Controlling for Context By Standardizing

Milo Schield

Abstract: Controlling for the context can be done by standardizing: adjusting the mixtures to convert a mixed-fruit comparison into an apples and apples comparison. This paper presents two simple manual techniques for standardizing percentages and rates when the predictor and confounder are both binary. Both involve the weighted average. One involves arithmetic calculation; the other involves a graphical approach. Students find the calculation easier to do, but the graphical technique helps them 'see' what controlling for (taking into account) really means.

Introduction:

Today's students want to engage in social issues. Most social issues involve statistics: averages, counts and rates. Most social statistics are crude statistics. They ignore the context; they don't take related factors into account. To really understand social statistics, students need to "see" the wider context; they need to "see" how to *take something into account*. Students get engaged in seeing that social statistics may have a *story behind the statistic.*

Most social statistics are observational and are influenced by factors that need to be controlled for. Controlling for related factors changes a crude statistic into an adjusted statistic: a crude comparison into an adjusted comparison.

Why is the Covid-19 infection rate much higher in Italy (1,333/M) than in the US (279/M)? [3/25]

- Older people are a bigger share of the population in Italy (23%) than in the US (17%).
- Population density is higher in Italy (533 per sq. mile) than in the US (94 per sq. mile).

To compare the infection rate of Italy with of the US, such confounders may need to be "adjusted" by including the wider context – by taking these related factors into account.

"Taking into account" and "controlling for" are mental.

Computer methods of controlling for confounders are powerful, but they may obscure the process. Manual methods using weighted average are easy to use. Using graphs, they can "show" students the key ideas in standardizing.

Table 1: Ways of mentally controlling for (taking into account) confounders

CONTROLLING FOR CONFOUNDERS			
Take into account (mental)			
	Can do by hand		*Calculator/Computer*
1	Select/Stratify	4	Linear Regression
2	Form Ratios	5	Logistic Regression
3	Standardize	6	Multivariate Regress

Standardizing (adjusting) requires a standard for matching the mixtures. Here are two standards:

1. *Standard-group matching* means selecting one group as the standard and adjusting the other group mixture to match that standard. (Cf. demography)
2. *Combined-group matching* adjusts both group mixtures to their combined mixture. (Cf. regression)

Calculations can be done algebraically or graphically. With two standards, there are four combinations.

Table 2: Ways of Controlling for Confounders

Controlling for:	Calculation	
Standardize on	**A**lgebraic	**G**raphical
1 Standard Group Mixture	1A	1G
2 Combined Group Mixture	2A	2G

The following tables and figures show both methods (algebraic and graphical) for both types of standardizing (standard and combined group). In each of the four cases, we use the same data: patients' death rates at two hospitals: City and Rural. Patients are classified as being in either good or poor condition. Thus, the predictor (hospital) and the confounder (patient condition) are both binary. This greatly simplifies things so students can better see what it means to control for – to take into account – a related factor.

1: Standardizing Rural on the Mixture at City:

1A Standard Group: Algebraically Adjust Rural Mix to Match City:

The table on the left shows the patients' death rates by hospital and patient condition. The mixture of patients by hospital is shown by the fractions in parenthesis. Weighted averages are on the right. The algebra is at the bottom.

The patients' death rate at the City hospital (5.5%) is 1.3 percentage points higher than that at the Rural hospital (4.2%). The table on the left side is a crude comparison: a mixed-fruit comparison and a Simpson's paradox. City has a lower death rate than Rural for patients in either condition. Yet City has a higher patients' death rate overall.

Table 3: Patients' Death Rates at City and Rural Hospital: Standardizing Rural on City

Patients' Death Rate (Mix: Fraction in this condition)			
	Patient Condition		
Hospital	Good	Poor	All
City	1% (0.1)	6% (0.9)	5.5%
Rural	3% (0.7)	7% (0.3)	4.2%
All: City	= 0.1*1% + 0.9*6%		1.3 points
All: Rural	= 0.7*3% + 0.3*7%		City higher

Match Rural Mix to City; Apply City Mix to Rural			
Patients' Death Rate (Mix: Fraction in this condition)			
Hospital	Good Cond.	Poor Cond.	All
City	1% (0.1)	6% (0.9)	5.5%
Rural	3% **(0.1)**	7% **(0.9)**	**6.6%**
All: City	= 0.1*1% + 0.9*6%		**-1.1 pts**
All: Rural	**= 0.1*3% + 0.9*7%**		City lower

The table on the right standardizes the demographics of the Rural Hospital to match those at City. The items in bold have been adjusted. The result is an adjusted comparison: an "apples and apples" comparison and a reversal. The Rural Hospital's adjusted patients' death rate (6.6%) is higher than that at City Hospital (5.5%): 1.1 points higher.

1G Standard Group: Graphically Adjust Rural Mix to Match City

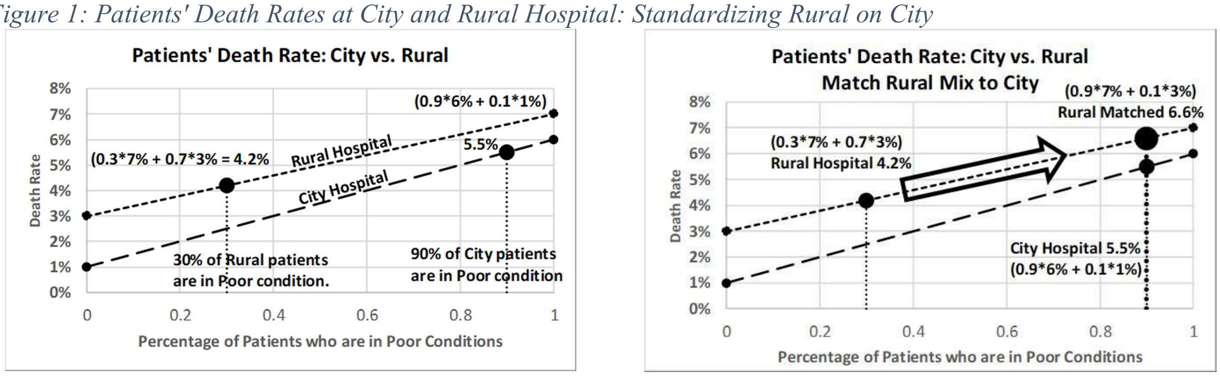
On these slides we repeat what we did on the previous slides – but now we are going to do it graphically.

The vertical axis is the patients' death rate. The horizontal axis is the mixture of patients' conditions: the percentage of patients who are in poor condition. On the right side of each graph (at 100%) are the patients in Poor condition along with their death rates at City and at Rural. On the left side (at 0%) are the patients in good condition along with their death rates at City and Rural. The diagonal lines connecting the patients' death rates for each hospital are the weighted average lines for groups of patients having various mixtures of patients in good and poor condition.

In the graph on the left, patients in poor condition are 90% of City hospital patients (30% of Rural hospital patients). The dots are the weighted averages. The crude death rate at Rural (4.2%) is 1.3 points lower than at City (5.5%)

But this is a crude comparison: a mixed-fruit comparison; an apples and oranges comparison. It is also a Simpson's paradox. City has a lower patients' death rate for patients in poor and in good condition, but higher rate overall.

Figure 1: Patients' Death Rates at City and Rural Hospital: Standardizing Rural on City



In the graph on the right, the mixture of Rural patients was adjusted to match the mixture at City. The death rate at Rural moves up the Rural weighted average line. The adjusted rate at Rural (6.6%) is now higher than at City (5.5%): a difference of 1.1 percentage points. This is identical to what we did algebraically in the previous table.

After '*controlling for*' (taking into account) patient condition, the patients' death rate at Rural hospital is now higher than that at City hospital. This adjusted comparison is now an apples and apples comparison.

2: Standardizing Both on their Combined Mixture:

We now repeat the process (algebra and graph) but standardize both on the combined mixture: assumed to be where 70% of patients are in poor condition.

2A Standardize Both on the Combined Mixture: Algebraic Method

Table 4: Patients' Death Rates at City and Rural Hospital: Standardizing Both on Combined Mixture

Patients' Death Rate (Mix: Fraction in this condition)			
	Patient Condition		
Hospital	Good	Poor	All
City	1% (0.1)	6% (0.9)	5.5%
Rural	3% (0.7)	7% (0.3)	4.2%
All: City	= 0.1*1% + 0.9*6%		1.3 points
All: Rural	= 0.7*3% + 0.3*7%		City higher

Match City & Rural Mixes to Combined Mix			
Patients' Death Rate (Mix: Fraction in this condition)			
Hospital	Good Cond.	Poor Cond.	All
City	1% **(0.3)**	6% **(0.7)**	**4.5%**
Rural	3% **(0.3)**	7% **(0.7)**	**5.8%**
All: City	**= 0.3*1% + 0.7*6%**		**-1.3 pts**
All: Rural	**= 0.3*3% + 0.7*7%**		City lower

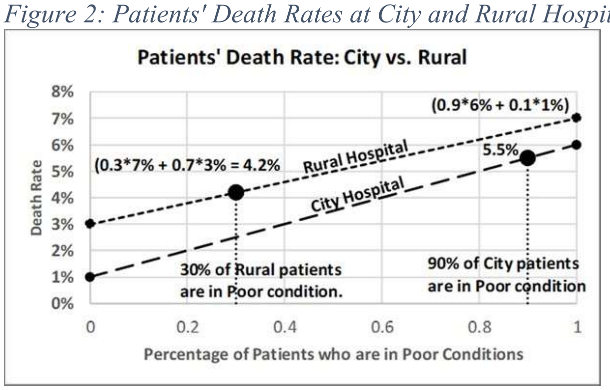
The table on the left is the same as the table on the left in Table 3. The table on the right is similar to the table on the right in Table 3, but with one difference: the patient mix at both hospitals is adjusted to the combined mixture.

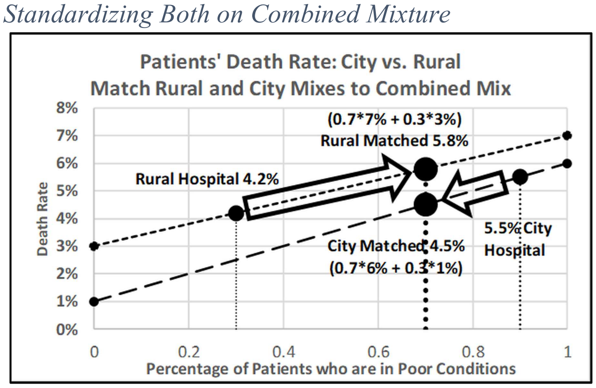
After controlling for patient condition, the patients' death rate at Rural (5.8%) is 1.3 percentage points higher than that at City (4.5%). This adjusted comparison is an 'apples and apples' comparison.

The weights are shown as decimal fractions. Some students get confused when the weights and rates are both shown as percentages.

2G Standardize Both on the Combined Mixture: Graphical Method

Figure 2: Patients' Death Rates at City and Rural Hospital: Standardizing Both on Combined Mixture





The graph on the left side of Figure 2 is identical to that on the left side of Figure 1. The graph on the right side of Figure 2 is similar to that on the right side of Figure 1 but with a slight difference: the patient mix at both hospitals is adjusted to the combined mixture.

As before, the graphical results are the same as the algebraic results. After controlling for patient condition, the patients' death rate at Rural (5.8%) is higher than that at City (4.5%): 1.3 percentage points higher.

This weighted-average graph is very unusual. The horizontal axis shows both the binary values (left and right extremes) and the prevalence of the right group in the combined mixture. Check out the references to Wainer (2002) and Schield (2006).

See Appendix A on the history of this graph.

3: Comparing the Results of the Different Standards:

Appendix B shows the results of adjusting the mixture at City to match that at Rural.[1]

Notice that the different standards resulted in these different separations after adjustment.

Mixture Standard	Separation*
Adjust Rural to match City	1.1
Use Combined Mixture	1.3
Adjust City to match Rural	1.7
* Measured in percentage points	

Why do the different choices give different separations? The algebra generates these numbers, but it doesn't show why there are differences.

The graphs show why. Notice that the weighted average lines for City and Rural are not parallel. The lines are closer together at the 90% for City, further apart at the 70% combined prevalence, and furthest apart at the 30% for Rural. So, standardizing on Rural gives the biggest separation, standardizing on City gives the smallest, and standardizing on combining will always be in between.

4: Comparing the Student and Teacher Perceptions of the Two Methods:

The algebraic and graphical methods give the same results for a given standard. But they present different problems for students and teachers.

PLUS: Compared to the graphical technique, the Algebraic techniques seem to be

- simpler for teachers to teach (graphs take time),
- simpler for students to do (graphs are tricky), and
- more applicable (applies to more than two groups).

MINUS: Compare to the Graphical techniques, the Algebraic techniques

- are calculation based (not visual. Students can't see), and
- are very sensitive to the wording (match, apply).

5: My Plan:

I taught combined-group graphs (#2G) for over 10 years. This last two years I taught standard-group algebra (#1A). In the future, I plan to start with *standard-group algebra* (#1A): it is simpler to teach and simpler to do. Then I will ask students to show their results using Wainer-Lesser graph (#1G). Given time, I will introduce *combined-group algebra* (#2A) and the associated Wainer-Lesser graph (#2G).

Doing both the algebraic and the graphical should help students to better *understand* what it means when a comparison has controlled for (taken into account) the influence of a related factor.

6: Conclusion:

If students can distinguish a crude and an adjusted comparison, and if they really understand what it means to control for (take into account) the influence of a related factor, they are well on their way to being statistically literate. Ideally, they will better understand that statistics are numbers in context.

[1] Appendix A presents the history of the graphical technique. Appendix C presents the comments posted at ECOTS.

Bibliography:

Jeon,]. W, Chung, !-1. Y., and Bae,]. S. (1987), "Chances of Simpson's Paradox," Journal of the Korean Statistical Society, J 6, 117-125.

Baker, Stuart. G., and Kramer, Barry. S. (2001), "Good for Women. Good for Men, Bad for People: Simpson's Paradox and the Importance of Sex-Specific Analysis in Observationa1 Studies," *Journal, of Women's Health and Gender Based Medicine*, 10, 867-872.

Lesser, Larry (2001). Representations of Reversal: An Exploration of Simpson's Paradox. 2001 Yearbook of the National Council of the Teaching of Mathematics. Copy at www.statlit.org/pdf/2001LesserNCTM.pdf

Schield, Milo and Tom Burnham (2003). Confounder-Induced Spuriosity and Reversal. 2003 ASA Proceedings of the Section on Statistical Education, pp. 3690-3697. Copy at www.statlit.org/pdf/2003SchieldBurnhamASA.pdf

Schield, Milo (2006). Understanding Confounding from Lurking Variables using Graphs. *STATS Magazine* ASA. Fall 2006, pp. 14-18. Copy at www.StatLit.org/pdf/2006SchieldSTATS.pdf

Schield, Milo (2020). Controlling for Context by Standardizing. Video of slides presented at ECOTS.
Slides at www.StatLit.org/pdf/2020-Schield-ECOTS-Slides.pdf
Video at www.StatLit.org/V/2020-Schield-ECOTS-Slides.pdf and at https://youtu.be/eTiBpLuNnyA

Tan, Arjun (2012). How I Created a Trapezoidal Display of Simpson's Paradox. Copy at http://www.statlit.org/pdf/2012-Arjun-Tan-Simpsons-Paradox.pdf

Wainer, H. (2002). "The BK-Plot: Making Simpson's Paradox Clear to the Masses." *CHANCE*, 15(3):60-62. www.statlit.org/Wainer.htm Copy at: www.StatLit.org/cp/Cornfield/2002-Wainer-Visual-Revelation-BK-plot.pdf

Acknowledgements:

Thanks to Marc Isaacson for the title of – and comments on – the slides. Thanks to the ECOTS for hosting my five-minute video presentation. And thanks to Roger Bissell for his contributions as my editor.

The Author:

Milo Schield is a Fellow of the American Statistical Association (ASA), an elected member of the International Statistical Institute (ISI) and the editor of. He is currently the President of the National Numeracy Network (NNN), a US representative of the International Statistical Literacy Project (ISLP) and a full Professor in Business Administration at Augsburg University.

Appendix A: The Origin of the Graphical Technique

Wainer (2002) aptly describes this value of this graphical technique in his title: "The B-K Plot: Making Simpson's paradox clear to the masses." Wainer ascribes this technique to Baker and Kramer (2001). Baker and Kramer noted that it had been independently discovered by Jeon and Bae (1987).

Lesser (2001) describes this technique as the "Trapezoidal Representation" and ascribe it to Tan: "Tan (1986) provides a trapezoidal representation of Simpson's paradox that is built only on the observation (Hoehn 1984) that "[t]he length of any line segment which is parallel to the two bases and has its endpoints on the nonparallel sides of a trapezoid is the weighted mean of the lengths of the two bases."

Tan (2012) describes the process by which he generated the trapezoidal display in 1986.

One might name this weighted-average graph the Hoen graph since he was arguably the first to demonstrate its utility. One might name it the Tan graph since he was the first to show the graph as applied to Simpson's paradox. One might name this graph the trapezoidal graph based on Tan's description of the shape. One might name it the Tan- Jeon-Bae-Baker-Kramer graph after all the known independent creators. Or one might name it the Wainer-Lesser graph after the two that did the most to publicize this unique graph to statistical educators. I favor the latter.

Appendix B: Standardizing City on the Mixture at Rural

To see the differences caused by the choice of a standard, consider standardizing on a single group but adjusting the mixture at City to match that at Rural (applying the mixture for Rural to City).

A: Standardizing on Rural: Algebraic Method

Table 5: Patients' Death Rates at City and Rural Hospital: Standardizing City on Rural

Patients' Death Rate (Mix: Fraction in this condition)			
	Patient Condition		
Hospital	Good	Poor	All
City	1% (0.1)	6% (0.9)	5.5%
Rural	3% (0.7)	7% (0.3)	4.2%
All: City	= 0.1*1% + 0.9*6%		1.3 points
All: Rural	= 0.7*3% + 0.3*7%		City higher

Match City Mix to Rural; Apply Rural Mix to City			
Patients' Death Rate (Mix: Fraction in this condition)			
Hospital	Good Cond.	Poor Cond.	All
City	1% **(0.7)**	6% **(0.3)**	**2.5%**
Rural	3% (0.7)	7% (0.3)	4.2%
All: City	**= 0.7*1% + 0.3*6%**		**-1.7 pts**
All: Rural	= 0.3*7% + 0.7*3%		City lower

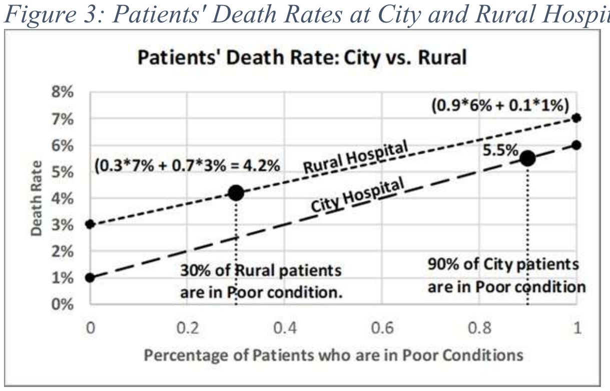
The table on the left is the same as before. The comparison of death rates is a crude comparison: an unadjusted comparison – a mixed-fruit comparison; an apples and oranges comparison.

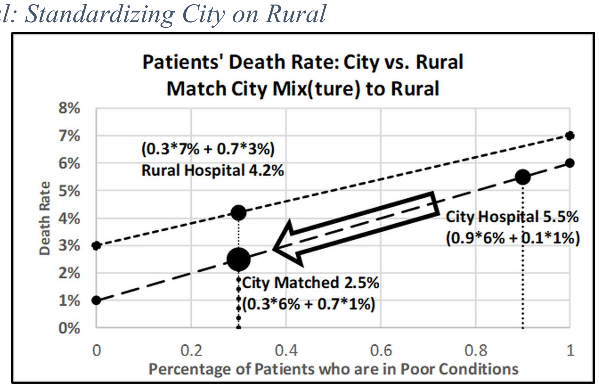
The table on the right standardizes the demographics of the City Hospital to match those at Rural. The items in bold have been adjusted. The result is an adjusted comparison: an "apples and apples" comparison. The Rural Hospital's adjusted patients' death rate (4.2%) is now higher than that of the City Hospital (2.5%): 1.7 points higher.

G: Standardizing on Rural: Graphical Method

The graph on the left is the same as before.

Figure 3: Patients' Death Rates at City and Rural Hospital: Standardizing City on Rural





In the graph on the right, the mixture of City patients was adjusted to match the mixture at Rural. The death rate at City moves down the City weighted average line. The rate at Rural (4.2%) is now higher than the adjusted rate at City (2.5%): a difference of 1.7 percentage points and an 'apples and apples' comparison.

Appendix C: Comments on the ECOTS poster presentation

Dennis Pearl

Thanks for a nice presentation Milo. Examining comparisons by taking potential confounders into account is at the heart of examining the strength of the evidence for most any proposition. I find the most difficult thing to teach is not so much how to do this when I suggest the confounders - but in teaching students to think of potential confounders themselves. Do you have any thought on how to train students in that skill?

> **Milo Schield**
>
> Teaching students to think hypothetically (hypothetical thinking) requires time -- a lot of time, a lot of examples! I don't know of any magic shortcut. That is your primary job in teaching students about confounding. I tell my students, your first word should be "Maybe". Your idea should be plausible, but in this statistics class we are not experts in other fields. Think boldly; think creatively. They've never done this before in any class -- certainly not in a math class. Finally, they don't have to be right. They are just thinking of things that a good study should have considered. If the study could have controlled for them -- but did not -- then they can say, "I'm not convinced." They don't have to prove the study wrong. Finally, they should try and think of two or three confounders -- different kinds of confounders. Hypothetical thinking is an art -- a skill. It takes repetition before it becomes natural. This spring, one of my students said, "This is the hardest class I've taken in college, but I think it will be the most valuable."

John Gabrosek

Thought-provoking. Interesting to see multiple ways to do the adjusting. I like your idea of doing the graphical and then seeing the algebra.

Megan Mocko

Hi Milo. This was a very interesting presentation. I am very curious how the students will do with the combined algebraic and graphical approach. I wonder if initially it will be more difficult for them but it results in them having a better understanding at the end. Thanks,

> **Milo Schield**
>
> You could be right. I hadn't planned to do both until I made this poster. I will try doing both this fall and see how it goes. Normally, I never write about something until I've tried it several times. Yes, I've done each separately. Doing them together will be a new adventure. I'll survey the students afterward and see what they think.

Damien Raftery

Thanks Milo for a very interesting presentation, offering greater depth to the way I currently teach Simpson's Paradox - and interesting to see the combination of visual and algebraic ways combined.