714%).

NCHS and AHRQ – would call these either a **1331 percentage point increase** (NCHS) or **1331% increase** (AHRQ), referring to the increase from (from 283% to 1714%).

CDC would call it a **22 percentage point increase**.

Table 13. Illustration from Albain (J Nat Cancer Inst 2009) Data on Survival Rates of White and Black Women for Various Types of Cancers, from Albains et al., with Disparities Measures

Type	Wh Surv Rate	Bl Surv Rate	W/B Ratio Surv	B/W Ratio Mort	Abs Df (pp) *	Odds Ratio	EES
premenopausal breast cancer	77%	68%	1.13	1.39	9	1.58	0.27
postmenopausal breast cancer	62%	52%	1.19	1.26	10	1.51	0.26
advanced ovarian cancer	17%	13%	1.31	1.05	4	1.37	0.18
advanced prostate cancer	9%	6%	1.50	1.03	3	1.55	0.21

Studies finding larger relative differences in survival for more survivable cancers (or among the young) are really about relative differences in mortality. See Mortality and Survival page Mortality/Survival Illustration subpage of Scanlan's Rule page.

Table 14. Rates of Births Attended by Skilled Measures for Highest and Lowest Quintiles in Columbia and Bangladesh, from WHO Handbook on Health Inequality Monitoring (2013), with Disparities Measures

Country	Highest Quintile Attend Rate	Lowest Quintile Attend Rate	H/L Ratio Attend	L/H Ratio No Attend	EES
Columbia	99.4%	83.7%	1.19	27.17	1.34
Bangladesh	50.6%	4.9%	10.33	1.93	1.67

WHO Handbook cites the 2005 NCHS monograph and seems to think it is following it. But relying on relative differences between rates finds largest disparity (BLUE) where NCHS would find smallest (RED), and vice versa, with starkly different interpretations

Spurious Contradictions

- **1. Escarce and McGuire APHA 2004**
  - racial differences in uncommon (increasing) procedures outcomes 1986-1997
  - found usually **decreasing relative differences** in receipt (but would have found usually increasing absolute differences)

- **2. Jha et al. NEJM 2005**
  - similar to no. 1 but for period 1992 to 2001
  - found **usually increasing absolute differences** (but would have found usually decreasing relative differences in receipt)

- **3. Trivedi et al. NEJM 2005**
  - examined racial differences in common (increasing) outcomes
  - found **usually decreasing absolute differences**

- **4. Le Cook et al. Med Care Res and Rev 2008**
  - titled "Measuring Trends in Racial/Ethnic Health Care Disparities"
  - relied on absolute differences in things it reported
  - Regarding studies 1 and 2 stated: "*The methods and data in [Jha et al.] were the same as Escarce and McGuire, except for the partial overlap in time periods. Assembly of a longer time series in Medicare would be necessary to reconcile the apparent differences in the findings of the two studies.*"

See Spurious Contradictions subpage of MHD.

Table 15: Varying Appraisals of the Comparative Degree of Employer Bias Using Different Measures of Disparities in Selection/Rejection Rates

Employer/ Setting	AG Sel Rate	DG Sel Rate	(1) AG/DG Ratio Selection	(2) DG/AG Ratio Rejection	(3) Abs Diff (pp)	(4) Odds Ratio
A	20.%	9.%	2.22 (1)	1.14 (4)	11 (4)	2.53 (1)
B	40.1%	22.7%	1.77 (2)	1.29 (3)	17(2)	2.29 (3)
C	59.9%	40.5%	1.48 (3)	1.48 (2)	19 (1)	2.19 (4)
D	90.%	78.2%	1.15 (4)	2.18 (1)	12 (3)	2.50 (2)

Approach 1 (relative favorable) (BLUE): A,B,C,D
Approach 2 (relative adverse) (RED): D,C,B,A (opposite of Approach 1)
Approach 3 (absolute difference) (GREEN): C,B,D,A
Approach 4 (odds ratio) (ORANGE): A,D,B,C (opposite of Approach 3)

- Is one employer more biased as to selection while another more biased as to rejection?
- Is one more biased in relative terms and another more biased in absolute terms?
- There can be only one reality as to the comparative ranking.

- Which is the correct ranking?
- As all rows are based on the same specifications as Table 2 and Figures 1 to 3 (EES = .5), there is no rational basis for distinguishing among them.
- Any measure that does distinguish among them is a flawed measure.

Table 16. Appraisals of the Differences in Outcome Disparities for AG and DG Applicants with Low and High Qualifications

Row #	Applicant Qualification	AG Sel Rate	DG Sel Rate	AG/DG Ratio Selection	DG/AG Ratio Rejection	Abs Diff (pp)	Odds Ratio
1	Very Low	20%	9.%	2.22	1.14	11	2.53
2	Low	40%	22.7%	1.77	1.29	17	2.29
3	High	59%	40.5%	1.48	1.48	19	2.19
4	Very High	90%	78.2%	1.15	2.18	12	2.50

Note: Some observers would read the smaller relative difference in selection rates (BLUE) among the highly qualified applicants (rows 3 and 4) as evidence that employers are less likely to rely on stereotypes when there are objective indicators of qualifications.

Table 17. Illustration of Contrasting Interpretations of Effects of Convictions on Callback Rates of Applicants by Race (based on Pager 2003)

Race	No Conviction (AG) CB Rt	Conviction (DG) CB Rt	AG/DG Ratio CB	DG/AG Ratio No CB	EES
White	34%	17%	2.00	1.26	0.54
Black	14%	5%	2.80	1.10	0.56

Note: This table reflect the alternative perspective (comparison of a factor's effects on different groups). Author drew inferences based on comparative size of relative differences in favorable outcomes (blue field). See the Criminal Record Effects subpage of SR for racial differences among those with and without criminal records and a later study with rather different results.

Table 18. Illustration of Problematic Nature of Representational Comparisons

DG Proportion of Pool	DG Proportion of Selections	AG/DG Ratio Selection
20%	10%	2.25
30%	20%	1.71
50%	30%	2.33
10%	5%	2.11
50%	25%	3.00

We cannot appraise the comparative likelihood that bias was involved because we cannot determine the actual selection rates. We need those to derive the EES.

.

Table 19. Effects of Prevalence of Outcome on Measures from IDEA Data Center Guide for Identifying "Significant Disproportionality" in Special Education (b5618a1)

(1) DG Prop Pool	AG Adv Rate	DG Adv Rate	(a) DG/AG Ratio Adv Rate	(b) Abs Df Btw Rates (pp)	*(2) DG Prop of Adv*	(c) Rel Df Bwt (1) and (2)	(d) Abs Df Btw (1) and (2)
20%	20%	36.7%	1.83	16.7	*31.4%*	57.2%	11.4
20%	10%	21.8%	2.18	11.8	*35.2%*	76.2%	15.2
20%	3%	8.4%	2.79	5.4	*41.1%*	105.6%	21.1
70%	20%	36.7%	1.83	16.7	*81.1%*	15.8%	11.1
70%	10%	21.8%	2.18	11.8	*83.6%*	19.4%	13.6
70%	3%	8.4%	2.79	5.4	*86.7%*	23.9%	16.7

See IDEA Data Center Disproportionality Guide subpage of Discipline Disparities page.

Table 20. Effects of DG Representation in Pool on Measures from IDEA Data Center Guide for Identifying "Significant Disproportionality" in Special Education (b5618a2)

(1) DG Prop Pool	AG Adv Rate	DG Adv Rate	(a) DG/AG Ratio Adv	(b) Abs Df Btw Rates (pp)	*(2) DG Prop of Adv*	(c) Rel Df Bwt (1) and (2)	(d) Abs Df Btw (1) and (2)
20%	10%	21.8%	2.18	11.8	*35.2%*	76.2%	15.24
30%	10%	21.8%	2.18	11.8	*48.2%*	60.9%	18.27
40%	10%	21.8%	2.18	11.8	*59.2%*	48.0%	19.20
50%	10%	21.8%	2.18	11.8	*68.5%*	37.1%	18.52
60%	10%	21.8%	2.18	11.8	*76.6%*	27.6%	16.56
70%	10%	21.8%	2.18	11.8	*83.6%*	19.4%	13.55
80%	10%	21.8%	2.18	11.8	*89.7%*	12.1%	9.70

Summary re Pay for Performance

- Reliance on absolute differences to measure healthcare disparities led to the perception in US (where increasing uncommon outcomes were examined) that P4P would tend to increase disparities and perception in UK (where increasing uncommon outcomes were examined) that P4P would tend to reduce disparities.
- Perception in US led Massachusetts to include a disparities element in its Medicaid P4P program, but to employ a measure that is more likely to increase than reduce disparities.

Table 21: Illustration Based on Werner et al. (*Circulation* 2005) Data on White and Black CABG Rates Before and After Implementation of CABG Report Card (see Comment on Werner)

Period	Wh Rt	Bl Rt	W/B Ratio CABG	RB/W Ratio No CABG	***Abs Df (pp)***	Odds Ratio	EES
1	3.6%	0.9%	4.00	1.03	***2.70***	4.11	0.58
2	8%	3%	2.67	1.05	***5.00***	2.81	0.48

Rather than find decreasing disparities according to the relative differences in receipt of CABG (as was probably the most common approach at the time), authors rely on absolute difference to find incentive program increases disparities. Study causes numerous researchers to recommend including disparities measure in pay-for-performance.

Table 22. Illustration of Changes in Absolute Differences over Time to Outcomes of Low (A) and High (B) Prevalence (Re Pay for Performance)

Outcome – Time	AG Fav Rt	DG Fav RT	Abs Df (pp)
A – Year One	20%	9%	11
A – Year Two	30%	15%	15
B – Year One	80%	63%	17
B – Year Two	90%	78%	12

Increases in low frequency favorable outcomes tend to increase absolute differences; improvements in high frequency favorable outcomes tend to increase absolute differences.

Table 23. Illustration of Absolute Differences at Low and High Performing Hospital as to Outcomes of Low (A) and High (B) Prevalence (Re Pay for Performance)

Hospital–Outcome	AG Fav Rt	DG Fav RT	Abs Df
Low Performing – A	20%	9%	11
High Performing – A	30%	15%	15
Low Performing – B	80%	63%	17
High Performing – B	90%	78%	12

Highlighted rows reflect situation of Massachusetts Medicaid pay for performance program. See page 21-24 of the Harvard Letter and Between Group Variance subpage of Measuring Health Disparities page.

Table 24. Varying Interpretations of Effects of Educational Improvements on Difference in Falling Below Basic and Reaching Advanced Level

Row No.	AG Fav Rate	DG Fav Rate	AG/DG Ratio Fav	DG/AG Ratio Adv	Abs Diff (pp)
1	10%	3.8%	2.67	1.07	0.06
2	20%	9.0%	2.22	1.14	0.11
3	80%	63.3%	1.26	1.83	0.17
4	90%	78.2%	1.15	2.17	0.12

Movement from Row 1 to Row 2 reflects increases in rates of reaching the advanced level; movement from Row 3 to Row 4 reflects increases in rates of reaching basic level.

Observers relying on absolute differences [GREEN] would tend to find (a) increase in former but (b) decrease in latter (approach in study discussed in Education Trust Glass Ceiling Study subpage of Educational Disparities page).

Observers relying on larger absolute difference [BLUE for 1 to 2; RED for 3 to 4] would tend to find (a) decrease in the former but (b) increase in the latter (as in study discussed in the McKinsey Achievement Gap Study subpage of the Educational Disparities page).

Table 25. Illustration of Effect on Standard Measures of Increase in Poverty Such as to Pull Into Poverty Everyone with Incomes Below 125 Percent of Poverty Line

Period	Perc of Pov Line	Prop AG Below	Prop DG Below	AG/DG Ratio Above	DG/AG Ratio Below	Abs Df (PP)	EES Ratio	EES
1	100	10.8%	24.7%	1.18	2.29	13.9	2.71	.55
2	125	14.9%	31.0%	1.23	2.08	16.1	2.57	.54

Common situation where observers relying on the larger relative differences (i.e., in poverty rates) would find decreasing disparity, while those relying on absolute differences (as is increasingly done) would find increasing disparities.

Table 26. Illustration of Change in Standard Measures of Increase in Poverty Such as to Pull Into Poverty Everyone with and Below 125 Percent of Poverty Line (alternative perspective)

Group	Initial Pov Rate	Final Pov Rate	Perc Fav Decrease	Perc Adv Increase	Abs Df (pp)	EES
White	10.8%	14.9%	4.6%	38.0%	4.1	0.20
Black	24.7%	31.0%	8.4%	25.5%	6.3	0.19

Same as prior table, but from the alternative perspective. Those relying on relative measures would say poverty increased more for whites, while those relying on absolute differences would say poverty increased more for blacks.

Table 27: Patterns of in Changes in Unemployment Rates by Race and Ethnicity (from 2011 Center for American Progress study)

Race	2007 Unempl Rate	2011 Unempl Rate	Perc Dec Employ	Perc Inc Unemploy	Abs Change (pp)	EES
Black	8.6%	15.8%	7.9%	83.7%	7.2	0.36
Hispanic	5.8%	12.9%	7.5%	122.4%	7.1	0.44
White	4.2%	8.7%	4.7%	107.1%	4.5	0.37

Minnesota Issues (1)

- Disparities by Subject subpage of Educational Disparities page discusses data in a study titled "Closing the Gap on Educational Disparities" by Daria Paul Dana of University of Minnesota at Mankato about proficiency rate disparities and discusses the patterns described here in that context.

- Subgroup Effects – Nonclinical subpage of the Scanlan's Rule page discusses a 2011 Minneapolis Urban League Study titled "Racial Disparities in Home Ownership," that, (a) in examining location effect on mortgage loan denial rates (the favorable outcome where whites had lower baseline rates) found a larger effect on whites than blacks, but, (b) in examining the location on home ownership (the favorable outcome on which blacks had lower rates than whites), found larger effects on blacks than whites. As generally discussed here, both patterns are to be expected in the circumstances. Compare discussion in Immunization Disparities of study that analyzed relative socioeconomic disparities (a) regarding receipt of any immunization in terms of relative differences in adverse outcomes and (b) regarding full immunization in terms of relative difference in favorable outcomes and discussion in McKinsey Achievement Gap Study subpage of the Educational Disparities of study that analyzed racial/ethnic differences (a) regarding basic proficiency in terms of relative differences in adverse outcomes and (b) regarding advanced proficiency in terms of relative differences in favorable outcomes. Neither study indicated awareness of reasons why general improvements would tend to increase (a) but reduce (b).
.

Minnesota Issues (2)

- See the "Mismeasure of Discrimination" (at 10 n.11) regarding a 2010 study by of housing issues in St. Paul (*Widening the Gap: How the Housing Crisis Deepened Racial Disparities in St. Paul and Hour to Fix it*), that focusing on relative differences in adverse outcomes or things that were functions of relative differences in adverse outcome reflected the mistaken impression that the housing crisis had increase those difference and that failed to recognize that the higher more vacant buildings were concentrated in poor neighborhoods the smaller tended to be both (a) the total number of vacant buildings and (b) the number of buildings in poor neighborhoods.

- See the Minneapolis Disparities and St. Paul Disparities subpages of the Discipline Disparities page regarding the way recent decreases in suspension rates in the cities were accompanied by increasing relative differences in suspension rates. Compare with California Disparities, Maryland Disparities, Los Angeles SWPBS, Denver Disparities, Beaverton OR Disparities, and Montgomery County MD Disparities subpages regarding similar patterns in the referenced jurisdictions.