



## An Applet for the Investigation of Simpson's Paradox

[Kady Schneiter](#)

Jürgen Symanzik

Utah State University

*Journal of Statistics Education* Volume 21, Number 1 (2013),  
[www.amstat.org/publications/jse/v21n1/schneiter.pdf](http://www.amstat.org/publications/jse/v21n1/schneiter.pdf)

Copyright © 2013 by Kady Schneiter and Jürgen Symanzik all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the authors and advance notification of the editor.

---

**Key Words:** Simpson's Paradox; Baker-Kramer plot; Graphics.

### Abstract

This article describes an applet that facilitates investigation of Simpson's Paradox in the context of a number of real and hypothetical data sets. The applet builds on the Baker-Kramer graphical representation for Simpson's Paradox. The implementation and use of the applet are explained. This is followed by a description of how the applet has been used in an introductory statistics class and a discussion of student responses to the applet.

### 1. Introduction

Simpson's Paradox was first introduced by [Yule \(1903\)](#) as “the fallacies that may be caused by the mixing of distinct records” (p. 132-133). [Simpson \(1951\)](#), without citing Yule, discussed the interpretation of interaction in contingency tables. In one of his examples, [Simpson \(1951\)](#) observed “that there is a positive association between treatment and survival both among males and among females; but if we combine the tables we again find that there is no association between treatment and survival in the combined population” (p. 241). [Blyth \(1972\)](#) provided an excellent mathematical description of this: For events A, B, and C (and the complements  $B^C$  and  $C^C$ ) it is possible to have  $P(A|B) < P(A|B^C)$  and simultaneously to have  $P(A|BC) \geq P(A|B^C C)$  and  $P(A|BC^C) \geq P(A|B^C C^C)$ . Blyth called this “Simpson's Paradox” (rather than “Yule's Paradox”), and the name has stuck.

[Moore, McCabe, and Craig \(2012\)](#) defined Simpson's Paradox as follows: “An association or comparison that holds for all of several groups can reverse direction when the data are combined to form a single group. This reversal is called **Simpson's paradox**” (p. 145). Unfortunately, this

definition does not fully explain the paradox. On the other hand, the mathematical description by [Blyth \(1972\)](#) is almost impossible to understand for students of introductory statistics courses and for non-statisticians. Thus, explaining Simpson's Paradox in an introductory statistics class is particularly challenging. More unfortunately, as [Lesser \(2002\)](#) pointed out, “some well-known introductory textbooks [...] do not mention Simpson's Paradox at all, some discuss it in a section marked ‘optional’ [...]” (para. 8).

Simpson's Paradox has been observed in many real-life situations. Some examples are in the context of studies on smoking habits and 20-year survival ([Appleton, French, and Vanderpump 1996](#)); in sex bias in graduate student admissions ([Bickel, Hammel, and O'Connell 1975](#)); in the effect of the health care provider in a respiratory disease study ([Bottle and Wakefield 2004](#)); in success rates in removing kidney stones, gender in a psychiatric hospital, and mortality and diabetes ([Julious and Mullee 1994](#)); in the seasonality of human births across the United States ([Knapp 1985](#)); in baseball season batting averages, in classroom tests for two students, and in hiring data ([Lesser 2005](#)); in female employment in the United States during World War II and survival rates for two different cancer treatments ([Mantel 1982](#)); in film ratings ([Moore 2006](#)); in a longitudinal study of growth of children in South Africa ([Morrell 1999](#)); in a multicenter epidemiological study on nosocomial infections ([Reintjes, de Boer, van Pelt, and Mintjes-de Groot 2000](#)); in National Assessment of Educational Progress (NAEP) data ([Terwilliger and Schield 2004](#)); in magazine renewal rates and on income tax rates ([Wagner 1982](#)); in the effect of minority contributions to the SAT score turnaround ([Wainer 1986](#)); in the decline of SAT-Math scores, on statewide mathematics performances, and on statewide death rates ([Wainer 1999](#)); in medical school admission rates ([Wainer and Brown 2004](#)); and in jury composition in New Zealand ([Westbrooke 1998](#)).

We will describe graphical tools and applets that can be used to examine Simpson's Paradox in Section 2. In Section 3, we will discuss the implementation of a new applet (SP applet) for Simpson's Paradox. Advantages of graphical and applet presentations of Simpson's Paradox are discussed in Section 4. In Section 5, we describe how the SP applet has been used in class and in homework assignments. Section 6 summarizes student responses to working with the applet and conclusions are given in Section 7.

## 2. Graphical Representations and Applets for Examining Simpson's Paradox

Several graphical tools have been developed to facilitate understanding of Simpson's Paradox. [Falk and Bar-Hillel \(1980\)](#) and [Wardrop \(1995\)](#) introduced plots that were characterized as “platform scale” representations ([Lesser 2004](#)). [Paik \(1985\)](#) introduced the so-called “circle graph” to visualize Simpson's Paradox. [Lesser \(2001\)](#) applied the platform scale representation, the circle graph, and a few other graphical representations (such as unit square and complex numbers representations) to one data set and discussed the advantages and disadvantages of each of these tools. A comparison of graphical representations of the same data set can also be found in [Lesser \(2005\)](#). [Shapiro \(1982\)](#) introduced plots of conditional probability coordinates to visualize Simpson's Paradox. [Schield \(1999\)](#) developed a graphical technique to illustrate percentage-point difference comparisons that may be used to identify necessary conditions for the paradox. [Rücker and Schumacher \(2008\)](#) showed how to combine scatterplots and line plots into so-called “overlay plots” to demonstrate Simpson's Paradox for two continuous variables

and a third categorical variable. [Jeon, Chung, and Bae \(1987\)](#) produced a graphical representation that was later introduced to the biomedical community by [Baker and Kramer \(2001\)](#). [Wainer \(2002\)](#) suggested that this plot, referred to as the “Baker-Kramer plot” or just “BK plot,” may help “making Simpson’s Paradox clear to the masses” (pg. 60). As [Lesser \(2004\)](#) pointed out, [Tan \(1986\)](#) created a similar trapezoidal plot based on a geometrical construction for the weighted mean ([Hoehn 1984](#)). An applet that facilitates dynamic investigation of Simpson’s Paradox via the Baker-Kramer plot is the central topic of this article.

Interactive statistical applets have been available since the mid-1990’s ([West and Ogden 1998](#)) and a variety of applets and virtual manipulatives for teaching statistics is available on the internet. Collections of statistics applets include the Statistics Online Computational Resource ([SOCR](#)) (n.d.) at UCLA, the Rossman/Chance Applet Collection ([Rossman and Chance n.d.](#)), the Rice Virtual Lab in Statistics (RVLS) ([Lane 2006](#)), National Council of Teachers of Mathematics (NCTM) illuminations ([2006](#)), Shodor Interactive ([SHODOR n.d.](#)), the National Library of Virtual Manipulatives (NLVM) ([Cannon, Dorward, Heal, and Edwards 2007](#)), and Statistics Applets ([Schneiter 2013](#)).

An advantage of using applets for teaching statistical topics is that, in contrast to a general software analysis or investigation package, they are often very specific. Applet interactions can be designed to enable focused investigations of desired content objectives. They can be used to run simulations without having to understand or to modify computer code; with the appropriate applet, changing the parameters of a simulation can be a matter of adjusting a slider or clicking a button.

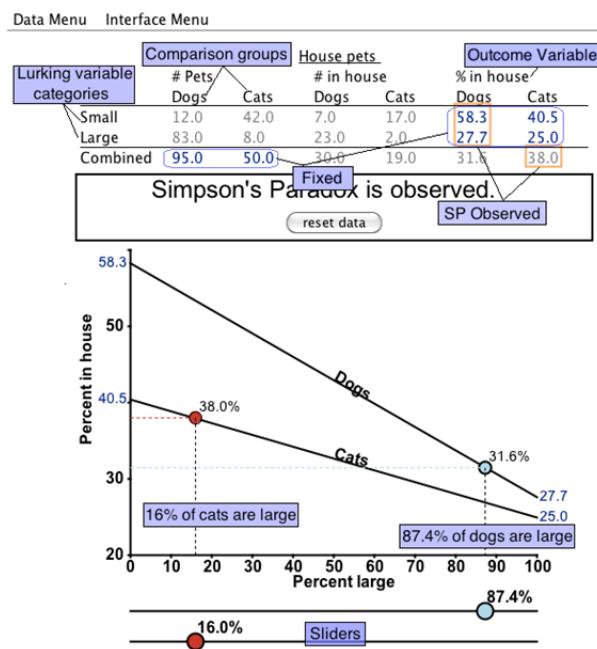
In addition to the applet described in this paper, we are aware of two other applets that facilitate investigation of Simpson’s Paradox. The first, from Cut-the-Knot (<http://www.cut-the-knot.org/Curriculum/Algebra/SimpsonParadox.shtml>), relies on vector algebra to explain the phenomenon. The underlying idea is similar to Lesser’s (2001) complex numbers representation. This explanation and applet is probably most useful for more advanced mathematics and statistics students rather than those in a first statistics course. Another applet (<http://www.uvm.edu/~dhowell/SeeingStatisticsApplets/HeteroSubSamp.html>) addresses Simpson’s Paradox for quantitative data in the context of correlation. This applet uses a scatter plot to show trends in each of two subsets of the data and then in the combined data, an idea similar to [Rücker and Schumacher’s \(2008\)](#) overlay plot. Additional applets and information regarding Simpson’s Paradox can also be found at [Schield’s \(2012\)](#) web page on standardizing. The approaches of these applets to provide an explanation of Simpson’s Paradox are all quite different and therefore might effectively be used complementarily to provide a broad overview of the topic in an upper level course.

### 3. Explanation of the Applet Implementation

We created the SP applet (<http://www.math.usu.edu/~schneit/CTIS/SP/>) to be accessible to students in introductory statistics courses. This applet facilitates dynamic investigation of the BK plot in the context of a number of real and hypothetical data scenarios.

The applet consists of two main components: a table and a plot. When a data set is loaded or reset, the table displays the observed data from the selected study. We will describe the applet features in the context of the hypothetical study comparing the percentages of pet dogs and cats kept in the house by their owners (see [Figure 1](#)). We have used a hypothetical example rather than one of the real data examples included in the applet because the pet data are particularly easy to understand. Since the context is familiar to most people, using the pet example to describe how the applet works helps to keep the explanation as simple as possible.

Columns of the table show the number of subjects in each of two comparison categories (number of dogs, cats), the number of these that meet a certain condition of interest (number of in-house dogs, cats), and the percentage in each category that meet that condition (percentage of in-house dogs, cats). These numbers are displayed in rows for each of two levels (small and large) of a lurking variable (size) and for the combined or overall data. As the user manipulates the applet, the total ('combined') number of observations for each comparison group (total number of dogs, cats in the study) is held fixed as are the percentages having the condition of interest for each comparison group and each classification of the lurking variable (i.e., the percentages of small in-house dogs, large in-house dogs, small in-house cats, and large in-house cats are fixed). The other cells of the table are updated as the user manipulates the plot. The fixed cell values are displayed in blue while the dynamic cell values are in gray.

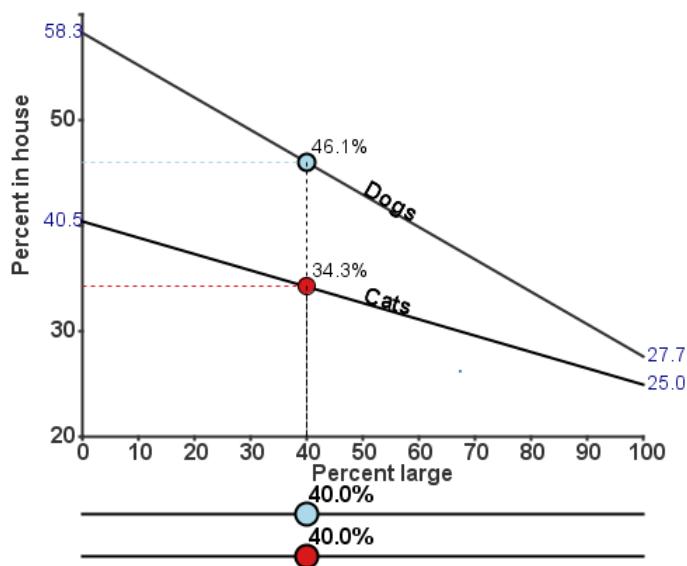


**Figure 1:** SP applet showing the data table in the top and the BK plot at the bottom.

In the plot, the horizontal axis corresponds to the percentage of observations in a specified level of the lurking variable (the percentage of large animals) while the vertical axis corresponds to the percentage of subjects having the condition of interest (in-house). The relationship between these variables is graphed for each comparison category (dogs and cats). There is a marker on each

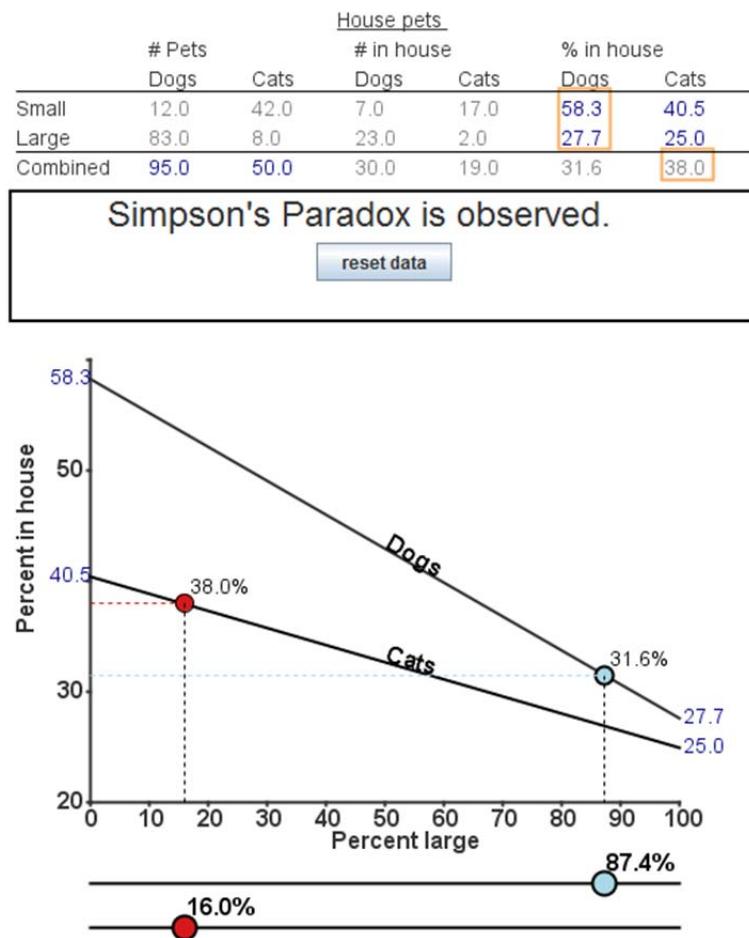
line that can be adjusted to highlight specific percentages. This marked percentage of subjects having the condition of interest is reflected in the ‘combined’ row of the table.

The applet has two interfaces, ‘explore’ and ‘test.’ In the ‘explore’ interface, sliders enable the applet user to adjust the highlighted percentage of observations that have the condition of interest. For example, if the sliders are both set to 40%, the display marks the case in which 40% of both cats and dogs are large. From the plot (Figure 2), we can see that if 40% of cats and dogs were large, 46.1% of dogs and 34.3% of cats would be kept in the house. The table is updated as the user adjusts the sliders (note that manipulating the percentages sometimes results in fractional counts of discrete objects in the table).



**Figure 2:** BK plot showing equal percentages of large pets.

Simpson’s Paradox is observed whenever the percentages of subjects in one comparison group that have the condition of interest are higher for both levels of the lurking variable, while the *overall* percentage of subjects having the condition of interest is higher for the other comparison group. For example, this would occur when there is a higher percentage of small in-house dogs than small in-house cats and a higher percentage of large in-house dogs than large in-house cats but a higher percentage of in-house cats overall. In the table, the relevant percentages are displayed in the last two columns and the higher percentages (comparing between groups, within rows) are boxed (see Figure 3). In the plot, the paradox is observed whenever the marker on the lower line corresponds to a higher percentage of the condition of interest than does the marker on the higher line. For the pet data, the percentage of in-house dogs is greater than the percentage of in-house cats for any specified percentage of large animals. Thus, Simpson’s Paradox occurs when the percentages of large cats and dogs are chosen such that there is a higher percentage of in- house cats than in-house dogs overall. The applet instructions do not include a detailed explanation of what is going on in the plot or table when Simpson’s Paradox occurs; this is intentionally left to the instructor to explain or to allow students to discover.



**Figure 3:** The SP applet when Simpson's Paradox is observed.

In the ‘test’ interface, the applet randomly selects percentages of the indicated level of the lurking variable for each of the comparison groups (e.g., the percentages of large dogs and large cats) and then constructs the table and places the markers in the plot based on these values. The user is asked to determine whether or not Simpson’s Paradox is observed. He/she can choose to view the table and the plot, the plot only, or the table only to make this determination.

#### 4. Advantages of Graphical and Applet Approaches

Simpson’s Paradox is often presented to students with tables as shown below ([Tables 1](#) and [Table 2](#)). These tables display data from the hypothetical pet study introduced in the previous section.

[Table 1](#) shows that cats are more likely than dogs to be kept in the house. However, suppose the pets are also classified according to size: small (weighing 14 pounds or less) and large (weighing over 14 pounds). [Table 2](#) shows the pet data with the added classification. From this table, we see that in each of the size categories, large and small, dogs are more likely to be kept in the house than are cats. Simpson’s Paradox is observed since the relationship between pet type and the percentage kept in the house is reversed when the additional factor, pet size, is taken into

consideration. This tabular presentation can be used to explain that Simpson's Paradox occurs when a relationship between variables is reversed when a lurking variable is included in the tabulation. However, to someone learning about the paradox, the table does not help much for understanding *why* the reversal occurred.

**Table 1:** Percentages of cats and dogs kept in the house.

Pet Type	# in House	# not in House	% in house
<b>Cat</b>	19	31	<b>38.0</b>
Dog	30	65	31.6

**Table 2:** Pet size, pet type and percentage in house.

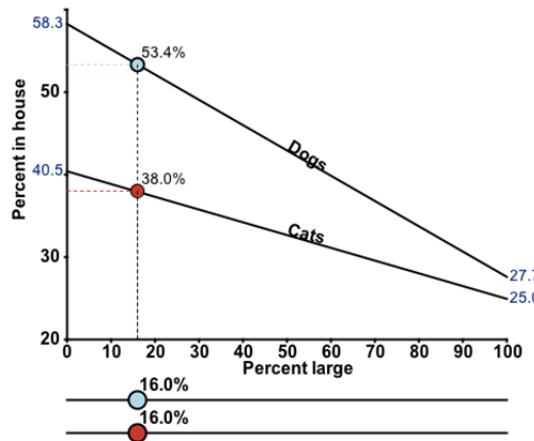
Pet Size	Pet Type	# in house	# not in house	% in house
Small	Cat	17	25	40.5
	<b>Dog</b>	7	5	<b>58.3</b>
Large	Cat	2	6	25.0
	<b>Dog</b>	23	60	<b>27.7</b>

A particular strength of the BK graphical representation is that it provides insight into the cause of Simpson's Paradox. The BK plot of the house pets data from the SP applet ([Figure 3](#)) shows all the percentages that are given in Tables 1 and 2 above. The observed percentages of in-house pets for the combined data are displayed on the markers on the plotted lines: 31.6% for dogs and 38.0% for cats. The percentages of small in-house pets of the two types are given on the vertical axis (considering the percentage of large pets at 0): 58.3% for dogs and 40.5% for cats.

Similarly, the percentages of large in-house pets are displayed at the far right of the graph (considering the percentage of large pets at 100): 27.7% for dogs and 25.0% for cats. The graphical display also shows percentages associated with the lurking variable, pet size, making it possible to see why the paradox occurs. That is, small pets are more likely to be kept inside than large pets (both lines have negative slopes, i.e., the in-house percentage decreases as the percentage of large pets increases) and cats are more likely to be small (the observed percentage of large cats was 16.0% compared to 87.4% of dogs). Furthermore, because the line corresponding to dogs is everywhere above the cat line, the plot shows that, based on these data, if the percentage of large pets were the same for dogs and cats, regardless of the actual percentage, more dogs would be in-house pets.

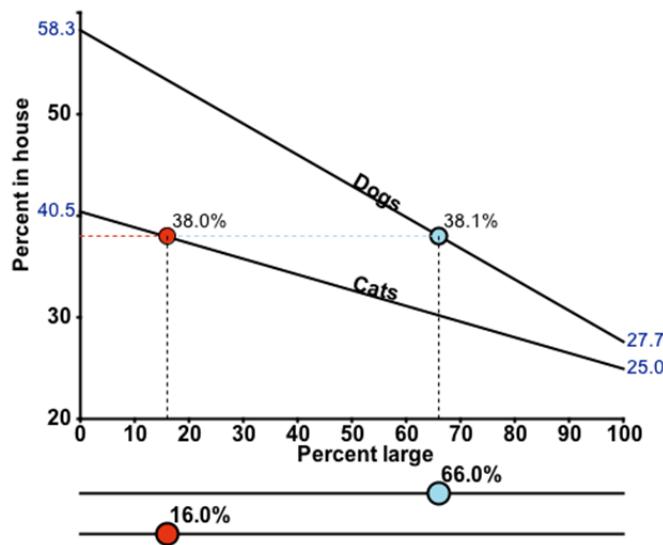
The advantage of an applet over the static graphical display is that it allows the user to adjust parameters of the data. With our SP applet, the user can change the percentages associated with the lurking variable. This facilitates observations such as this: if the slider corresponding to the percentage of large dogs is moved to 16.0% and the other slider is unmoved, so that there is the

same percentage of large dogs and large cats, more dogs would be kept in the house (see [Figure 4](#)).



**Figure 4:** Equal percentages of large pets, shown in the BK plot.

The sliders can also be manipulated to ascertain the conditions under which Simpson's Paradox will occur. For example, if the percentage of large cats is held at 16.0%, for what percentages of large dogs would Simpson's Paradox not occur? By moving the slider until the percentage of in-house pets is about the same for cats and dogs, we can see that Simpson's Paradox is not observed if between 0% and 66.0% of dogs are large (see [Figure 5](#)).



**Figure 5:** For what percentages does Simpson's Paradox not occur?

In addition to the “House Pets” data, the SP applet includes the following data sets: “Baker-Kramer” data ([Wainer 2002](#)) compares survival for patients receiving two different treatments; “Berkeley Admissions” data ([Bickel et al. 1975](#)) compares university admissions rates for male and female applicants; “Florida Death Penalty” data ([Radelet and Pierce 1991](#)) compares percentages of death penalty sentences for black and white defendants in murder trials;

“Airlines” data ([Moore, McCabe and Craig 2009](#)) compares percentages of delayed flights for two airlines; “Civil Rights Act” data ([Matheson 1999](#)) compares percentages of Democrats and Republicans in favor of the act; and “Smoker Survival” data ([Vanderpump et al. 1996](#)) compares percentages of smokers and non-smokers surviving at a 20 year follow up. The “Civil Rights Act,” “Smoker Survival,” and “House Pets” data sets were added in response to user feedback as described in Section 6.

## 5. Class Use of the SP Applet

The SP applet was class-tested in an “Introduction to Statistical Methods” course (Stat 2000) in the Spring 2012 semester, taught at Utah State University by the second author of this article. The course was taught in a large lecture format with additional recitation sessions each week. The course closely followed the textbook from [Moore, McCabe, and Craig \(2012\)](#) and the PowerPoint slides provided by the publisher. A series of applets from the first author’s “Statistics Applets” web page (<http://www.math.usu.edu/~schneit/CTIS/>) have been used to demonstrate some of the concepts. At the time Simpson’s Paradox was introduced to the students, they had already worked with the applets for “Mean and Median,” “Standard Deviation,” and “Correlation” from this web site in their homework assignments. Other applets from this collection were used later in the semester. For more information on these applets, see [Kohler, Schneiter, and Thatcher \(2010\)](#) and [Schneiter \(2008, 2011\)](#). The Stat 2000 course web page that includes the syllabus, homework assignments, and additional information is freely accessible at [http://www.math.usu.edu/~symanzik/teaching/2012\\_stat2000/stat2000.html](http://www.math.usu.edu/~symanzik/teaching/2012_stat2000/stat2000.html).

Overall, 73 students were enrolled in this class. Seventy students were still present at the date of the Final Exam. Students were predominantly white, with a few of Hispanic, Asian, and Arabic descent. Gender was split almost evenly between males and females. Students were all undergraduates and from all class ranks (freshman, sophomore, junior, and senior) with sophomores (about 37%) the dominant rank. Because this is a service course required in many departments, students came from 27 different majors. Animal/Dairy/Veterinary Science (21 students, or about 28%), Human Movement Science (8 students, or about 11%), Environmental Studies and Wildlife Science (each 6 students, or each about 8%) and Recreation Resource Management (5 students, or about 7%) were the most frequent majors. The remaining 22 majors were represented by one or at most two students. Two students had double majors and only one student majored in Statistics.

Simpson’s Paradox was introduced in class and the SP applet was introduced using the “Airlines” data. As mentioned above, this data set contains information on the percentage of delayed flights for two airlines, Alaska Airlines (AA) and America West Airlines (AW). The lurking variable is the origination city of the flight; in particular, the user can adjust the percentage of flights originating in Phoenix for each airline. Both key features of the applet, the data table and the BK plot, were explained in detail and several examples were shown to demonstrate whether Simpson’s Paradox is present or not present in a given situation. Thereafter, two iClicker questions were asked: Each question dealt with the “Airlines” data available in the applet. One focused on interpreting the table (Question 12, [Figure 6](#)) and the other on interpreting the plot (Question 13, [Figure 7](#)). In both questions, students were shown two

situations and asked to decide in which cases SP was observed. The results are displayed in [Table 3](#).

After students responded to each question, the instructor explained the correct answer and answered related questions. During the semester, a total of 39 iClicker questions were asked, many related to applets. Question 13 had by far the highest percentage (98%) of correct answers throughout the entire semester.

**Table 3:** Percentages of correct answers for iClicker questions 12 and 13.

Question	Correct Response	Number of respondents		Number of correct answers		Percentage of correct answers
12	C	58		48		82.8%
13	B	58		57		98.3%

Flight Delays for Alaska Airlines and America West						
	# Flights		# Delayed		% Delayed	
	AA	AW	AA	AW	AA	AW
Combined	<b>792.0</b>	<b>6066.0</b>	71.4	621.6	9.0	10.2
Los Angeles	514.8	2183.8	57.1	315.0	<b>11.1</b>	<b>14.4</b>
Phoenix	277.2	3882.2	14.3	306.6	<b>5.2</b>	<b>7.9</b>
Flight Delays for Alaska Airlines and America West						
	# Flights		# Delayed		% Delayed	
	AA	AW	AA	AW	AA	AW
Combined	<b>792.0</b>	<b>6066.0</b>	64.8	479.0	8.2	7.9
Los Angeles	403.9	0.0	44.8	0.0	<b>11.1</b>	<b>14.4</b>
Phoenix	388.1	6066.0	20.0	479.0	<b>5.2</b>	<b>7.9</b>

Is Simpson's Paradox observed?

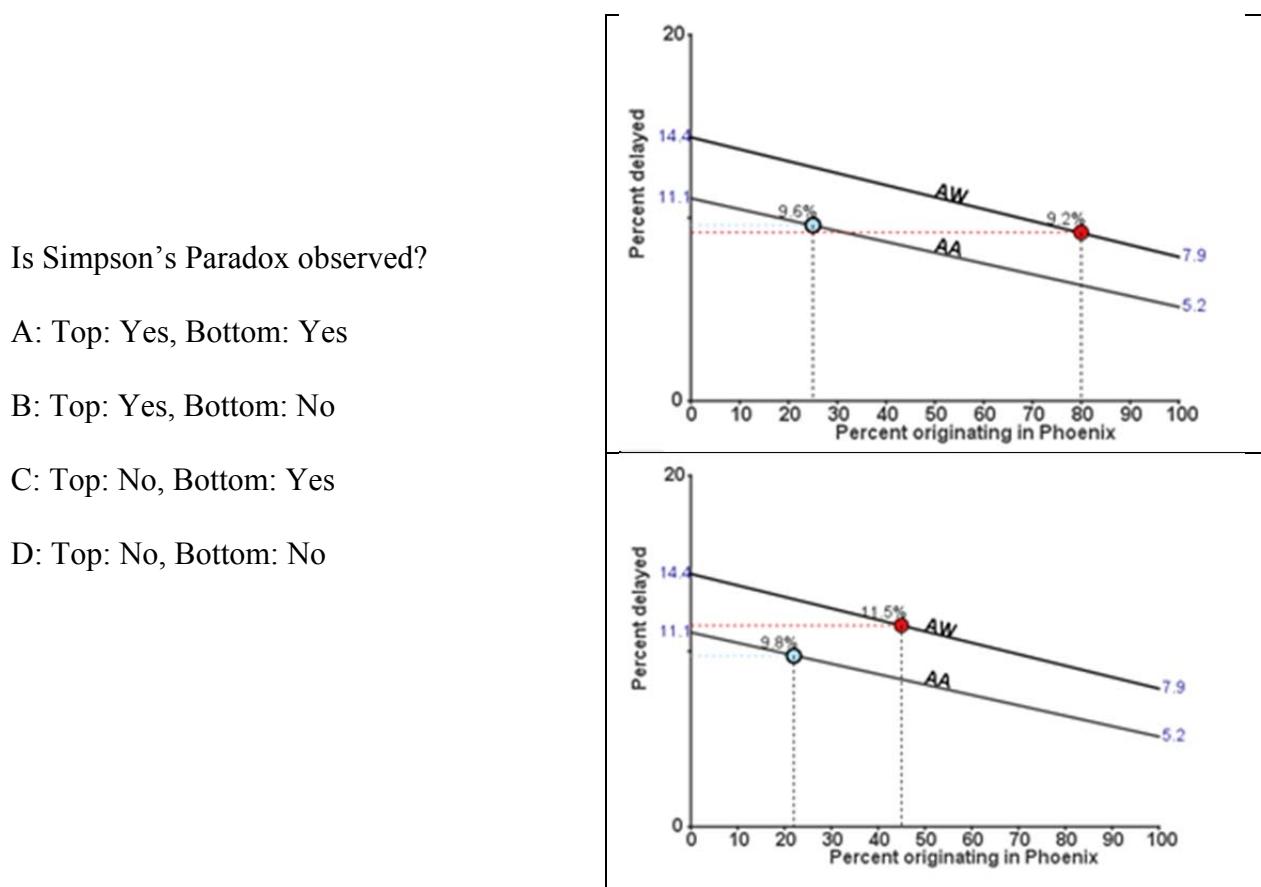
A: Top: Yes, Bottom: Yes

B: Top: Yes, Bottom: No

C: Top: No, Bottom: Yes

D: Top: No, Bottom: No

**Figure 6:** iClicker Question 12.

**Figure 7:** iClicker Question 13.

A subsequent homework assignment (Homework 6, see Appendix) included questions related to Simpson's Paradox and required students to work with the SP applet. The relevant question consisted of 5 parts (a-e), the first three of which (a-c) required students to indicate whether or not Simpson's Paradox would be observed under specified conditions. The remaining two parts required students to supply the conditions under which the paradox would be observed. Sixty-three students completed the homework assignment. On parts a-c, three students did not indicate whether or not Simpson's Paradox was observed. Of the remaining 60 students, 53 correctly indicated when Simpson's Paradox was observed and five of these added an explanation that SP can't occur when there is only one category of the lurking variable. Of the seven students who did not correctly identify Simpson's Paradox, only one recorded the correct percentages but drew the wrong conclusions. One other gave a response that showed that he/she didn't understand the question at all and the remaining five had various errors relating to reading the table or identifying the relevant percentages.

No data were collected from students to ascertain whether they were able to generalize their understanding of Simpson's Paradox to contexts beyond the applet. This would be an important avenue for follow-up research. Additionally, we would like to engage in further study to compare

understanding of the paradox between students who were and students who were not exposed to the applet.

## 6. Student Response

Following collection of the homework assignment, students were asked to complete a short survey regarding the SP applet. Students' answers were anonymous and voluntary. The 63 students who completed the homework assignment were asked to complete the survey; 60 students did so.

Students responded to several prompts that asked them to indicate their level of agreement (strongly disagree, disagree, agree, strongly agree) with statements related to whether they enjoyed using the applet or found it useful. The statements and the percentages that either agreed or strongly agreed with these are indicated in the table below.

**Table 4:** Summary of students' responses to applet use survey.

Statement	% agree or strongly agree
I understand more about the topic than I did before I used the applet.	78
Overall I enjoyed using the applet.	70
I would like to use other applets like this or involving other concepts and subjects.	78
The directions for the applet were easy to follow.	80
I used the applet to look at things besides those specifically asked for in the assignment.	50

The results indicate that the majority of students found working with the applet to be useful and enjoyable. The last statement included above "I used the applet to look at things besides those specifically asked for in the assignment" was included to gauge whether or not the applet might stimulate users to do independent exploration. Since the assignment was completed out of class time, there was ample time for students to do so if they desired. No instructions or incentives were given to do this. Fifty percent of the students reported that they did use the applet to look at things that were not required for the assignment.

Students were also asked to comment on what could be done to make the applet more useful or easier to use. The three prompts they were asked to address were "If I could change the applet to make it more effective I would change...," "If I could change the applet to make it easier to use I would change...," and "I would have learned more about the topic if the applet had...." As there was considerable overlap in the responses to these questions, we will consider them as a group. Of the 90 comments and suggestions given, 30 (33%) were that nothing should be changed, the applet worked well, etc. Another 33 were irrelevant or did not give suggestions (e.g., "use Excel" or "it's really easy if you teach it clearly"). Of the remaining 27 comments, four suggested that more data sets be added, five asked that the table presentation be improved, four were concerned with slider usability, two requested further instructions in the applet interface itself, three wanted more explanation of when Simpson's Paradox is observed, and eight asked for more detail in the accompanying instructions.

In response to these comments, we added three data sets (“Civil Rights Act,” “Smoker Survival” and “House Pets”) and improved the table organization by rearranging rows and differentiating between static and dynamic text with color. We have also included in the instructions a detailed description of how the applet works and what can be observed from it. Some of the student comments regarding the instructions would more appropriately be addressed by an instructor using the applet. For instance, the instructions do not say explicitly what is going on in the plot or the table when Simpson’s Paradox is observed. Some instructors might choose to point this out explicitly (i.e., “Simpson’s Paradox is observed when the higher percentage in the *Combined* row of the table corresponds to a different comparison group than do the higher percentages in the rows corresponding to lurking variable classification” or “Simpson’s Paradox is observed when the marker on the lower line is higher than the marker on the upper line”) but others might wish to leave it to their students to discover this.

A continuing shortcoming of the table display is that, for certain settings of the sliders in the plot, fractional values are given for counts of discrete objects. This is due to the nature of the sliders and the plot data which dictate how the table changes. As this does not affect the interpretation of Simpson’s Paradox, we recommend that concerned instructors use this as a discussion point with their students or that they focus their discussions on the percentages rather than the counts. We hope to be able to address this issue in future updates.

## 7. Conclusion

As described in this paper, the SP applet was used for in-class demonstrations during which the students had access to iClickers. This is a useful presentation scenario because it enables the instructor to receive immediate feedback from students; however, iClickers are by no means necessary for effective use. We recommend that the applet be introduced by the instructor in class and that students are then either given computer time at school to work with the applet or are asked to work with the applet for a homework assignment. Allowing students time to work with the applet on their own or in small groups will enable them to make connections that may not be obvious through direct instruction. Furthermore, during this time, students may explore beyond what they would do in a more constrained classroom environment. As indicated in [Table 4](#) above, 50% of the student who used the applet said that they used the applet to look at things beyond those required by the assignment. There was no external incentive given for students to do this but, as there were also no time constraints, students were able to explore data sets or scenarios of interest to themselves.

The details of interpretation of the applet displays are not stated explicitly in the applet instructions. Under ‘Points to Ponder’ in the instructions, the following prompts are given:

- Describe what you see in the table when Simpson’s Paradox is observed.
- Describe what you see in the plot when Simpson’s Paradox is observed.
- When adjusting the sliders, at what point does Simpson’s Paradox appear?
- What do the conditions you observe in the applet tell you about what’s going on in the data when Simpson’s Paradox occurs?

These are included without further detail leaving it to the instructor to either supply the details or to allow students to use the applet to discover the connections on their own. The last prompt

above is especially important. It is possible that students will be able to identify when Simpson's Paradox occurs by looking at the table or plot but not make the connection to what is actually going on in the data. For instance, a student might recognize that Simpson's Paradox is observed 'when the marker on the lower line in the plot is higher than the marker on the higher line' without being able to explain what is going on in the data that causes this to occur. Thus instructors who use the applet are encouraged to discuss the connection with their students, ideally after students have been given time to consider this on their own.

Students have responded favorably to working with the applet indicating that they found it enjoyable to use as well as instructive. The vast majority of students indicated that they felt that the applet positively influenced their understanding of Simpson's Paradox and that they would like to see more applets in their instruction. Half were motivated to explore Simpson's Paradox beyond the requirements of the assignment.

The BK plot, one of the existing graphical displays for examining Simpson's Paradox, can help students to observe *why* the paradox occurs, not just *when* it occurs. Looking at the BK plot dynamically via the SP applet facilitates deeper exploration by enabling students to adjust parameters, to formulate and test conjectures, and to test their understanding of Simpson's Paradox.

---

## Appendix

### Homework #6

#### **Exercise 2) Simpson's Paradox** (5 points):

Revisit the "Simpson's Paradox" applet at <http://www.math.usu.edu/~schneit/CTIS/SP/>. In the "Data Menu", select the "Berkeley Admissions" data set. These data have been extensively discussed in the literature, see for example the section called the "Berkeley gender bias case" at [http://en.wikipedia.org/wiki/Simpson's\\_paradox](http://en.wikipedia.org/wiki/Simpson's_paradox). Note that numbers from Departments A & B have been combined into the "Easy" component and numbers from Departments C to F have been combined into the "Hard" component in the applet (and wikipedia seems to list an incorrect number as the number of male applicants in Department F typically is listed as 373 and not as 272).

Based on the applet, answer the following questions:

- (a) If 0% of women and 0% of men would apply to harder admittance majors, what would be the overall percentage of men and the overall percentage of women that would be admitted? Is Simpson's Paradox observed in this situation?
- (b) If 100% of women and 100% of men would apply to harder admittance majors, what would be the overall percentage of men and the overall percentage of women that would be admitted? Is Simpson's Paradox observed in this situation?
- (c) If 75% of women and 25% of men would apply to harder admittance majors, what would be the overall percentage of men and the overall percentage of women that would be admitted? Is Simpson's Paradox observed in this situation?
- (d) If 60% of women would apply to harder admittance majors, about what percentage of men would have to apply to the harder admittance majors so that Simpson's Paradox barely is observed. You should narrow this percentage down to a 1%-interval, say between 8% to 9% or between 92% and 93% (just giving two incorrect answers here).
- (e) If 15% of men would apply to harder admittance majors, about what percentage of women would have to apply to the harder admittance majors so that Simpson's Paradox barely is observed. You should narrow this percentage down to a 1%-interval, say between 8% to 9% or between 92% and 93% (just giving two incorrect answers here).

#### **Solutions for Exercise 2:**

- (a) If 0% of women and 0% of men would apply to harder admittance majors: 79.7% of women and 62.4% of men would be omitted; Simpson's Paradox is **not** observed.
- (b) If 100% of women and 100% of men would apply to harder admittance majors: 26.5% of women and 25.6% of men would be omitted; Simpson's Paradox is **not** observed.

(c) If 75% of women and 25% of men would apply to harder admittance majors: about 39.8% of women and about 53.2% of men would be omitted; Simpson's Paradox **is observed.**

(d) If 60% of women would apply to harder admittance majors, between 39% and 40% of men would have to apply to harder admittance majors so that Simpson's Paradox is barely observed.

(e) If 15% of men would apply to harder admittance majors, between 42% and 43% of women would have to apply to harder admittance majors so that Simpson's Paradox is barely observed.

---

## References

Appleton, D.R., French, J.M., and Vanderpump, M.P.J. (1996), "Ignoring a Covariate: An Example of Simpson's Paradox." *The American Statistician*, 50(4), 340-341.

Baker, S.G., and Kramer, B.S. (2001), "Good for Women, Good for Men, Bad for People: Simpson's Paradox and the Importance of Sex-Specific Analysis in Observational Studies." *Journal of Women's Health & Gender-Based Medicine*, 10(9), 867-872.

Bickel, P.J., Hammel, E.A., and O'Connell, J.W. (1975), "Sex Bias in Graduate Admissions: Data from Berkeley." *Science*, 187(4175), 398-404.

Blyth, C.R. (1972), "On Simpson's Paradox and the Sure-Thing Principle." *Journal of the American Statistical Association*, 67(338), 364-366.

Bottle, A., and Wakefield, J. (2004), "Controlling for Provider of Treatment in the Modelling of Respiratory Disease Risk near Cokeworks." *Statistics in Medicine*, 23(20), 3139-3158.

Cannon, L., Dorward, J., Heal, R., and Edwards, L. (2007), "National Library of VirtualManipulatives (NLVM)." <http://nlvm.usu.edu/en/nav/vlibrary.html>. Accessed February 3, 2013.

Falk, R., and Bar-Hillel, M. (1980), "Magic Possibilities of the Weighted Average." *Mathematics Magazine*, 53(2), 106-107.

Hoehn, L. (1984), "A Geometrical Interpretation of the Weighted Mean." *The College Mathematics Journal*, 15(2), 135-139.

Jeon, J.W., Chung, H.Y., and Bae, J.S. (1987), "Chances of Simpson's Paradox." *Journal of the Korean Statistical Society*, 16(2), 117-127.

Julious, S.A., and Mullee, M.A. (1994), "Confounding and Simpson's Paradox." *British Medical Journal*, 309(6967), 1480-1481.

Knapp, T.R. (1985), "Instances of Simpson's Paradox." *The College Mathematics Journal*, 16(3), 209-211.

Kohler, B.R., Schneiter, K., and Thatcher, K. (2010), "Combining Professional Development with Resource Development: Design of Integrated Technology Lessons in Statistics." *Journal of the Research Center for Educational Technology*, 6(2), 31-47.

Lane, D.M. (2006), Rice Virtual Lab in Statistics (RVLS) <http://onlinestatbook.com/rvls.html>. Accessed February 3, 2013.

Lesser, L.M. (2001), "Representations of Reversal: An Exploration of Simpson's Paradox." In AA Cuoco, FR Curcio (eds.), *The Roles of Representation in School Mathematics*, pp. 129-145. National Council of Teachers of Mathematics, Reston, VA. Available online at <http://www.statlit.org/pdf/2001LesserNCTM.pdf>.

Lesser, L.M. (2002), "Letter to the Editor." *Journal of Statistics Education*, 10(1). [http://www.amstat.org/publications/jse/v10n1/lesser\\_letter.html](http://www.amstat.org/publications/jse/v10n1/lesser_letter.html).

Lesser, L.M. (2004), "Wainer, H., and Brown, L. M. (2004), "Two Statistical Paradoxes in the Interpretation of Group Differences: Illustrated with Medical School Admission and Licensing Data," *The American Statistician*, 58, 117-123: Comment by Lesser." *The American Statistician*, 58(4), 362.

Lesser, L.M. (2005), "Illumination through Representation: An Exploration Across the Grades." *The Statistics Teacher Network*, 66, 3-5. <http://www.amstat.org/education/stn/pdfs/stn66.pdf>.

Mantel, N. (1982), "Simpson's Paradox in Reverse." *The American Statistician*, 36(4), 395.

Matheson, V.A. (1999, June 22), "A Closer Look at the Civil Rights Act Tally." *The Washington Times*, A18.

Moore, D.S., McCabe, G.P., and Craig, B.A. (2009), *Introduction to the Practice of Statistics* (Sixth Edition). W. H. Freeman and Company, New York, NY.

Moore, D.S., McCabe, G.P., and Craig, B.A. (2012), *Introduction to the Practice of Statistics* (Seventh Edition). W. H. Freeman and Company, New York, NY.

Moore, T.L. (2006), "Paradoxes in Film Ratings." *Journal of Statistics Education*, 14(1). <http://www.amstat.org/publications/jse/v14n1/datasets.moore.html>.

Morrell, C.H. (1999), "Simpson's Paradox: An Example from a Longitudinal Study in South Africa." *Journal of Statistics Education*, 7(3). <http://www.amstat.org/publications/jse/secure/v7n3/datasets.morrell.cfm>.

National Council of Teachers of Mathematics (NCTM) (2006), "Illuminations."  
<http://illuminations.nctm.org/>. Accessed February 3, 2013.

Paik, M. (1985), "A Graphic Representation of a Three-Way Contingency Table: Simpson's Paradox and Correlation." *The American Statistician*, 39(1), 53-54.

Radelet, M.L., and Pierce, G.L. (1991), "Choosing Those Who Will Die: Race and the Death Penalty in Florida." *Florida Law Review*, 43(1), 1-34.

Reintjes, R., de Boer, A., van Pelt, W., and Mintjes-de Groot, J. (2000), "Simpson's Paradox: An Example from Hospital Epidemiology." *Epidemiology*, 11(1), 81-83.

Rossman, A.J., and Chance, B.L. (n.d.), "Rossmann/Chance Applet Collection."  
<http://www.rosmachance.com/applets/index.html>. Accessed February 3, 2013.

Rücker, G., Schumacher, M. (2008), "Simpson's Paradox Visualized: The Example of the Rosiglitazone Meta-Analysis." *BMC Medical Research Methodology*, 8(34).  
<http://www.biomedcentral.com/1471-2288/8/34>.

Schield, M. (1999), "Simpson's Paradox and Cornfield's Conditions." In 1999 Proceedings of the Section on Statistical Education, pp. 106-111. American Statistical Association, Alexandria, VA.

Schield, M. (2012), "Standardizing." <http://www.statlit.org/Standardizing.htm>. Accessed February 3, 2013.

Schneiter, K. (2008), "Two Applets for Teaching Hypothesis Testing." *Journal of Statistics Education*, 16(3). <http://www.amstat.org/publications/jse/v16n3/schneiter.html>.

Schneiter, K. (2011), "Buffon's Coin Problem and Beyond." *Teaching Statistics*, 33(2), 34-37.

Schneiter, K. (2013), "Statistics Applets." <http://www.math.usu.edu/schneit/CTIS>. Accessed February 3, 2013.

Shapiro, S.H. (1982), "Collapsing Contingency Tables - A Geometric Approach." *The American Statistician*, 36(1), 43-46.

SHODOR: A National Resource for Computational Science Education (n.d.), "Interactivate: Activities." <http://www.shodor.org/interactivate/activities/>. Accessed February 3, 2013.

Simpson, E.H. (1951), "The Interpretation of Interaction in Contingency Tables." *Journal of the Royal Statistical Society, Series B*, 13(2), 238-241.

SOCR: Statistics Online Computational Resource (n.d.), <http://www.socr.ucla.edu/SOCR.html>. Accessed February 3, 2013.

Tan, A. (1986), "A Geometric Interpretation of Simpson's Paradox." *The College Mathematics Journal*, 17(4), 340-341.

Terwilliger, J., and Schield, M. (2004), "Frequency of Simpson's Paradox in NAEP Data." American Educational Research Association. Presented at AERA Summer Conference, <http://www.statlit.org/pdf/2004TerwilligerSchieldAERA.pdf>.

Vanderpump, M.P.J., Tunbridge, W.M.G., French, J.M., Appleton, D., Bates, D., Clark, F., Grimley Evans, J., Rodgers, H., Tunbridge, F., and Young, E.T. (1996), "The Development of Ischemic Heart Disease in Relation to Autoimmune Thyroid Disease in a 20-Year Follow-up Study of an English Community." *Thyroid*, 6(3), 155-160.

Wagner, C.H. (1982), "Simpson's Paradox in Real Life." *The American Statistician*, 36(1), 46-48.

Wainer, H. (1986), "Minority Contributions to the SAT Score Turnaround: An Example of Simpson's Paradox." *Journal of Educational Statistics*, 11(4), 239-244.

Wainer, H. (1999), "Simpson's Paradox." *Chance*, 12(2), 43-44.

Wainer, H. (2002), "The BK-Plot: Making Simpson's Paradox Clear to the Masses." *Chance*, 15(3), 60-62.

Wainer, H., and Brown, L.M. (2004), "Two Statistical Paradoxes in the Interpretation of Group Differences: Illustrated with Medical School Admission and Licensing Data." *The American Statistician*, 58(2), 117-123.

Wardrop, R.L. (1995), "Simpson's Paradox and the Hot Hand in Basketball." *The American Statistician*, 49(1), 24-28.

West, R.W., and Ogden, R.T. (1998), "Interactive Demonstrations for Statistics Education on the World Wide Web." *Journal of Statistics Education*, 6(3).

<http://www.amstat.org/publications/jse/v6n3/west.html>.

Westbrooke, I. (1998), "Simpson's Paradox: An Example in a New Zealand Survey of Jury Composition." *Chance*, 11(2), 40-42.

Yule, G.U. (1903), "Notes on the Theory of Association of Attributes in Statistics." *Biometrika*, 2(2), 121-134.

---

Kady Schneiter, Ph.D.  
Utah State University  
Department of Mathematics and Statistics

3900 Old Main Hill  
Logan, UT 84322  
[kady.schneiter@usu.edu](mailto:kady.schneiter@usu.edu)  
(435) 797-2820  
(435) 797-1822 (fax)

Jürgen Symanzik, Ph.D.  
Utah State University  
Department of Mathematics and Statistics  
3900 Old Main Hill  
Logan, UT 84322  
[symanzik@math.usu.edu](mailto:symanzik@math.usu.edu)

---

[Volume 21 \(2013\)](#) | [Archive](#) | [Index](#) | [Data Archive](#) | [Resources](#) | [Editorial Board](#) | [Guidelines for Authors](#) | [Guidelines for Data Contributors](#) | [Guidelines for Readers/Data Users](#) | [Home Page](#) |  
[Contact JSE](#) | [ASA Publications](#)