Statistical Literacy: Scanlan's Paradox

Milo Schield, Augsburg University. Minneapolis, MN

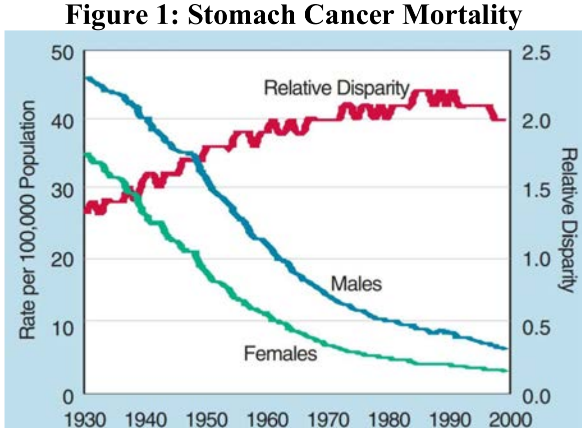
Abstract: When dealing with binary outcomes, two groups and two situations or times, a counter-intuitive phenomena emerges. As the percentages of one outcome decrease, the associated risk ratio and relative difference tend to increase when using the larger as the numerator. As the percentages of the complimentary outcome increase, the associated risk ratio and relative difference tend to decrease when using the larger as the numerator. Scanlan first identified this situation in 1987. Since then he has written numerous articles trying to increase awareness of these confusing effects. Others have named this result Scanlan's rule. Since this result is so surprising, I call it Scanlan's Paradox. Scanlan described it this way: "The rarer an outcome, the greater *tends* to be the relative difference in experiencing it and the smaller *tends* to be the relative difference in avoiding it." This paper presents and analyzes examples of the Scanlan Paradox. When some things improve, other things tend to worsen. This paper reviews Scanlan's 'tendency' and identifies the conditions under which Scanlan's Paradox will occur. Scanlan's call to find a better measure of disparity is strongly endorsed.

Keywords: statistical education, social statistics

1. Introduction

Group disparities are a hot topic. Despite improvement in reducing adverse outcomes, group disparities seem to be increasing. In Figure 1, stomach cancer mortality decreased for both men and women.

Figure 1: Stomach Cancer Mortality



Both rates decreases, but the risk ratio increased. Harper and Lynch (2006). In 1930, men were 1.4 times as likely to die from stomach cancer as women (2.2 times in 1990). As things got absolutely better (lower death rates) for men and women, things got worse for men relative to women. As the rates decreased, the risk ratio increased.

This unexpected result is a part of what I call Scanlan's paradox: a decrease in rates tends to increase their relative disparity.

1.1 Measuring Disparities:

Measuring disparities looks simple. The disparity in Figure 1 was just the risk ratio: the ratio of the rates. That is one of the three most common methods of comparison: subtraction (difference), division (ratio) or a combination (relative difference).

The difference is easy to calculate. As you can see from Figure 1, the difference must be less than the larger of the two rates. So as the larger rate decreases, the difference tends to decrease. Those arguing that the disparity is decreasing, may use the difference as evidence. But others recognize that the decreasing difference is strongly influenced by both groups approaching their limits. As such it is may not be a very good measure of disparity.

The simple ratio and the relative difference may seem equally useful. The simple ratio is often named the risk ratio (RR). In Figure 1, RR = Male Rate / Female Rate. The relative difference (RD) is (Male Rate – Female Rate)/Female Rate: RD = RR-1. Mathematically, they are very similar. But they have several important differences.

1. Their comparisons give different results. If men are 2.1 times as likely to die from stomach cancer as women, then men are 110% more likely to die from stomach cancer than women.
2. Their comparisons support different interpretations. Suppose the risk ratio was cut from RR=2 to RR = 1: a 50% reduction. But the relative difference is cut from RD=1 to RD=0: a 100% reduction. Either way the disparity is eliminated: RR = 1 or RD=0 means no disparity. Only the relative difference gives a mathematical result that aligns with the change in disparity.
3. Describing change in disparity is simpler with the relative difference. With the risk ratio (RR), the change in disparity must be measured as moving toward or away from unity (the goal). Whereas with the relative difference (RD), the change in disparity is readily described as increasing or decreasing since the goal is zero.
4. Comparing relative differences typically gives bigger ratios than comparing risk ratios. Suppose RR1 = 6 and RR2 = 2, so the ratio is three. But RD1 = 5 and RD2 = 1, so the ratio is five.

Describing the change in disparity is greatly simplified by always comparing the larger rate with the smaller. This way the risk ratio is always greater than one; the relative difference is always positive. So RD = RR-1 for RR>1; RD = 1-R for RR < 1. Scanlan used this convention.

Despite the excellent argument for using the relative difference, this paper uses risk ratio for several reasons. First, many technically-trained adults confuse "times more than" with "times as much as". Second, journalists are unlikely to form a comparison over time. They are more likely to quote the risk ratios before and after a change. Third, risk ratio is about seven times as common as relative difference in Google books.[1] Finally, the focus of Scanlan's paradox is not on the size of the change but on the opposing directions of the changes.

[1] Google Ngrams. Search "risk ratio, relative difference". https://books.google.com/ngrams

1.2 Mathematical vs. Evaluative

Generally speaking there are two distinct ways of describing disparities: mathematical and evaluative. Scanlan uses the mathematical. Journalist, social activists and politicians tend to use the evaluative. This paper will use both. The benefit of the evaluative is that it focuses attention on the impact of the mathematical relationships. There are two problems with the evaluative. First, it is too easy to include good and bad in the definitions of the terms. If passing an exam is good, then maybe we should change things so that more people pass. What becomes desirable or favorable becomes a political or social necessity or requirement. Second, the term 'disparity' is used as though the measure of disparity was obvious. One point made by Scanlan and this paper is that measuring disparity in this situation is not obvious.

1.2 Scanlan's Paradox: General Conditions

Scanlan's paradox emerges under a very specific situation. It must involve binary outcomes (succeed or fail), it must involve two different times (before or after) or situations (two criteria), and it must involve at least two groups. Comparing these group percentages at a single time or situation is as a risk ratio or a relative difference is straightforward. Comparing these risk ratios or relative differences of time is straightforward. But as will be shown, measuring disparities is more complex than comparing numbers.

The paradox shown in Figure 1 is just a part of Scanlan's paradox. To see the complete paradox, consider the following examples.

2. More Examples of Scanlan's Paradox

There are two approaches involved in changing group disparities: the reality changes or the rule for the outcome changes. Observing a change in reality was demonstrated in the Introduction. Changing the rule is demonstrated here.

There are two different kinds of situations involved in measuring group disparities: outcomes are ordered on a quantitative scale (e.g., a Normal distribution) or outcomes are multinomial (different activities). Both are demonstrated here.

2.1 Change the Rule: Reduce Group Disparities in Failing Exam

Consider two groups, A and B, whose scores are both normal distributions with standard deviations of one. The center of B is 0.5 standard deviations below the center of A. [2]

Evaluative: Passing the exam is at the right (higher scores); failing the exam at the left (lower scores). In Figure 2, the cutoff for passing is the center of the A distribution. Given this cutoff, 50% of those in the A distribution fail. Of those in the B distribution, 70% fail: the 50% below the center of the B distribution and the 20% between the centers of the two distributions. The B/A risk ratio for failure is 1.4 (70% / 50%).

[2] Scanlan (1994) used this example.

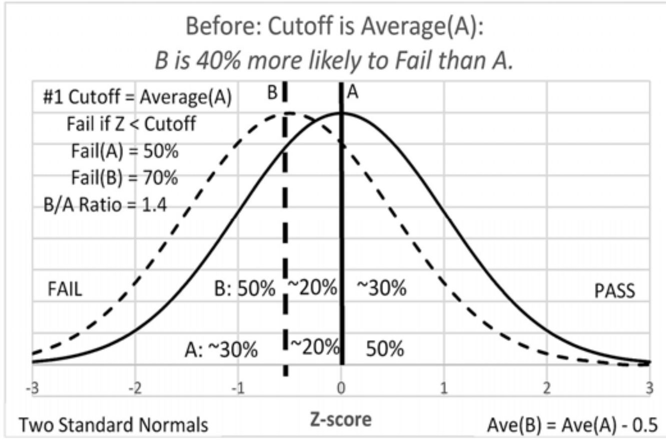


Figure 2: Two Standard Normal Distributions: Before Lowering the Cutoff

In Figure 3, the cutoff for passing is lowered to the center of the B distribution. Now just 30% of those in the A distribution fail (50% of those in the B distribution). The new B/A risk ratio for failure is 1.67 (50% / 30%).

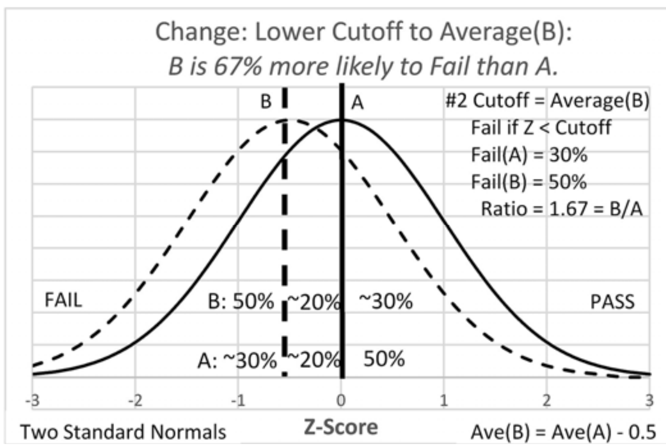


Figure 3: Two Standard Normal Distributions: After Lowering the Cutoff

In this case, lowering the minimum rate to pass increased the B/A disparity ratio for failures. Before lowering the cutoff, those in B were 50% more likely to fail than were those in A. After lowering the cutoff for passing, those in B were now 67% more likely to fail than were those in A.

Seeing this graphical relationship in a tabular form may be helpful. See Table 1.

Table 1: The numbers on the left side apply to those that failed. The top two rows and the first two columns contain the failure rates for those in the A and B distributions. In the first column, those failing are below the center of the A distribution: 50% of those in A (70% of those in B) with a B/A risk ratio of 1.4 for failures. In the second column, those failing are below the center of the B distribution: 30% of those in A (50% of those in B) with a B/A risk ratio of 1.7 for failures.

The right two columns contain four new measures of change. In the top row, the failure rate for those in A dropped (improved) by 40%: from 50% to 30%. In the second row, the failure rate of those in B dropped (improved) by 29%: from 70% to 50%. In the third row, the B/A disparity ratio for failures increased (worsened) by 19%: from 1.4 to 1.67. In the fourth row

In the bottom row, assuming an equal A-B mixture, the percentage of those failing who are B increased above 50%: a worsening for B relative to the 50-50 goal.[3] This increasing percentage of B among the failures is a second part to Scanlan's paradox.

The decrease in the cutoff makes those in A and B both better off – for failures taken absolutely. But lowering the cutoff made those in B worse off – relative to A.

Table 1: Pass and Fail Rates for Two Normal Distributions

Fail %	Z < Ave(A)	Z < Ave(B)	Change	Result
A	50%	30%	-40%	A better
B	70%	50%	-29%	B better
Ratio B/A	1.4	1.7	19%	**B/A worse**
P(B|Fail)*	58%	63%	7%	**B worse**

Pass %	Z > Ave(A)	Z > Ave(B)	Change	Result
A	50%	70%	40%	A better
B	30%	50%	67%	B better
Ratio A/B	1.7	1.4	-16%	**A/B better**
P(B|Pass)*	38%	42%	11%	**B better**

* Assumes 50-50 split in A and B

Table 1: The numbers on the right side contains the pass rates in the top two rows and the first two columns. These pass rates are just the complements of the failure rates.

Since the pass rate for B is less than that for A, the A/B risk ratios are both greater than one. The A/B ratio decreased by 16% moving downward from above toward unity. Finally (assuming equal sized groups), the percentage of those passing who are B increased toward 50%: an improvement for B. On this basis, these four changes in pass rates are improvements.

What might one conclude from these pass rates. As we lower the cutoff for passing, the disparity ratio for passing moves toward unity. So, if we want a disparity ratio to approach unity, we should lower the pass ratio. It seems to make sense.

But what happens when we view the change on the failure side? The opposite happens: the disparity ratio for failure moves further from unity.

If moving the disparity ratios toward unity is the goal, then the movement for success is good while the movement for failure is bad.

This particular example has one key weakness. In this case, the absolute differences (20 percentage points) are the same on both sides, but that is coincidental.

Scanlan (2000) summarized this tendency mathematically as follows: "the less prevalent a condition, the greater the disparity in experiencing the condition. A corollary is that the less prevalent the condition, the larger will be the proportion of those experiencing the condition [that is] comprised by the more susceptible group."

Since Scanlan's paradox involves the directions of the change in the size of a percentage difference, the magnitude of the change is incidental: the direction of the change in magnitude is key.

[3] The direction of change in P(B|Fail) is the same regardless of the A-B mixture. See Appendix E

2.2 Change the Rule: Reduce Racial Disparities in School Suspensions:

Evaluative: Racial disparities may be a sign of systemic racism. If a racial disparity is due to racism, then eliminating that disparity is a reasonable goal for those trying to eliminate racism.

School suspensions were six times as likely for blacks as for whites in the St. Paul public schools.[4] In Minnesota, a third of all school exclusions are for minor incidents: talking back, eye rolling or swearing. Students with disabilities make up 14% of all K-12 students; 43% of suspensions and expulsions. If these actions or conditions were more common among blacks than among whites, then eliminating these as a basis for suspension was expected to lower the black-white disparity ratio for suspensions.

Change the rule: The St. Paul teachers "took racial equity training, the district narrowed the types of behaviors that were to result in suspension, and principals were instructed to keep kids in class when possible."

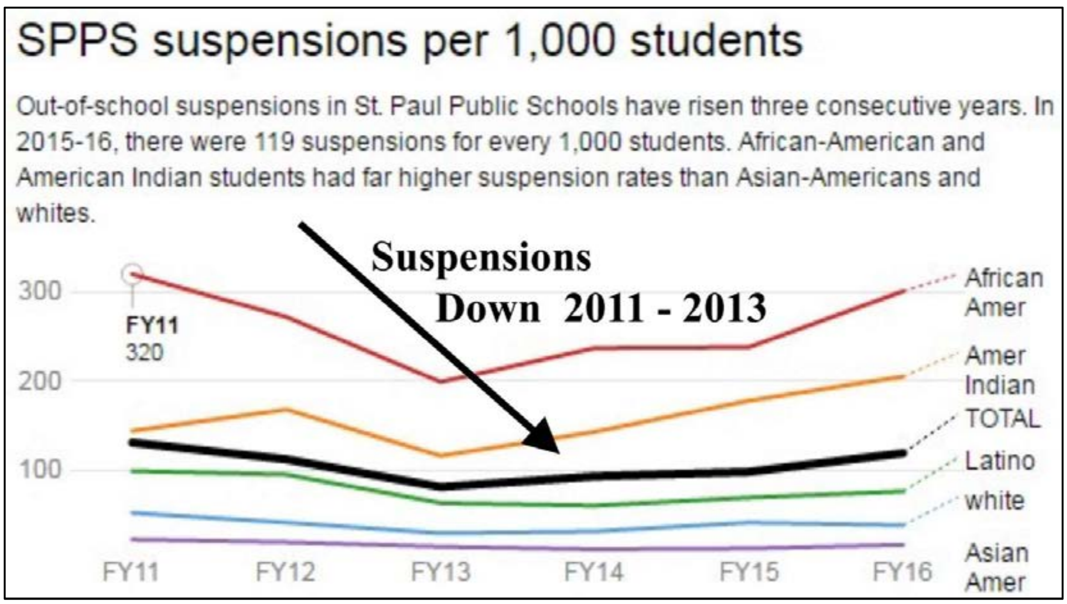


Figure 4: St. Paul Public School Suspensions

The results? Good news! Suspension rates were cut: blacks cut 37%; whites cut 44%.

Mission accomplished? Not quite!

The racial disparity ratio for suspensions increased! The black-white ratio of suspension rates increased: from 6.2 to 7.6! Now blacks are almost 8 times as likely to be suspended as whites. The disparity ratios increased for Latinos and American Indians as well.

[4] www.twincities.com/2018/06/29/st-paul-schools-to-scrutinize-student-suspensions-under-human-rights-agreement/

Table 2: Disparity Ratios of Suspension Rates Relative to Whites

	FY11	FY12	FY13	FY14	FY15
Latino	1.9	**2.3**	**2.2**	1.9	1.7
American Indian	2.8	**4.1**	**4.0**	**4.6**	4.3
Afro-American	6.2	**6.6**	**6.9**	**7.6**	5.8
Afro-American	Bad	Worse	Worse	Worse	Better

This increase is "journalistically significant". It makes headlines. It is politically and socially significant. Superintendents and principals can be fired for these negative results.

Evaluative summary: Decreasing the suspension rates for both blacks and whites ended up increasing the black-white disparity ratio for suspensions. Once again, trying to improve things by decreasing the suspension rates for blacks and whites ended up making things worse for blacks – relative to whites.

2.3 Other Examples
For more examples of Scanlan's paradox, see Appendix A.

3. James Scanlan: His Examples and His Descriptions

3.1 James Scanlan

James Scanlan is an Attorney at Law in Washington DC "specializing in the use of statistics with respect to employment discrimination litigation and compliance." The following is his mathematical description of what I call Scanlan's paradox.

In a great many places since 1987, Scanlan had described the pattern that he maintains is inherent in most risk distributions whereby *the rarer an outcome the greater tends to be the relative differences in experiencing it and the smaller tends to be the relative difference in avoiding it*. Usually, he has done in so in explaining that a belief reflected in much social science, and underlying many civil rights enforcement policies, that reducing an adverse outcome would be expected to reduce relative racial and other demographic differences in rates of experiencing it is incorrect.

That is, he maintains, while reducing an outcome tends to reduce relative differences in the corresponding opposite outcome, it tends to increase relative differences in the outcome itself. Scanlan maintains that there are many pernicious consequences of leading the public to believe that reducing adverse outcomes will tend to reduce measures of racial or other demographic differences when the policies in fact tend to increase those measures.

Since 2006, Scanlan has been explaining the way that absolute differences and odds ratios tend to change solely because the overall prevalence of an outcome changes, maintaining both that there is no value in employing standard measures to quantify demographic differences without consideration of the way the measure tends to change solely because the overall prevalence of an outcome changes and that a sound measure of a demographic difference reflected by the two groups' outcome rates must remain unchanged when there occurs a general change in the overall prevalence of an outcome such as that effected by the lowering of a test cutoff.

In a number of places since 2008, Scanlan has promoted probit d' as a measure meeting that criterion, while acknowledging its limitations and recommending approach along the same lines while informed by a sound understanding of the actual shapes of the underlying distributions. See especially "Race and Mortality Revisited" (Scanlan, 2014) at 337. See especially Scanlan (2014) at 337. See also Scanlan (2016a, 2016b).

Scanlan (1987) identified the elements of Scanlan's paradox. Scanlan is a prolific author with over 50 articles on this subject. His thinking goes far beyond what is presented in this paper. For more details see his website: www.jpscanlan.com. In particular, review Scanlan (2006, 2012, 2014, 2016a and 2020).

3.2 Scanlan Examples

The following are quotes from Scanlan (2020) and are based on what the data show:

a. Income and credit score data (showing that the lower an income or credit score requirement, the greater tend to be relative racial differences in failure to meet the requirement while the smaller tend to be relative racial differences in meeting it)

b. Life tables (showing that the higher the age, and thus the lower are overall rates of reaching it [a given age], the greater tend to be relative racial and gender differences in failure to reach it, while the smaller tend to be relative racial and gender differences in reaching it)

c. NHANES data on systolic blood pressure and folate level (showing that general improvement in control of systolic blood pressure and general increases in folate tend to increase relative differences in the adverse outcomes (hypertension, low folate) while reducing relative differences in the corresponding favorable outcomes (avoiding hypertension, adequate folate)

d. Framingham Study data (showing how improving risk profiles in a way that generally reduces heart attack risks tends to increase relative gender differences in heart attack risk while reducing relative gender differences in risks of avoiding heart attacks);

e. Health numeracy data (showing that the lower the health numeracy level, the greater tend to be relative differences between rates at which insured and uninsured persons fail to reach it, while the smaller tend to be relative differences between the rates at which they reach it.)

f. Literacy data (showing that the lower the level of reading proficiency, the larger the relative racial difference in failing to reach the level and the smaller the relative racial difference in reaching it)

g. Truancy data (showing that greater the level of truancy, the larger is the relative racial difference in rates of reaching the level)

h. California prison data (showing that the greater number of convictions of incarcerated persons, the larger is the relative racial difference in reaching or exceeding the number and the smaller is the relative racial difference in rate of failure to reach the level).

The following are from other sources:

i. As mortality declines, relative differences in survival tend to decrease while relative differences in mortality tend to increase.

j. As health-care receipt rates increase, relative differences in receipt tend to decrease while relative differences in non-receipt tend to increase.

k. Lowering credit score requirements tends to reduce relative differences in applications accepted, while increasing relative differences in applications rejected.

l. As hiring and promotion rates increase, relative differences in rate of being hired and promoted tend to decrease, while relative differences in not being hired or promoted tend to decrease.

The phrase 'Scanlan's rule' was first used by Bauld, Day and Judge (2008).

3.3 Scanlan's statements describing these situations:

Scanlan has used various expressions to describe these effects and the associated paradox. These expressions may be difficult to understand on first reading or hearing.

Scanlan (2005):

> "the rarer an outcome, the greater tends to be the relative (percentage) difference between the rates at which advantaged and disadvantaged groups experience the outcome and the smaller tends to be the relative difference between rates at which such groups avoid the outcome."

Scanlan (2017a):

> "Reducing an outcome and thereby increasingly restricting it to those most susceptible to it, while tending to reduce relative differences in rates of avoiding the outcome (i.e., experiencing the opposite outcome), will tend to increase relative differences in the outcome itself; correspondingly, reducing the outcome, while tending to increase the proportions groups more susceptible to the outcome make up of persons avoiding the outcome, will tend also to increase the proportions such groups make up of persons experiencing the outcome itself."

Scanlan (2020):

> "the rarer an outcome the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative difference in avoiding it (or, more precisely put, the more the outcome is restricted to the most susceptible part of the overall population the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative difference in avoiding it." "It should be kept in mind that it does not matter whether the issue is discussed in terms of a favorable outcome or the corresponding adverse outcome or whether the focus is on whether the outcome is decreasing or increasing."

4. Conditions for Scanlan's Paradox

Scanlan's use of "tends" motivates statisticians to look for the conditions under which his paradox occurs. Despite Scanlan's sustained and prodigious output, only one has focused on identifying the conditions.

Keppel et al (2005) reviewed the methodology for generating health statistics and presented an example of Scanlan's Paradox. Unfortunately, they give no basis for any particular choice. See Appendix H for more details.

Thomas and Hettmansperger (2017) provide a distribution-based analysis of Scanlan's paradox thatias most helpful. Consider their abstract:

> Abstract: Risk ratios are distribution function tail ratios and are widely used in health disparities research. Let A and D denote advantaged and disadvantaged populations with cdfs FA(x) and FD(x) respectively, FA(x) ≤ FD(x). Consider a selection setting where those selected have x > c a critical value. Scanlan observed in empirical data that as c is lowered the failure ratio FR(c) = FD(c)/FA(c) and success ratio SR(c) = [1 – FD(c)] / [1 – FA(c)] can both be increasing with decreasing c, a surprising result Scanlan calls Heuristic Rule X (HRX).

Figure 5 clearly shows the results for Scanlan's shifted Normal example. As the cutoff, x, decreases, the Success Ratio (SR) and its inverse (1/SR) both approach unity (better), while the Failure Ratio (FR) diverges from unity (worse). The condition, FA(x) ≤ FD(x) for all x, is shown to be sufficient to generate the first part of Scanlan's paradox. But it may not be necessary.

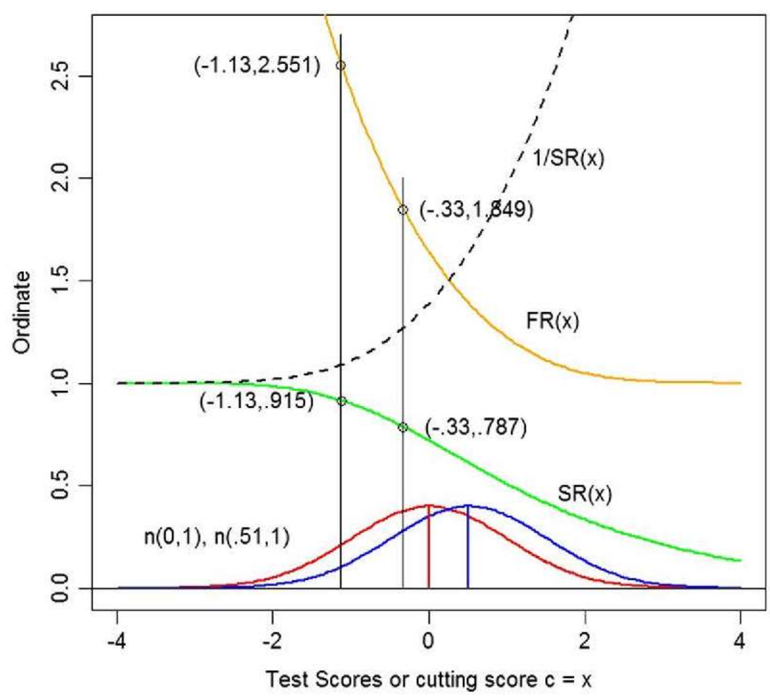


Figure 5: Two Shifted Normal Distributions versus the cutoff (c=x)

5. General Conditions for Scanlan's Paradox

The general condition for Scanlan's paradox must include continuous distributions and multinomial groups. Ideally, it will not include any reference to the distribution or the multinomial groups.

The general condition is really quite simple. In order to keep a ratio of two values unchanged, one must multiply or divide both by the same constant. If one divides the denominator by more than the numerator, a positive ratio will increase. Here is a summary of the general conditions for two of the three parts of the Scanlan's paradox.

Suppose there are just two outcomes (favorable and adverse), two times (situations) where the rates change, and two groups (Advantaged and Disadvantaged) where the disadvantaged have a higher fraction who experience the adverse outcome. If the percentage who experience the adverse outcome decreases for both Advantaged and Disadvantaged, and if the relative decrease is greater for the Advantaged than for the Disadvantaged, then one side of Scanlan's paradox will occur:

1. the Disadvantaged/Advantaged ratio of adverse rates will increase (a worsening),
2. among those with adverse outcomes, the percentage who are disadvantaged will increase (a worsening)

These conditions are not sufficient for the risk ratio of the favorable outcome to decrease. Consider this example.

In the top panel, the Before failure rates are cut equally: by 25%. RR is unchanged among the fails, and decreases among the pass. In the second panel, the Advantaged rate is cut by 35%: more than the 25% cut for the Disadvantaged rate. The RR for adverse (fail) increases and the RR for favorable (pass) decreases. This illustrates Scanlan's tendency.

Table 3: RR(Pass) varies with Size of Cuts in Adverse Rate

Fail	Before	After	Change	**Pass**	Before	After	Change
Adv	10%	**7.5%**	**-25%**	Adv	90%	92.5%	3%
Dis	20%	15%	-25%	Dis	80%	85%	6%
RR(Fail)	2.00	2.00	0%	RR(Pass)	**1.13**	**1.09**	-3%

Fail	Before	After	Change	**Pass**	Before	After	Change
Adv	10%	**6.5%**	**-35%**	Adv	90%	93.5%	4%
Dis	20%	15%	-25%	Dis	80%	85%	6%
RR(Fail)	2.00	2.31	15%	RR(Pass)	**1.13**	**1.10**	-2%

Fail	Before	After	Change	**Pass**	Before	After	Change
Adv	10%	**4.371%**	**-56%**	Adv	90%	95.6%	**6.25%**
Dis	20%	15%	-25%	Dis	80%	85%	**6.25%**
RR(Fail)	2.00	3.43	72%	RR(Pass)	**1.125**	**1.125**	0%

Fail	Before	After	Change	**Pass**	Before	After	Change
Adv	10%	**2.5%**	**-75%**	Adv	90%	97.5%	8%
Dis	20%	15%	-25%	Dis	80%	85%	6%
RR(Fail)	2.00	6.0	200%	RR(Pass)	**1.13**	**1.15**	2%

In the third panel, the fail rate for the advantaged is cut even more (56%). The RR for adverse (fail) increases as before, but now the RR for favorable (pass) is unchanged. Note that this occurs when the percent change in RR(pass) for both groups are equal. In the bottom panel, the fail rate for the advantaged is cut even more (75%). The RR for adverse (fail) increases even more, but now the RR for favorable increases.

If the percentage changes in rates are equal, then the risk ratios remain unchanged. If the risk ratios are formed using the larger as the numerator, then (a) if the percentage reduction is bigger in the lower rate, then the risk ratio increases, and (b) if the percentage increase is smaller in the higher rate, then the risk ratio decreases.

6. Appendices

Appendix A presents more examples of Scanlan's paradox. Appendix B summarizes Scanlan's earliest statement of what this paper calls Scanlan's paradox. Appendix C illustrates the conditions for Scanlan's paradox. Appendix D derives the mathematical conditions for the adverse disparity ratio to remain unchanged, to increase or to decrease. Appendix E shows that whatever happens to the percentage who are B assuming equal-sized groups, happens to a population with any mixture of A and B. Appendix F shows that as the B/A ratio increases, so does the share of the adverse who are B. Appendix G derives the conditions under which the relative risk for the increasing outcomes is unchanged. Appendix H presents an example of Scanlan's paradox that was present in the CDC document authored by Keppel et al (2005). Appendix I presents Scanlan's paradox in terms of blocks as an alternative to the percentage table presentation. Appendix J reviews the properties of the Odds ratio. Appendix K reviews the properties of Probit and Logit. Appendix L examines comparisons using Probit differences and Odds ratios.

7. Better Measure of Disparity

Scanlan has searched for a better measure of disparity: one that is less dependent on the small adverse rates. One might consider the Odds ratios. See Appendix J. But, for two small percentages, the Odds ratio approaches the ratio of the two percentages. As such it is extremely sensitive to these small percentages. Thus, it may add little.

Scanlan considered a Probit generated difference in standardized means. This comparison of means seemingly avoids the influence of small percentages. But it applies a particular distribution to the data. There is no reason to presume that the actual distributions will match the Probit distribution. Therefore, this model may not apply in many situations.

Scanlan (2017b) questioned the typical measures of disparity this way:

> "It is also important to understand that an increase in the relative difference in the adverse outcome does not mean that a disparity has increased in some meaningful sense any more than the reduction in the relative difference in the favorable outcome means that a disparity has decreased in a meaningful sense. Rather, the problem is that neither relative difference is a useful indicator of the strength of the forces causing the outcome rates of two groups to differ (or, as we might otherwise put it, the size of the difference in the circumstances of two groups reflected by their outcome rates). That is quite important to recognize as we endeavor to understand the causes of disparities and determine whether they are growing larger or smaller over time or are larger in one setting than another."

If using ratios to measure disparities can generate a problem, one possibility is to abandon that approach and return to percentage point differences. Looking for the perfect solution may have become the enemy of the good.

Suppose there is no good measure of the overall change in disparity. Perhaps we are left by saying that neither the decrease in disparity for those with favorable outcomes nor the increase in disparity for those with adverse outcomes is the entire story. If so, then the focus may need to shift. Instead of trying to reduce disparities by changing the standards, the focus should shift to reducing disparities by changing the difference in distributions or by seeing if the disparities are related to other factors than discrimination, sexism, etc.

8. Recommended Actions

Scanlan (2015) urged the American Statistical Association

> (a) "to form a committee to explore the ways analyses by statisticians and others of demographic and other differences in outcome rates are fatally undermined as a result of the failure to recognize patterns by which standard measures of differences between outcome rates tend to be systematically affected by the frequency of an outcome," and
> (b) "to formally advise arms of the United States government that … reducing the frequency of an outcome tends to increase [not reduce] relative differences in rates of experiencing the outcome."

Further research is needed to understand the condition under which the risk ratio for avoiding risk is unchanged. See Appendix G.

9. Conclusion

We can now state the conditions for Scanlan's paradox more precisely.

> If the prevalence of one of the binary outcomes decreases for both groups, if the more susceptible rate is the rate that is largest in experiencing this decrease, if risk ratios are formed using the larger rate as the numerator, and if the percentage reduction is larger in the less susceptible rate than in the more susceptible rate, then (a) the larger will become the risk ratio in experiencing the decrease and (b) the smaller will become the risk ratio in avoiding the decrease (provided the percentage increase in the less susceptible rate is smaller than that in the more susceptible rate).

Scanlan's paradox is important. Scanlan's call for further analysis by concerned statisticians and his request that the ASA advise governments on this matter should be supported.

Any one dealing with group disparities involving binary outcomes should understand Scanlan's paradox. As adverse outcomes decrease, more instances of Scanlan's paradox are to be expected.

Statistical educators should be aware of the Scanlan paradox and be prepared to teach it to students who will use social statistics. By doing this, teacher may influence the reporting by journalists and the choice of goals by policy makers.

10. Acknowledgment

Thanks to James Scanlan for identifying the existence and nature of this phenomena and for his unrelenting efforts in showing how it applied in a wide variety of circumstances.

11. REFERENCES

Bauld, Linda, Patricia Day and Ken Judge (2008). Off Target: A Critical Review of Setting Goals for Reducing Health Inequalities in the United Kingdom. *International Journal of Health Services*. https://doi.org/10.2190/HS.38.3.d

Harper, Sam and John Lunch (2006). Methods for Measuring Cancer Disparities: Using Data Relevant to Healthy People 2010 Cancer-Related Objectives. Center for Social Epidemiology and Population Health. University of Michigan. Copy at https://seer.cancer.gov/archive/publications/disparities/measuring_disparities.pdf

Keppel, Kenneth, et al (2005). Methodological issues in measuring health disparities. National Center for Health Statistics. *Vital Health Stat* 2(141). 2005. Copy at https://stacks.cdc.gov/view/cdc/6654 or https://pubmed.ncbi.nlm.nih.gov/16032956/.

Scanlan, James (1987). The 'Feminization of Poverty; is Misunderstood. *Cleveland Plain Dealer*. Nov 11, 1987. Reprinted as the "The 'Feminization of Poverty' Issue. Copy at http://www.jpscanlan.com/images/Poverty_and_Women.pdf

Scanlan, James (1994). 'Divining difference'. *Chance* 1994, 7(4):38-39, 45. http://jpscanlan.com/images/Divining_Difference.pdf

Scanlan, James (2000). "Race and Mortality," Society (Jan./Feb. 2000) (reprinted in Current, Feb. 2000). http://www.jpscanlan.com/images/Race_and_Mortality.pdf

Scanlan, James (2006). 'Can We Actually Measure Health Disparities?' *Chance.* www.jpscanlan.com/images/Can_We_Actually_Measure_Health_Disparities.pdf

Scanlan, James (2012): Misunderstanding Statistics Leads to Misguided Law Enforcement. *Amstat News*. magazine.amstat.org/blog/2012/12/01/misguided-law-enforcement

Scanlan, James (2014). 'Race and Mortality Revisited'. *Society* magazine. http://jpscanlan.com/images/Race_and_Mortality_Revisited.pdf. See also at http://link.springer.com/article/10.1007%2Fs12115-014-9790-1#page-1

Scanlan, James (2015). Letter to American Statistical Association. http://jpscanlan.com/images/Letter_to_American_Statistical_Association_Oct._8,_2015_.pdf

Scanlan, James (2016a): Mismeasure of Health Disparities. *Journal of Public Health Management Practice*, 2016, 22(4), 415–419. Copy at www.jpscanlan.com/images/The_Mismeasure_of_Health_Disparities_JPHMP_2016_.pdf, https://journals.lww.com/jphmp/Fulltext/2016/07000/The_Mismeasure_of_Health_Disparities.14.aspx

Scanlan, James (2016b). Comments of James Scanlan for Commission on Evidence-Based Policymaking https://www.regulations.gov/document?D=USBC-2016-0003-0135

Scanlan, James (2017a). United States Exports Its Most Profound Ignorance About Racial Disparities to the United Kingdom," *Federalist Society Blog* (Nov. 2, 2017) https://fedsoc.org/commentary/blog-posts/united-states-exports-its-most-profound-ignorance-about-racial-disparities-to-the-united-kingdom

Scanlan, James (2017b). "Measuring Discipline Disparities," Written testimony for U.S. Commission on Civil Rights Briefing "The School to Prison Pipeline: The Intersection of Students of Color with Disabilities" (Dec. 8, 2017). Copy at http://jpscanlan.com/images/Measuring_Discipline_Disparities_.pdf

Scanlan, James (2020). Obesity Illustration. Reference page with many illustrations. http://jpscanlan.com/scanlansrule/obesityillusration.html

Thomas, H., Hettmansperger, T. Risk ratios and Scanlan's HRX. *J Stat Distrib* App 4, 27 (2017). https://doi.org/10.1186/s40488-017-0071.

TABLE OF CONTENTS FOR THE APPENDICES:

Appendix A: More Examples of Scanlan's Paradox

Change the Reality: Reduce Disparities in Smoking Rates during Pregnancy

In 2004, Scotland tried (1) to reduce smoking during pregnancy and (2) to reduce the disparity between those in affluent versus deprived geographic areas. They did this by introducing educational programs and support groups. The reduction in smoking during pregnancy was significant: 8% in the most deprived areas, 17% in the most affluent. See Bauld et al (2008) for details. The 13.5% in Table 4 was estimated from the figures in Bauld.

Table 4: Percentage of Pregnant Women who Smoked during Pregnancy

Smoke	2003	2004	% change	Change
Disadvantaged	35.5%	32.5%	-8.38%	Better
Advantaged	**13.5%**	11.2%	-16.91%	Better
Disadv/Adv	2.63	2.90	10.27%	**Worse**

Unfortunately, this greater reduction among those in the affluent areas meant that the risk ratio widened by 10.3%. Mathematically: this is an example of Scanlan's paradox. Evaluatively: making things better for each group (advantaged and disadvantaged) made things worse for the disadvantaged – relative to the advantaged.

Reality Changes: Infant Mortality Rate get Lower

Consider infant mortality. This includes all deaths of newborn children during their first year after birth. Infant mortality and infant survival are complements of live births during the first year after birth. Infant mortality is different from childhood mortality. Childhood mortality includes all deaths of newborn children during the first five years after birth.

The right side of Table 5 shows that infant survival rates were higher for both whites and black in 1997 than they were in 1983. On this basis, both groups were better off over time.

In comparing infant survival rates over time, the black-white ratio moved from 0.990 to 0.992; the white-black ratio moved from 1.010 to 1.009. Both ratios moved closer to one so the black-white disparity in infant survival decreased. Blacks improved relative to whites in infant survival. Assuming a 50-50 mixture of whites and blacks, the percentage of infant survivors who are black increased toward 50%: an improvement for blacks.

In terms of infant survival over time, blacks were better off both absolutely and relative to whites as a percentage of those among the survivors.

Table 5: Infant Mortality and Survival: 1997 vs 1983

Mortality Rate (Deaths per 1,000 live births)

US Infants	1983	1997	Change	Result
White	9.7	6.0	-38%	W better
Black	19.2	14.2	-26%	B better
Ratio B/W	2.0	2.4	20%	**B worse**

Survival Rate (Survivors per 1,000 live births)

US Infants	1983	1997	Change	Result
White	990.3	994.0	0.4%	W better
Black	980.8	985.8	0.5%	B better
Ratio W/B	1.010	1.008	-0.1%	B better

Scanlan (2000). Race and Mortality. Society 37(2):29-35

The left side of Table 5 shows that infant mortality rates were lower for both whites and blacks in 1997 than they were in 1983. Both groups were better off.

On the right side of Table 5, the White-Black risk ratio decreased. If unity is the goal, then blacks were better off relative to whites in terms of survival.

On the left side of Table 5, the black-white ratio of infant mortality increased from 2.0 to 2.4. If unity is the goal, then the black-white disparity ratios in infant mortality worsened.

So relative to whites over time, blacks were better off in infant survival, but worse off in infant mortality. To repeat, making some things better can make other things worse.

Lower the Maximum Poverty Level: Change the Rule

"The Census Bureau determines poverty status by using an official poverty measure (OPM) that compares pre-tax cash income against a threshold that is set at three times the cost of a minimum food diet in 1963 and adjusted for family size." [5]

This example compares the results using two different poverty levels for the same population. The first is at 100% of the official poverty level (the second is at 50%),

In Table 6, the right side shows that poverty-avoided rates were higher for whites and black at the 50% level than at the 100% level. On this basis, both groups were better off.

The black-white ratio of poverty-avoided rates moved from 0.76 to 0.89; the white-black ratio of poverty-avoided ratio moved from 1.31 to 1.1 2. Both ratios moved closer to one so the black-white disparity improved. In terms of poverty -avoided, blacks improved relative to whites. Assuming a 50-50 black-white mixture, the percentage of those avoiding poverty who were black increased toward 50%: an improvement for blacks.

In terms of poverty-avoided under the new lower standard, blacks were better off absolutely and relative to whites, and as a percentage of poverty avoiders who were black.

Table 6: Family Poverty: Lower from 100% of the Poverty Level to 50%

Poverty Sustained (Income below ↓ Poverty Line)

US 1990	$<100%*	$<50%**	Change	Result
White	11%	4%	-64%	W better
Black	32%	14%	-55%	B better
Ratio B/W	3.0	3.7	24%	**B worse**

* Income below 100% (** 50%) of the poverty line

Poverty-Avoided (Income above ↑ Poverty Line)

US 1990	$>100%*	$>50%**	Change	Result
White	89%	96%	8%	W better
Black	68%	86%	26%	B better
Ratio W/B	1.31	1.12	-14%	B better

Scanlan (2000). Race and Mortality. Society 37(2):29-35

In Table 6, the left side shows that the poverty-sustained rates were lower under the 50% standard than under the 100% standard. On this basis, both groups were better off.

In comparing the below-poverty rates, the black-white ratio increased from 3.0 to 3.7. In terms of poverty-sustained, blacks were better off absolutely but worse off relative to whites. In terms of poverty-avoided, blacks were better off absolutely and better off relative to whites.

So, relative to whites during the decrease in the poverty level, blacks were better off in terms of poverty avoided, but worse off in terms of poverty sustained. Making some things better can make other things worse.

[5] https://www.irp.wisc.edu/resources/how-is-poverty-measured/

Appendix B: Excerpts from the Feminization of Poverty Issue

The following are quotes from Scanlan (1987). [Italics added by Schield]

"… ever since the phrase "the feminization of poverty" was first coined by sociologist Diana Pearce in 1978, the concern for the perceived increase in the proportion of the poor who are in families headed by women has been the major theme in [the] discussion of poverty in America."

A 1983 article by Washington Post columnist Judy Mann … illustrates the way the theme was pursued in the ensuing years. Stating that "[w]omen have become much more economically vulnerable in the past 30 years than is generally understood," she noted:

> "In 1959 only 14.8% of [whites] below the poverty level were [in families] headed by women; in 1980, more than a quarter of them were. The figures for black families are even more staggering; 24.4% of [blacks] below the poverty line in 1959 were [in families] headed by women, but 58.6% of them were by 1980."

More recently, conservative commentators have emphasized similar changes between 1959 and 1984, blaming those increases on feminism, welfare and easy divorce. Generally ignored, however, in this preoccupation with the provocative are certain critical features of the feminization of poverty. Most significant is the fact that *as a rule the feminization of poverty varies inversely with the amount of poverty, including the amount of poverty in female-headed families*.

A STATISTICAL DISTINCTION

That is, when there is much poverty, female-headed families will comprise a certain proportion of the poor; *as poverty decreases female-headed families, being those most susceptible to poverty, will comprise an increasing proportion of the poor, even as the poverty rate for such families is also declining*.

Thus, the major reason for the dramatic increase in the feminization of poverty, which actually occurred between 1959 and the middle 1970's, was an unprecedented reduction in poverty that included a dramatic reduction in the poverty of female-headed families. Among whites, for example, between 1959 and 1974, as the overall poverty rate declined from 18% to 9%, the poverty rate for persons in female-headed families dropped from 40% to 28%. Though far less poverty prone than in 1959, female-headed family members had almost doubled their representation among the poor (from 15% to 27%) while their representation among the white population had grown by only about a quarter. …

Table 7: Poverty Rates of White Families [Table created by Schield]

White Families	1959	1984	Change	Result
Female HOH	40%	28%	-30%	F better
ALL Families	18%	9%	-50%	All better
RR: Female/All	2.2	3.1	-40%	**F worse**
P(Female|Poverty)*	15%	27%	80%	**F worse**

Scanlan (1988). Poverty and Women. * Actual

It is important to understand that the tendency of a decrease in poverty to increase the feminization of poverty does not simply reflect that female-headed families do not share fairly in the reduction of poverty, as certain features of the data might suggest. Rather, *it is* *in the nature of normal distributions that a group that is poorer on average will make up a larger proportion of each increasingly more poverty prone group.*

In 1979, for example, female-headed families comprised 17% of persons with incomes between 150% and 125% of the official poverty line, 20% of persons between 125% and 100% of the poverty line, 25% of persons between 100% and 75% of the poverty line, and 35% of persons below 75% of the poverty line. Thus, it can be seen that when there are changes in the total amount of poverty, there will be changes in the proportion that female-headed families comprise of the poor without any actual change in the relationship of the income status of female-headed families to that of other persons.

The same underlying phenomenon manifests itself in other mathematical relationships that similarly misleadingly suggest a change in the relative well-being of two groups having different income distributions. *Whenever there is a decrease in poverty, the poorer group will have a smaller percentage decrease in its poverty rate than other groups, and the ratio of the poverty rate of the poorer groups to that of other groups will increase.*

For example, using the 1979 data discussed above, were there a general reduction in poverty such that only the persons previously below 75% of the poverty line remained in poverty, the poverty rate for female-headed families would be reduced by 26% (from 34.4% to 25.3%), while the poverty rate for all other persons would be reduced by 36% (from 9.6% to 6.1%), and the ratio of the poverty rate in female-headed families to that of other persons would increase from 3.6 to 1 to 4 to 1.

Thus, when the National Advisory Council in its 1980 report cited as an illustration of the "deepening inequality between men and women, that "in 1967, a woman heading a family was about 3.8 times more likely to be poor than a man heading one, [but by] 1977, after more than a decade of antidiscrimination efforts, she was about 5.7 times more likely to be poor," it was noting a change the direction of which was inexorably compelled by an underlying benign truth–namely, that the economic circumstances of male and female family heads improved measurably during this period.

The same properties of normal distributions will also tend to create the impression that female-headed families are less affected by increases in poverty–such as those observed in the years after 1979–whether or not such is actually the case. *These mathematical principles apply as well to a variety of comparisons between groups, such as black/white unemployment and black/white infant mortality. In 1983, for example, white and black infant mortality rates each reached an all-time low; correspondingly, the ratio of black to white infant mortality reached an all-time high.*

This is not to deny significant changes in the relative economic status of female-headed families. At various times such changes no doubt have occurred… But in the numerous commentaries that speak as if there have been real changes in the relative well-being of female-headed families, none indicates a complete understanding of the underlying functional relationships much less carries out the complex analysis required to separate the real from the apparent.

Appendix C: Conditions for Scanlan's Paradox

What are the conditions for the Scanlan effects? While decreasing the rate of an outcome *tends* to result in an increased disparity ratio, we want to know what is necessary. The shape and center of the distributions is critical. Rather than analyze the impact of various shapes, centers and standard deviations with a wide variety of changes in cut-points, we can summarize the impact of all of these combinations into three situations.

Table 8: Three Situations, Tabular Data: Overview

CutAdv < CutDis				CutAdv = CutDis				CutAdv > CutDis			
Negative	Before	After	Change	**Negative**	Before	After	Change	**Negative**	Before	After	Change
Adv	10%	**6%**	-40%	Adv	10%	**5%**	-50%	Adv	10%	**4%**	-60%
Dis	20%	10%	-50%	Dis	20%	10%	-50%	Dis	20%	10%	-50%

All three tables involve a negative or adverse outcome. All three tables have the same two rows: Advantaged (ADV) and Disadvantaged (DIS). The Disadvantaged group has the higher percentage who experience the adverse outcome. All three tables have the same two columns: the results Before and After the change.

All three tables have the same four data cells. Three of the data cell values are identical in each table. Only the values in the upper-right corner change. The reduction or cut can be obtained from the two values in a given row.

In Table 8, the titles of the three sub-tables summarize the situations. In the left table, the cut to the advantaged (40%) is less than the cut to the disadvantaged (50%). In the center table, the cuts are equal. In the right table, the cut to the advantaged (60%) is greater than the cut to the disadvantaged.

In the following tables, we examine the results of these three situations. Now we look at the results for the negative or adverse outcome (on the left) and the results for the positive or desired outcome on the right. The top three rows and columns are the same as we studied in the previous figures

Table 9: Three Situations, Tabular Overview: AdvCut Less Than DisCut

Cut**Adv**antaged(40%) < Cut**Dis**advantaged(50%) *points

Negative	Before	After	Change	Result	**Positive**	Before	After	Change	Result
Adv	**10%**	**6%**	-40%	Better	Adv	90%	94%	4%	Better
Dis	**20%**	**10%**	-50%	Better	Dis	80%	90%	13%	Better
Dis-Adv*	10	4	-60%	Better	Adv-Dis*	10	4	-60%	Better
Dis/Adv	2.0	1.7	-17%	Better	Adv/Dis	1.13	1.04	-7%	Better
P(Dis|-)	67%	63%	-6%	Better	P(Dis|+)	47%	49%	4%	Better

- Among Negatives: equal-size groups   + Among Positives: equal-size groups

The values in the four upper-left cells are just the complements of those in the left table. In Table 9 the table on the left contains the adverse or negative outcomes; the table on the right contains the desirable or positive outcomes. As noted in the title, the cut to the adverse outcomes is greater for the advantaged (60%) than for the disadvantaged (40%).

The far-right column summarizes the results of the change. The changes in the first two rows are both better because the changes involve less negative and more positive.

The bottom four row present four different ways of summarizing the relationship between advantaged and disadvantaged before and after the change.

- Dis-Adv: This is the difference in rates (measured in percentage points). This difference decreases in both tables. A smaller difference (disparity) is better.
- Dis/Low and Adv/Dis: The disparity ratios are calculated before and after the change in both tables. Approaching one is better; moving away from one is worse.
- P(Dis|-): percentage of negatives who are disadvantaged. P(Adv|+): percentage of positives who are disadvantaged. Both values assume that advantaged and disadvantaged have equal sizes. In this case, approaching 50% is better; moving away from 50% is worse.

By each of these 12 measures, things are getting better for this kind of reduction.

Table 10 has the same layout as Table 9. The only difference in the four bold values is this: the cut to the Advantaged now equals the cut to the Disadvantaged.

Table 10: Three Situations, Tabular Overview: AdvCut Equals DisCut

Cut**Adv**antaged(50%) = Cut**Dis**advantaged(50%) *points

Negative	Before	After	Change	Result
Adv	**10%**	**5%**	-50%	Better
Dis	**20%**	**10%**	-50%	Better
Dis-Adv*	10	5	-50%	Better
Dis/Adv	2.0	2.0	0%	**Same**
P(Dis|-)	67%	67%	0%	**Same**

- Among Negatives: equal-size groups

Positive	Before	After	Change	Result
Adv	90%	95%	6%	Better
Dis	80%	90%	13%	Better
Adv-Dis*	10	5	-50%	Better
Adv/Dis	1.13	1.06	-6%	Better
P(Dis|+)	47%	49%	3%	Better

+ Among Positives: equal-size groups

Since the layout in this table is the same as that in the previous table, we don't need to examine all the details. The big change is in the bottom three rows in the left (negative) table. Instead of being positive, the changes are neutral: the values after the change are identical to those before the change.

Table 11 has the same layout as Table 9 and Table 10. The only difference in the four bold values is this: the cut to the Advantaged is now larger than the cut to the Disadvantaged.

Table 11: Three Situations, Tabular Overview: AdvCut More than DisCut

Cut**Adv**antaged(60%) > Cut**Dis**advantaged(50%) *points

Negative	Before	After	Change	Result
Adv	**10%**	**4%**	-60%	Better
Dis	**20%**	**10%**	-50%	Better
Dis-Adv*	10	6	-40%	Better
Dis/Adv	2.0	2.5	25%	**Worse**
P(Dis|-)	67%	71%	7%	**Worse**

- Among Negatives: equal-size groups

Positive	Before	After	Change	Result
Adv	90%	96%	7%	Better
Dis	80%	90%	13%	Better
Adv-Dis*	10	6	-40%	Better
Adv/Dis	1.13	1.07	-5%	Better
P(Dis|+)	47%	48%	3%	Better

+ Among Positives: equal-size groups

The big changes are in the bottom two rows in the left (negative) table. Instead of being positive or neutral, the changes are negative or worse. These is Scanlan's paradox. Table 11 illustrates – but does not prove – that the associated condition is necessary for Scanlan's paradox effects. Those proofs are contained in Appendix D and Appendix F. .

Appendix D: Conditions for Scanlan's Paradox

Claim: If the adverse outcomes of advantaged and disadvantaged are being reduced, then the adverse relative difference (the adverse risk ratio) must increase if the advantaged have a greater relative reduction in prevalence than [have] the disadvantaged (more susceptible). If the relative reductions are identical, then the adverse disparity ratio will be unchanged.

1. P(Low,T1): Percentage of Low who have adverse outcomes before the change
2. P(High, T2): Percentage of High who have adverse outcomes after the change.
3. Assume: P(High,T1) - P(Low,T1) > 0 so P(High,T1) / P(Low,T1) > 1.
4. rCut(Low) = [P(Low,T1) – P(Low, T2)] / P(Low, T1). Note: rCut() > 0.

5. Ratio(Tk) = P(High,Tk)/P(Low,Tk)
6. Ratio(T2) – Ratio(T1) = [P(High,T2)/P(Low,T2)] – [P(High,T1)/P(Low,T1)]

Ratio(T2) – Ratio(T1) > 0 ** Scanlan's paradox conclusion **

7. if [P(High,T2)/P(Low,T2)] > [P(High,T1)/P(Low,T1)]
8. if [P(High,T2)/P(High,T1)] > [P(Low,T2)/P(Low,T1)]
9. if [1 – rCut(High] > [1 – rCut(Low)]
10. if rCut(High) < rCut(Low) or if rCut(Low) > rCut(High). QED.

Appendix E: Prevalence of disadvantaged among those with adverse outcomes

Claim: If the advantaged have a greater relative reduction in adverse prevalence than have the disadvantaged, then the fraction of High among those with adverse outcomes will increase regardless of the fraction of subjects who are High.

1. Let f = fraction of subjects who are Highs.
2. Let P(High, 1|Adverse) = f*P(High,T1) / [f*P(High,T1) + (1-f)*P(Low, T1)]
3. Let P(High, 2|Adverse) = f*P(High,T2) / [f*P(High,T2) + (1-f)*P(Low, T2)]

4. Let J(High) = P(High,T2) / P(High,T1). Let J(Low) = P(Low,T2) / P(Low, T1).
5. P(High, 1|Adverse) = f*P(High,T1) / [f*P(High,T1) + (1-f)*P(Low, T1)]
6. P(High, 2|Adverse) = f*J(Dis)*P(High,T1)
   / [f*J(Dis)*P(High,T1) +(1-f)*J(Adv)*P(Low, T1)]
7. P(High, 2|Adverse) = f*P(High,T1) /
   {f*P(High,T1) + (1-f)*[J(Low)/J(High)]*P(Low, T1)}
8. If we have a Scanlan effect, then J(Low) < J(High). If so, then
9. P(High, 2|Adverse) - P(High, 1|Adverse)
   = f*P(High,T1) / {f*P(High,T1)+(1-f)*[J(Adv)/J(Dis)]*P(Low, T1)}
   minus f*P(High,T1) / [f*P(High,T1) + (1-f)*P(Low, T1)].

For a Scanlan effect, J(Adv) < J(Dis).

Numerators are identical. If J(Adv)/J(Dis) < 1, then denominator of P(High,T2|Adverse) is less than that of P(High,1|Adverse). So P(High,T2|Adverse) > P(High,T1|Adverse).

This is true regardless of the value of f. So, if the prevalence of the disadvantaged group among the adverse outcome increases as the cutoff is lowered for equal size groups, then it will increase regardless of the relative size of the two groups..

Appendix F: As B/A Ratio Increases, so does Share of Adverse who are B

Claim: If the B/A disparity ratio increases, then the B/(A+B) prevalence must increase.

1. B(T2)/A(T2) > B(T1)/A(T1)
2. 1 + B(T2)/A(T2) > 1 + B(T1)/A(T1) — Add 1 to both sides
3. [A(T2)+B(T2)]/A(T2) > [A(T1)+B(T1)]/A(T1) — Convert one
4. A(T1) / [A(T1)+B(T1)] > A(T2) / [A(T2)+B(T2)] — Get sum in the denominator
5. {[A(T1)+B(T1)]–B(T1)}/[A(T1)+B(T1)]>{[A(T2)+B(T2)]–B(T2)}/[A(T2)+B(T2)]
6. 1 – {B(T1) / [A(T1)+B(T1)]} > 1 - {B(T2) / [A(T2)+B(T2)]}
7. – B(T1) / [A(T1)+B(T1)] > - B(T2) / [A(T2)+B(T2)] — Eliminate minus sign
8. B(T2) / [A(T2)+B(T2)] > B(T1) / [A(T1)+B(T1)] — QED

Appendix G: Adverse Conditions Generating Equal favorable Risk Ratios

Solve for the adverse (decreasing) conditions that create equal values for the favorable (increasing) risk ratios.

Table 12: Solving for Equal Risk Ratios for Increasing Outcomes

Decrease	Before	After	Fraction
Adv	A1	**A2**	A3=A2/A1
Dis	D1	D2	D3=D2/D1
Dis/Adv	D1/A1	D2/A2	

Increase	Before	After
Adv	1-A1	1-A1*A3
Dis	1-D1	1-D1*D3
Adv/Dis	(1-A1)/(1-D1)	(1-A1*A3)/(1-D1*D3)

1. (1-A1)/(1-D1) = (1-A1***A3**) / (1-D1*D3)
2. (1-A1) (1-D1*D3) = (1-A1***A3**) (1-D1)
3. 1-A1-D1*D3 +A1*D1*D3 = 1-A1***A3**-D1+A1***A3***D1
4. A3*(A1-A1*D1) = A1-D1 -D1*D3 + A1*D1*D3
5. A3 = (A1-D1 + D1*D3 -A1*D1*D3) / (A1-A1*D1)

Here are the five combinations of the behavior of the two sets of relative risk (assuming the larger number always appears in the numerator). The critical value is the value (A3) shown above.

	CHANGE IN RISK RATIO BY OUTCOME	Adverse Outcome**	Favorable Outcome*
#	**ACTION ON THE ADVERSE (DECREASING) RATES**	**RR: DIS/ADV**	**RR: ADV/DIS**
1	Disadvantaged rate decreases more than Advantaged	Decreases	Decreases
2	Disadvantaged decreases at same rate as Advantaged	Unchanged	Decreases
3	*Advantaged decreases little more than the Disadvantaged*	*Increases*	*Decreases*
4	Advantaged rate decreases to critical value.	Increases	Unchanged
5	Advantaged decreased by more than critical value	Increases	Increases

**Adverse outcome: rates are decreasing. *Favorable outcome: rates are increasing.

Figure 6: Change in Risk Ratio for adverse outcome.

Situation #3 seems to be the most common. More analysis is needed to understand this situation.

Appendix H: Excerpts from CDC (2005) Methodological Issues …

This CDC report mentions Scanlan as reference (6).

> "This report discusses six significant issues that should be considered in measuring disparities:… third, *measurement of disparity in terms of favorable or adverse events (6,7)*; ...sixth*, choosing whether or not to consider the order inherent in the domains with ordered categories when calculating summary measures of disparity (6)*. These measurement choices affect the way a disparity is expressed, including the size and direction of the disparity." [Italics added]

p. 5: Measuring Disparity in terms of adverse events

> Most health-related indicators can be expressed either in terms of favorable events or in terms of adverse events. A favorable event or characteristic is considered desirable and is promoted through public health action. An adverse event or characteristic is considered undesirable, and reduction or elimination is promoted through public health action.
>
> Conclusions about changes in disparity over time also depend on whether an indicator is expressed in terms of favorable or adverse events. (6) Table D [Table 13] shows the percentage of women who *had* a mammogram during the past 2 years and the percentage of women who *did not have* a mammogram during the past 2 years for non-Hispanic white and Hispanic women in 1990 and 1998 (10). The *simple difference* between the percentage of non-Hispanic white and Hispanic women who *had* a mammogram during the past 2 years increased from 7.5 percentage points in 1990 to 7.8 percentage points in 1998. The simple difference also increased from –7.5 percentage points to –7.8 percentage points for women who *did not have* a mammogram during the past 2 years.

Table 13: Use of Mammography 1998 vs 1990. Hispanic vs. White [CDC Table D]

1990 versus 1998: Percentage of women over 40 by use of mammography: Hispanic vs. Non-Hispanic white

Did not have	1990	1998	Change	
White nonH	47.3%	32.0%	-32%	W better
Hispanic	54.8%	39.8%	-27%	Hispanic better
W-H (pts)	-7.5	-7.8	4.0%	**W-H worse**
RD: (W-H)/H	-13.7%	-19.6%	43.2%	**W/H worse**

HAD	1990	1998	Change	
White nonH	52.7%	68.0%	29%	White better
Hispanic	45.2%	60.2%	33%	Hispanic better
W-H (pts)	7.5	7.8	4.0%	**W-H worse**
RD: (W-H)/H	16.6%	13.0%	-21.9%	W/H better

National Center for Health Statistics. US 2004 With Chartbook on Trends in the Health of Americans

> However, changes in the relative differences between Hispanic and non-Hispanic white women were not the same for both favorable and adverse events. The *percentage difference* between non-Hispanic white and Hispanic women who *had* a mammogram within the past 2 years decreased from 16.6 percent in 1990 to 13.0 percent in 1998. A decrease in disparity is indicated because the percentage difference moved closer to 0. The percentage difference between non-Hispanic white and Hispanic women who *did not have* a mammogram within the past 2 years increased from –13.7 percent in 1990 to –19.6 percent in 1998. The percentage difference moved further from 0, indicating an increase in disparity.
>
> In this example, both the magnitude and direction of change in the relative measure of disparity depend on whether the indicator is expressed in terms of favorable or adverse events. <snip> Similar results can occur when comparisons are made across different indicators, geographic areas, or populations.

Appendix I: Scanlan's Paradox Explained using Blocks for Categories

The conditions for Scanlan's paradox have been presented in algebraic and tabular forms. This section uses blocks. Consider school suspensions where the prevalences are shown by cause (minor and major) for each group (Advantaged and Disadvantaged). This data is realistic, but not real. It was chosen to illustrate the effect of different mixtures.

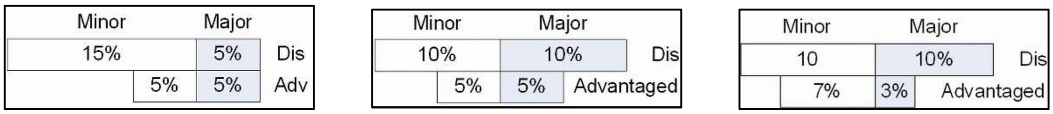


Figure 7: Three suspension Scenarios: Different Mixture of Minor and Major Infractions

Initially both minor and major infractions are grounds for suspension. All three scenarios start with 20% of the disadvantaged being suspended; 10% of the advantaged. This gives a disadvantaged-advantaged disparity ratio for suspensions of two-to-one.

In each case, the criteria for a suspension change. Minor infractions are no longer grounds for suspension. The resulting suspension rates involve just the major infractions.

For the situation on the left: Advantaged rate cut LESS than Disadvantaged rate:
- Infractions are cut by 50% for the advantaged (from 10% to 5%); infractions are cut by 75% for the disadvantaged (from 20% to 5%).
- The disadvantaged-advantaged disparity ratio for suspensions *decreased* from two-to-one (20%/10%) to one-to-one (5%/5%): a 50% decrease.
- For equal sized groups, the share of the suspended that were disadvantaged *dropped* from 67% (20%/30%) to 50% (5%/10%).
- If the disadvantaged rate is cut by a higher fraction than the advantaged rate, then the disadvantaged group is *better off in all three ways*.

For the situation in the center: Advantaged rate cut THE SAME as Disadvantaged rate:
- Infractions are cut by the same fraction (50%) for both groups.
- The disadvantaged-advantaged disparity ratio for suspension is *unchanged* at two-to-one.
- For equal sized groups, the share of the suspended that were disadvantage was *unchanged* at 67% (20% / 30%) initially and 67% (10%/15%) afterwards.
- If the disadvantaged rate is cut by the same fraction as the disadvantaged rate, then the disadvantages are *better off absolutely, but are unchanged otherwise*.

For the situation on the right: Advantaged rate cut MORE than Disadvantaged rate:
- Infractions are cut by 70% for the advantaged (50% for the disadvantaged).
- The disadvantaged-advantaged disparity ratio for suspensions *increased* from two-to-one to more than three to one: a 50% increase.
- For equal sized groups: the percentage of the suspended that were disadvantaged *increased* from 67% (20% / 30%) to 77% (10%/13%).
- In terms of suspensions, after versus before: if the advantaged rate is cut by a higher fraction than the disadvantaged rate, then the disadvantaged are *better off absolutely* but *worse off by the other two measures*

Scanlan's paradox occurs only in the last (right-most) situation: the adverse rates are cut by fractionally more for the Advantaged than for the Disadvantaged.

Appendix J: Comparing Odds Ratios

To make an 'apples and apples' comparison, we need a measure that doesn't have a ceiling. Instead of being bounded between zero and one as are part-whole percentages, we need a measure that transforms the range from zero to one into the range from zero to infinity. Furthermore, the measure should not make any assumption about the shape of a distribution. The Odds Ratio has that property.

Compare adverse 40% with 20% (favorable 60% with 80%) where 40% is Treatment:
OR1a: (40/20)/(60/80) = 2.67
OR1b: (60/80)/(40/20) = 0.33 Exchange Favorable and Adverse

OR2a: (40/20)/(60/80) = 2.67
OR2b: (20/40)/(80/60) = 0.67 Exchange Treatment and Control groups.

So, the Odds Ratio is not invariant when exchanging the favorable and adverse percentages (their compliments) or when exchanging the treatment and control group percentages.

The Odds Ratio is invariant after making both exchanges:

- OR3a: (40/20)/(60/80) = 2.67 where 40 and 20 are adverse (40 is treatment group)
- OR3b: (20/40)/(80/60) = 0.67 where treatment and control groups are exchanged.
- OR3c: (80/60)/(20/40) = 2.67 where favorable and adverse are exchanged

- OR4a: (40/20)/(60/80) = 2.67 where 40 and 20 are adverse (40 is treatment group)
- OR4b: (60/80)/(40/20) = 0.33 where favorable and adverse are exchanged.
- OR4c: (80/60)/(20/40) = 2.67 where treatment and control groups are exchanged

Appendix K: Using Probit and Logit

Given starting values for two frequencies or proportions, Probit gives the same standardized mean differences (d) and correlation coefficients (r) with opposite signs in two exchanges:

- Exchanging proportions for Treatment and Control groups
- Exchanging proportions for Favorable and Adverse outcomes (compliments)

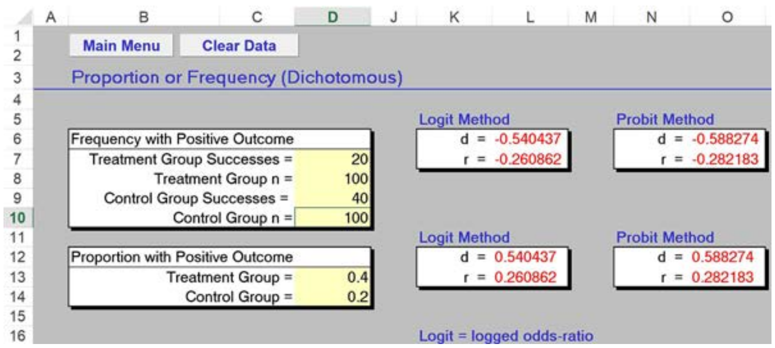


Figure 8: Logit and Probit Case 1

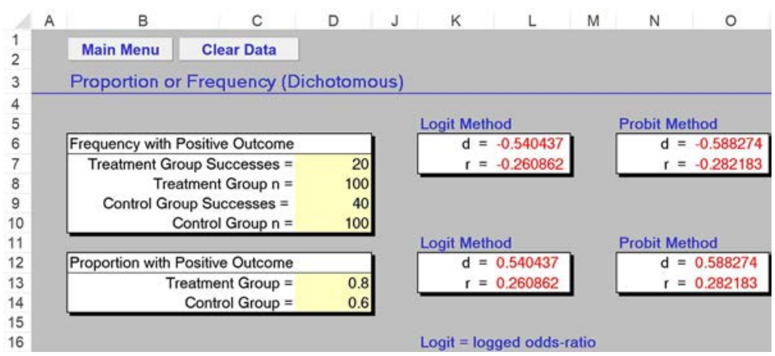


Figure 9: Logit and Probit Case 2

Logit generates different starting values (differences and correlations) than Probit given the same frequencies or proportions. But these differences and correlations are preserved in both exchanges (but with opposite signs in each).

These results are generated by the downloadable ES-Calculator.xls file made available on a webpage of Professor David B. Wilson of George Mason University. "The Effect Size Determination Program is for use with the book *Practical Meta-Analysis* written by Mark W. Lipsey and David B. Wilson and published by Sage." "This program computes standardized mean difference effect sizes (d) and correlation coefficients (r) from summary statistics." Weblink: http://mason.gmu.edu/~dwilsonb/ma.html

Appendix L: Comparisons Using Probit and Odds Ratio

Reconsider two of the cases presented previously.

Table 14: Infant Mortality vs. Survival with Probit and Odds Ratios

US Infant Mortality Rate (Death per 1,000 live births)

Deaths	1983	1997	Change	Result
White	**9.7**	**6.0**	-38%	W better
Black	**19.2**	**14.2**	-26%	B better
P(B|Death)*	66%	70%	6%	**B/W worse**
Ratio B/W	1.98	2.37	20%	**B/W worse**
OR: B/W	2.00	2.39	19%	**B/W worse**
Probit d	0.27	0.32	20%	**B-W worse**

Scanlan (2000). Race and Mortality. Society 37(2):29-35

US Infant Survival Rate (Survivors per 1,000 live births)

Suvivors	1983	1997	Change	Result
White	990	994	0.4%	W better
Black	981	986	0.5%	B better
P(B|Living)*	49.76%	49.79%	0.1%	W/B better
Ratio W/B	1.010	1.008	-0.1%	W/B better
OR: W/B	2.00	2.39	19%	**W/B worse**
Probit d	0.27	0.32	20%	**W-B worse**

* Assume 50-50 split white and black

A decrease in the death rate, an increase in the survival rate:

- Odds Ratio: Moved black-white further from unity for favorable and adverse
- Probit: Increased Probit distance for adverse and favorable.

Table 15: Family Poverty Sustained vs. Avoided with Probit and Odds Ratios

Poverty Sustained (Income below ↓ Poverty Line)

US 1990	$<100%*	$<50%**	Change	Result
White	**10.7%**	**3.9%**	-64%	W better
Black	**31.9%**	**14.4%**	-55%	B better
P(B|⇩Pov)	75%	79%	5.1%	**B/W worse**
Ratio B/W	2.98	3.69	24%	**B/W worse**
OR: B/W	3.91	4.15	6%	**B/W worse**
Probit d	0.772	0.700	-9%	B/W better

* Income below (above) 100% (** 50%) of the poverty line

Poverty-Avoided (Income above ↑ Poverty Line)

US 1990	$>100%*	$>50%**	Change	Result
White	89.3%	96.1%	8%	W better
Black	68.1%	85.6%	26%	B better
P(B|⇑Pov)	43%	47%	8.9%	B/W better
Ratio W/B	1.31	1.12	-14%	W/B better
OR: W/B	3.91	4.15	6%	**W/B worse**
Probit d	0.772	0.700	-9%	W/B better

Data: Scanlan (2000). Race and Mortality. *Society* 37(2):29-35

A decrease in the poverty line generates a decrease in the poverty sustained rate and an increase in the poverty avoided rate:

- Odds Ratio: Moved further from unity for favorable and adverse
- Probit: Decreased Probit distance for adverse and favorable.

The odds ratio approaches the risk ratio as the percentages involved approach zero. That can be seen in Table 14 for the deaths. Thus, the Odds ratio is closely related to the Black/White ratio among the deaths (the disadvantaged/advantaged ratio among the adverse outcome).

The Probit standardized mean difference is more complex. In Table 14, it worsens for both adverse and favorable. In Table 15, it improves for both adverse and favorable. Since the Probit is imposing a normal distribution on each pair of data points, this kind of variation is not unexpected.

Can the differences in Probit in these two cases reflect an underlying difference? Table 14 is based on a change in the underlying reality. Table 15 is based on a change in the rulea. Can Probit discern these differences? That seems unlikely. All Probit sees are the four rates. It seems that both types of change could generate the same rates.