NBER WORKING PAPER SERIES

CATCHING CHEATING STUDENTS

Steven D. Levitt
Ming-Jen Lin

Working Paper 21628
http://www.nber.org/papers/w21628

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2015

We thank the professor teaching this course for providing data. We thank Eric Andersen, Dai-Rong Chen, Yue-Shuan Chun, Jason Lai, and Dhiren Patki for their excellent research assistance. We thank Ministry of Science and Technology, Taiwan for financial support (103-2628-H-002 -001). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Catching Cheating Students
Steven D. Levitt and Ming-Jen Lin
NBER Working Paper No. 21628
October 2015
JEL No. K42

ABSTRACT

We develop a simple algorithm for detecting exam cheating between students who copy off one another's exam. When this algorithm is applied to exams in a general science course at a top university, we find strong evidence of cheating by at least 10 percent of the students. Students studying together cannot explain our findings. Matching incorrect answers prove to be a stronger indicator of cheating than matching correct answers. When seating locations are randomly assigned, and monitoring is increased, cheating virtually disappears.

Steven D. Levitt
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
slevitt@midway.uchicago.edu

Ming-Jen Lin
Department of Economics
National Taiwan University
No. 1, Sec. 4, Roosevelt Rd.
Taipei 10617, Taiwan
mjlin@ntu.edu.tw

Introduction

Student cheating is a perennial issue. In recent years, 70 students in a New York City top public high school were caught using smart phones to cheat on state exams (Baker 2012), and cheating scandals have rocked Harvard, Stanford, Dartmouth, and the Air Force Academy, just to name a few.[1]

These well publicized scandals are only the tip of the iceberg. McCabe (2005) surveyed 8,000 college students in the U.S. and Canada, finding that 11% of them admit to "copying from another student on an exam without their knowledge," 10% say they have "helped someone else cheat on test", and 9% acknowledge copying from another student "with their knowledge."

In spite of this apparent widespread cheating, there has been little academic attention devoted to the detection of cheating.[2] Zitzewitz (2012) surveys more than 100 pages in the emerging field of forensic economics, not one of which addresses student cheating.[3]

This paper develops an algorithm for identifying cases of students copying off one another's exam answers. We test this algorithm using data from a course taught at a top university in which the professor suspected cheating may have occurred. We find compelling evidence of cheating involving at least 10% percent of the class's 242 students on a midterm exam. When seating positions were randomly assigned and monitoring was increased for the final exam, almost all evidence of cheating disappears. We are able to rule out that the observed correlations in answers across students who voluntarily sit next to each other is due to studying

[1] Harvard University admitted that "about 125 students might have worked in groups on a take-home final exam." Roughly seventy students were forced to withdraw from the University (Perez-Pena 2013). Similar numbers of students were involved at Darmouth and the Air Force Academy (Frosch 2015, Associated Press 2015a). In March 2015, Stanford University Provost John Etchemendy sent a letter to the faculty expressing concerns over allegations of widespread cheating (Associated Press 2015b).

[2] Levitt and Jacob (2003) develop a set of tools for analyzing *teacher* cheating, some of which we build upon in this paper.

[3] Organizations such as the Educational Testing Service (ETS), provider of the SAT and GRE exams, no doubt have developed techniques for detecting cheating, but to the best of our knowledge, these tools have never been made publicly available.

together, as opposed to cheating on the exam because of an unusual experiment carried out an advance of the final exam. Students seated themselves voluntarily, with the expectation that the seats they chose would be the ones in which they would take the exam. These seating choices were recorded. Prior to the actual test, however, students were randomly reassigned to different seats. Thus, we are able to observe the patterns in correlations among students who wanted to sit together, but then were not allowed to.

The remainder of the paper is structured as follows. Section I describes the background of the cheating incident that we analyze. Section II presents a simple, reduced form regression analysis of the cheating patterns. Section III develops and implements a more systematic algorithmic approach to the problem and also considers alternative explanations for correlations like studying together. Section IV concludes.

Section I: The Cheating Incident and the Data

In spring 2012, 242 students registered in an introductory natural sciences course at a top American university. The course had three midterms and a final exam.[4] All of the exams were multiple choice, with four possible answers per question. Students recorded their answers on Scantron sheets. There was no punishment for incorrect guesses, i.e. a wrong answer yielded no points, as did leaving the question blank. The average percentage correct on the exams fell in the range of 75 to 85.

The first three midterms were held in a classroom with nearly every seat occupied. A single TA proctored the exam. During the third midterm, a student came to the TA reporting suspicions that another student had been cheating. The proctor did not take any action regarding the cheating during the midterm, but did report this information to the professor after the exam. In response, the professor sent out an email, saying that "cheating is morally wrong," and

[4] Students were required to take only two of the three midterms. The midterms had 50 questions each; the final exam had 80 questions.

encouraged students to admit their wrongdoing. No students voluntarily came forward, although a second student said she had also witnessed cheating. This prompted the professor to once again call for student confessions, bolstered by the threat that he was going to contact us -- the authors of this paper -- and have us catch the cheaters.[5] Again, no one came forward, and the professor did indeed contact us two weeks before the final.

The data available to us include students' answers to each question on all four exams, as well as seating information for the third midterm and the final exam. In addition, we were able to carry out an unusual experiment involving the final exam. Students entered the exam room and selected their own seats, as was usual practice. These seating choices were recorded. Before the exam actually began, however, students were shuffled into randomly assigned seats for the test taking. This provides us with the opportunity to observe correlations in answer patterns among students who would have liked to sit together (and perhaps studied together), but were then separated.

Further steps were taken to prevent cheating on the final exam. Unlike the first three exams where only one TA served as a proctor, four proctors were present during the final. Finally, the professor created two different versions of the final exam; the questions in both versions were the same, but the order in which they were asked was different. Students randomly received one of the two different versions of the test.

Section II: Reduced form detection of cheating

We begin our analysis of possible cheating with a simple reduced form regression approach in which the unit of observation is a *pair* of students on a particular exam. For each possible pair of students, we calculate the number of questions for which those students gave the same *correct* answer and also the number of questions for which those students gave the same

[5] In his email, the professor warned the students "[they are] extremely good at catching cheating if you have read *Freakonomics*." Apparently, none of the cheaters had read Freakonomics.

incorrect answer. If the number of common answers is high, this may be an indication of cheating, although of course there may be other explanations as well.

Copying from a student to one's left or right is the simplest way to cheat. Thus, the key variable of interest in the regression is an indicator variable that is equal to one if the students sit next to each other and takes the value of zero otherwise.[6] Given the room setup, it is difficult to cheat from two seats away with an empty seat in the middle, but triplets of students might effectively cheat. Therefore, we include an indicator for students whose seating pattern is "student 1 – empty seat – student 2," as well as an indicator if the seating pattern is "student 1 – some other student – student 2." Cheating from in front or behind another student was not easily done, so we would not expect this type of proximity to lead to elevated numbers of shared answers due to cheating. On the other hand, if there are other factors that lead to correlation in student answers who sit near each other (e.g. because they study together, or good students congregate near the front of the room), then the back-front indicator will capture these effects. In some specifications we also control for the gender composition of the pair, whether they are part of the same academic department.[7]

Table 1 shows the results of these regressions, using as the outcome variable the number of shared incorrect answers across the two students. The results in columns 1 and 2 correspond to the third midterm in which cheating is suspected. Columns 3 through 6 reflect the final exam. In columns 3 and 4, the right hand side variables associated with seat locations are the initial, voluntary seats occupied by the students; columns 5 and 6 are the assigned seats given to the students – where they actually sat when taking the final exam. The odd columns do not include any controls; the even columns include controls.

Students who sit next to one another on the midterm have an additional 1.1 shared incorrect

[6] Because we only have seating charts for the third midterm and the final, our analysis is restricted to these two tests.

[7] For gender we include dummies for both female, one female-one male, and two males. Each student is assigned to an academic department within the University (e.g. engineering or arts and sciences).

answers. This estimate is highly statistically significant. Note that, because the typical student gets most of the questions correct, the mean number of shared incorrect answers across all pairs of students is only 2.34. Thus, students who set next to each other have roughly twice as many shared incorrect answers as would be expected by chance. In contrast, we see no evidence that trying to sit next to one another in the final -- but being relocated before the test began – leads to shared incorrect answers. If, for instance, studying together were the cause of correlated answers among people who choose to sit next to each other, then we would expect more shared incorrect answers even after those students are relocated prior to the test. Students who actually sit next to each other in the final – after having been relocated and in the presence of heavy monitoring – also show no evidence of correlated incorrect answers in columns 5 and 6.

Pairs of students sitting two seats away with another student in between also have elevated levels of matching incorrect answers, on the midterm only. If there is an empty seat in between, however, then the correlation in incorrect answers disappears. This suggests that triplets of students may have worked together to cheat. There is no impact on shared answers among students sitting front to back on any of the tests.

Table 2 is identical to Table 1, except that the dependent variable is the number of shared correct answers. The pattern of coefficients is quite similar to the previous table. Students sitting next to each other during the midterm have an extra 1.2 shared correct answers. This coefficient is statistically significant at the .05 level. Students who are two seats apart with another student in between on the midterm once again have positive coefficients (but this time statistically insignificant). None of the other seating variables are coefficients are particularly predictive; if anything, sitting two seats apart with an empty seat in the middle reduces the number of shared correct answers.

Section III: A more systematic algorithmic approach to detecting cheating

The regressions above show that students who sit next to each other tend to have an increased number of both correct an incorrect answer pairs on average, but for identifying likely cheaters, it is the abnormality of the answers of a particular pair of students which is critical. In order to identify unlikely occurrences of matching answers, we first need to establish a baseline expectation with respect to the expected number of matching answers for any pair of students. To do so, we begin by modeling a student's answer on a particular question on either the third midterm or the final exam using a multinomial logit of the form:

$$p_{sa} = \Pr(y = a) = \begin{cases} \dfrac{e^{X_s \beta_a}}{1 + \sum_{a=1}^{3} e^{X_s \beta_a}} & if\ a < 4 \\ \dfrac{1}{1 + \sum_{a=1}^{3} e^{X_s \beta_a}} & if\ a = 4 \end{cases} \tag{1}$$

Where *s* indexes students and *a* reflects the answer the student gave to that question. There are four possible answers to each question, *a*=(1,2,3,4). In our basic specifications, we use the student's percentage correct on each of the midterms and the final to predict the answers a student gives to a particular question on the third midterm or the final. In computing the student's percentage correct, we exclude the results for that particular question.[8]

Let $\widehat{p_{ia}^{q}}$ denote the estimated probability that student *i* gives answer *a* on question *q*. Further, denote *a* as the correct answer, and *b*, *c*, and *d* as incorrect answers. If two students *i* and *j* answer a particular question independently then the probability that they both choose answer *a*, conditional on the variables included as controls in the multinomial logit, is given by $\widehat{\mathrm{p}_{\mathrm{ia}}^{\mathrm{q}}} \times \widehat{\mathrm{p}_{\mathrm{ja}}^{\mathrm{q}}}$. For each pair of students *ij*, the expected number of matching right and wrong answers is given by:

$$E(matching\ right\ answers) = \sum_{q} \widehat{\mathrm{p}_{\mathrm{ia}}^{\mathrm{q}}} \times \widehat{\mathrm{p}_{\mathrm{ja}}^{\mathrm{q}}} \tag{2}$$

$$E(matching\ wrong\ answers) = \sum_{q} \widehat{\mathrm{p}_{\mathrm{ib}}^{\mathrm{q}}} \times \widehat{\mathrm{p}_{\mathrm{jb}}^{\mathrm{q}}} + \widehat{\mathrm{p}_{\mathrm{ic}}^{\mathrm{q}}} \times \widehat{\mathrm{p}_{\mathrm{jc}}^{\mathrm{q}}} + \widehat{\mathrm{p}_{\mathrm{id}}^{\mathrm{q}}} \times \widehat{\mathrm{p}_{\mathrm{jd}}^{\mathrm{q}}} \tag{3}$$

[8] An argument could be made for using only the student's performance on the final exam as a control variable, due to cheating concerns on the midterms. Empirically, our results are little changed if we do that, or if we add more covariates such as a gender dummy.

We then compute two potential indicators of cheating based on unexpected concordance of answer patterns: (1) the residual between the observed and predicted number of matching *correct* answers (Δ_c), and (2) the residual between the observed and predicted matching *incorrect* answers (Δ_i). A priori, it is uncertain which of these two measures will be most predictive empirically.

We estimate the probabilities implied in model (1) using data for the 214 students that took both the third midterm as well as the final.[9] We then create a data set of all possible student pairs (22,791=214*213/2) for which we compute the number of matching correct and incorrect answers as well as the expected number of matching correct and incorrect answers using equations (2) and (3) respectively. Finally, we compute Δ_c and Δ_i as the difference between the observed and predicted number of matching correct and incorrect answers respectively.

Figure 1 shows a scatter plot of Δ_c against Δ_i on the third midterm, where cheating was suspected. Each symbol in the plot represents a pair of students in the class. Red triangles correspond to pairs of students who sat next to each other during the third midterm; blue squares are pairs of students who had one seat in between them, with that seat occupied; black circles are a 1% sample of all other pairs of students. The further to the right a pair is in the figure, the greater the number of unexpected shared incorrect answers. The higher up a pair is in the figure, the greater the number of excess correct answers. The red and blue symbols are greatly overrepresented in the Northeastern parts of the figure, consistent with cheating. Although pairs of students sitting next to each other represent only one half of one percent of all possible pairs of students, four of the six right-most data points are pairs of students who were next to each other. Although the pattern is less clear in the vertical dimension, the single greatest anomaly on shared correct answers is a pair seated next to one another. Some blue squares are

[9] We drop questions which all students answer correctly, as they provide no information. We also drop a handful of cases where exactly one student gave a particular answer on a question because of non-convergence of the multinomial logit estimation.

near the far right of the figure, but the pattern is much less obviously extreme than for the red.

For purposes of comparison, Figures 2 and 3 mirror Figure 1, but for the seating positions originally selected on the final (Figure 2) and the actual seating positions on the final (Figure 3). In stark contrast to Figure 1, there is no visual evidence that students wishing to sit next to each other or actually sitting next to one another have unusual answer patterns. This supports the interpretation of cheating on the midterm.

Figure 4 provides a more systematic means of capturing the degree to which sitting next to one another produces anomalous patterns on the midterm. The horizontal axis on Figure 4 captures ranges of excess incorrect answers (Δ_i), with the right-most columns corresponding to the highest (most suspicious) values. The height of the columns represents the hazard rate of the frequency with which students who sit next to each other produce outcomes in that range, compared to all pairs of students. Standard error bands are shown in red. If students sitting side-by-side look similar to randomly chosen pairs of students, then the hazard rates in all columns would be equal to one. If the hazard rate in a particular column is two, then that means that students sitting next to each other are twice as likely to generate answers in that range.

As can be seen in the figure, pairs of students who sit next to each other in the midterm produce answer patterns that diverge greatly from random pairs of students. (Note that the hazard rates are presented on a log scale because the differences in the tails are so large.) Focusing first on the right-most column of the table, which represents outcomes in the top one-tenth of one percent in terms of unexpected matching incorrect answers, pairs of students who sit next to each other are roughly 62 times as likely to fall into this category as a random pair. The null hypothesis that students sitting next to one another are no more likely than chance to be in this category is easily rejected. In the next column, reflecting outcomes in the 99.5-99.9th percentiles, students who sit next to each other are approximately 12 times more frequently represented than would be expected by chance. This result is also highly statistically significant.

For none of the other columns can we reject the null hypothesis of a hazard rate of one for students who sit next to each other. By chance, we would have expected one half of one percent of pairs of students sitting next to each other to appear in one of the two right most columns, or less than one pair. In actuality, about 9 percent of all left-right pairs (18 individual students) show up in the extreme tail.

Figures 5 and 6 present parallel results for the initial and actual seatings of the final exam. In contrast to Figure 4, there is no evidence that sitting next to another student is associated with large jumps in shared incorrect answers. We cannot reject the null hypothesis of a hazard rate of one for any of the columns on interest.

From this, we conclude that the overrepresentation of shared wrong answers by students seated next to each other on the midterm is plausibly attributed to cheating, implying that upwards of ten percent of the students cheated on the midterm in a manner that is detectable using statistics.

Figures 7-9 present the same set of results, but for matching correct answers rather than for matching incorrect answers. Four pairs—seven individual students because one student shows up in two pairs—appear in the right two columns. Of these, two of the pairs also would have been labeled cheaters based on anomalies in their incorrect answers. Thus, matching correct answers add relatively little to the potential cheating detection. Moreover, unlike for incorrect answers, there is an overrepresentation in the right most column on the final exam, both for those who wanted to sit together, but weren't allowed to, and for the students who were randomly assigned to sit next to each other. The former perhaps suggests that students who study together have correlated knowledge, although it is interesting that their incorrect answers do not overlap in an exceptional way.[10] The fact that randomly assigned students who

[10] Further calling into question this conjecture is that we surveyed students after the final exam the following year and asked them who else they studied with. We saw no excess weight in the right tail of either either shared correct or incorrect answers for those students who said they studied together.

sat next to each other also have correlated correct answers likely points to cheating in spite of the randomization. If the extra weight in the tails on the final is indeed due to cheating, then that suggests four students cheated on the final, still much lower than on the midterm.

Figures 10-15 are identical to Figures 6-9, except that they report patterns for students who sat one seat away from another student with a student sitting in the seat between them. On the midterm, relative risk ratios of roughly ten are present for the right-most category for both incorrect and correct answers. No such anomalies are present for such student pairs on the final exam, either in the seats they initially chose or in the randomized seats to which they were assigned. Doing calculations like those above, eight students sitting with a seat in between them are identified as likely cheaters. Of those eight, five were identified as cheaters in the analyses above that compared students sitting next to one another; three of them would have been missed.

Section IV: Conclusion

It is not surprising that students cheat – they have strong incentives to do so, and the likelihood of getting caught is low. What is perhaps more surprising, is that so little effort is devoted to catching cheating students. In this paper, we develop a simple algorithm for detecting cheating. In the particular setting in which we apply that algorithm, we conclude that more than 10 percent of the students in the class appeared to have cheated in a manner blatant enough to be detected by our approaches. For the most extreme examples, which leave the cheating students in the top one tenth of one percent of the distribution, the false positive rate (i.e. cases in which students are falsely accused of cheating) are likely to be quite small since students sitting next to one another are 62 times more likely than chance to find their way into this category.

Perhaps the best supporting evidence for our claims of cheating (and also, perhaps a powerful explanation as to why so little effort is invested in detecting cheaters), comes from what happened after we carried out our analysis. Based on our initial findings, the professor in the class forwarded the names of the six most suspicious pairs of students to the Dean's office, an investigation was initiated, and a student judiciary court hearing was scheduled.[11] Before the hearing could occur, four of the twelve students confessed. Despite these admissions, the Dean's office nonetheless cancelled the investigation the day before the student court hearing due to pressure from parents. While this precluded any further admissions of guilt, the professor withheld grades of the presumptive guilty pairs until the first day of the next semester which resulted in scholarship disqualification. Notwithstanding this punitive action, none of the twelve accused students complained or sought redress.

[11] Our initial detection algorithms were not as good as those we eventually developed; that is the reason only 12 students were identified.

Reference

Associated Press. “Dartmouth Students Charged with Cheating.” *The New York Times, US.* 8 January, 2015.

--- “Unusual Amount of Cheating Suspected at Stanford University.” *The New York Times, US.* 28 March 2015.

Baker, Al. “At Top School, Cheating Voids 70 Pupils’ Tests.” *The New York Times, NY/Region,* 9 July, 2012.

Frosch, Dan. “18 Air Force Cadets Exit Over Cheating.” *The New York Times, Education,* 2 May, 2007.

McCabe, Donald(2005). “Cheating Among College and University Students: A North American Perspective,” *International Journal for Academic Integrity*, Volume 1(**1**) (http://www.ojs.unisa.edu.au/ journals/ index.php/IJES/article/viewFile/14/64).

Jacob, Brian, and Steven Levitt (2003a). “Catching Cheating Teachers: The Results of an Unusual Experiment in Implementing Theory,” *Brooking-Wharton Papers on Urban Affairs*, pp185-209.

Jacob, Brian, and Steven Levitt (2003b). “Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics*, vol. 118(**3**), pp. 843-877.

Perez-Pena, Richard. “Students Accused of Cheating Return Awkwardly to a Changed Harvard.” *The New York Times*, *Education*, 16 Sept, 2013.

Perez-Pena, Richard and Jess Bidgood. “Harvard Says 125 Students May Have Cheated on a Final Exam.” *The New York Times*, *Education*, 30 Aug, 2012.

Perez-Pena, Richard. “Students Disciplined in Harvard Scandal.” *The New York Times,* *Education*, 1 Feburary 2013.

Zitzewitz, Eric (2012). “Forensic Economics,” *Journal of Economic Literature,* vol. 50(**3**), pp.731-69.

Table 1: Spring 2012 - Matching Incorrect Answers Amongst Pairs

Test	2012 Midterm 3		2012 Final Pre		2012 Final Post	
	(1)	(2)	(3)	(4)	(5)	(6)
Left-Right Pair	1.105***	1.104***	0.096	0.046	0.111	0.077
	(0.255)	(0.254)	(0.198)	(0.198)	(0.190)	(0.185)
Front-Back Pair	-0.062	-0.062	0.039	0.025	0.039	0.026
	(0.149)	(0.148)	(0.158)	(0.159)	(0.158)	(0.159)
Two Apart: Middle Student	0.614*	0.611*	-0.149	-0.195	-0.078	-0.117
	(0.272)	(0.271)	(0.193)	(0.190)	(0.214)	(0.215)
Two Apart: No Middle Student	0.147	0.119	-0.567	-0.538	-0.397	-0.359
	(0.229)	(0.222)	(0.347)	(0.352)	(0.363)	(0.370)
Constant	2.340***	2.316***	3.981***	3.962***	3.980***	3.961***
	(0.013)	(0.019)	(0.014)	(0.020)	(0.014)	(0.020)
Controls	No	Yes	No	Yes	No	Yes
N	19110	19110	22578	22578	22578	22578
r2	0.003	0.010	0.000	0.023	0.000	0.023

Table 2: Spring 2012 - Matching Correct Answers Amongst Pairs

Test	2012 Midterm 3		2012 Final Pre		2012 Final Post	
	(1)	(2)	(3)	(4)	(5)	(6)
Left-Right Pair	1.203*	1.204*	-0.124	-0.082	-0.388	-0.363
	(0.546)	(0.547)	(0.747)	(0.753)	(0.759)	(0.763)
Front-Back Pair	0.580	0.578	0.191	0.214	0.190	0.213
	(0.456)	(0.453)	(0.705)	(0.706)	(0.705)	(0.706)
Two Apart: Middle Student	0.762	0.770	0.486	0.568	0.129	0.190
	(0.540)	(0.545)	(0.883)	(0.890)	(0.863)	(0.865)
Two Apart: No Middle Student	-0.636	-0.550	-2.233	-2.211	-1.168	-1.104
	(0.834)	(0.852)	(1.152)	(1.163)	(1.483)	(1.510)
Constant	31.176***	31.154***	50.750***	50.638***	50.751***	50.639***
	(0.035)	(0.052)	(0.057)	(0.084)	(0.057)	(0.084)
Controls	No	Yes	No	Yes	No	Yes
N	19110	19110	22578	22578	22578	22578
r2	0.001	0.009	0.000	0.003	0.000	0.003

Figure 1: 2012 Midterm 3 - Residual Matching Answers for Student Pairs

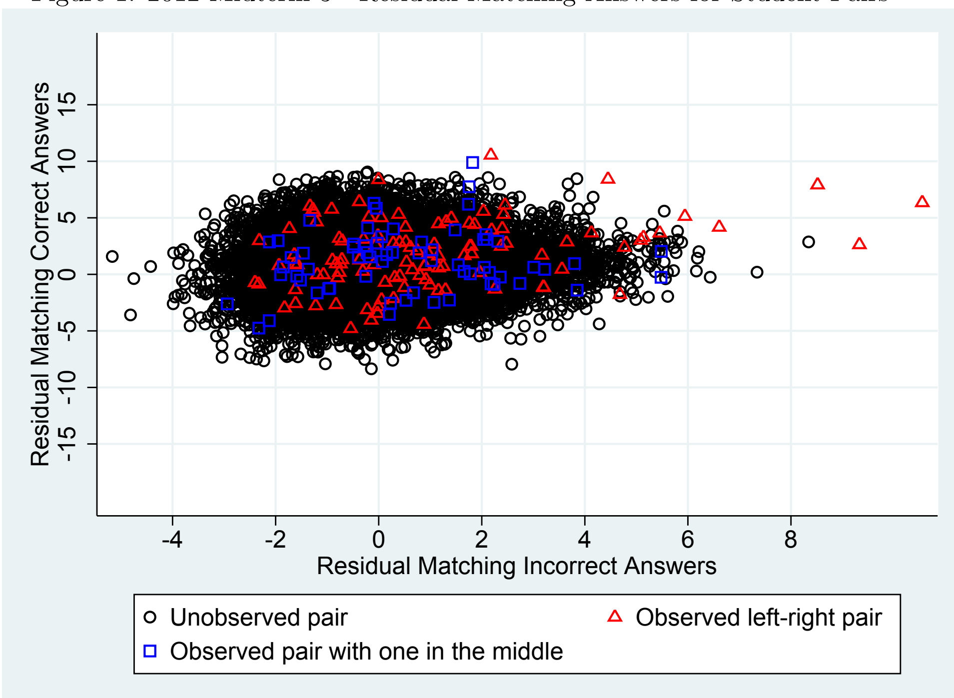

Figure 2: 2012 Final Pre-Rand - Residual Matching Answers for Student Pairs

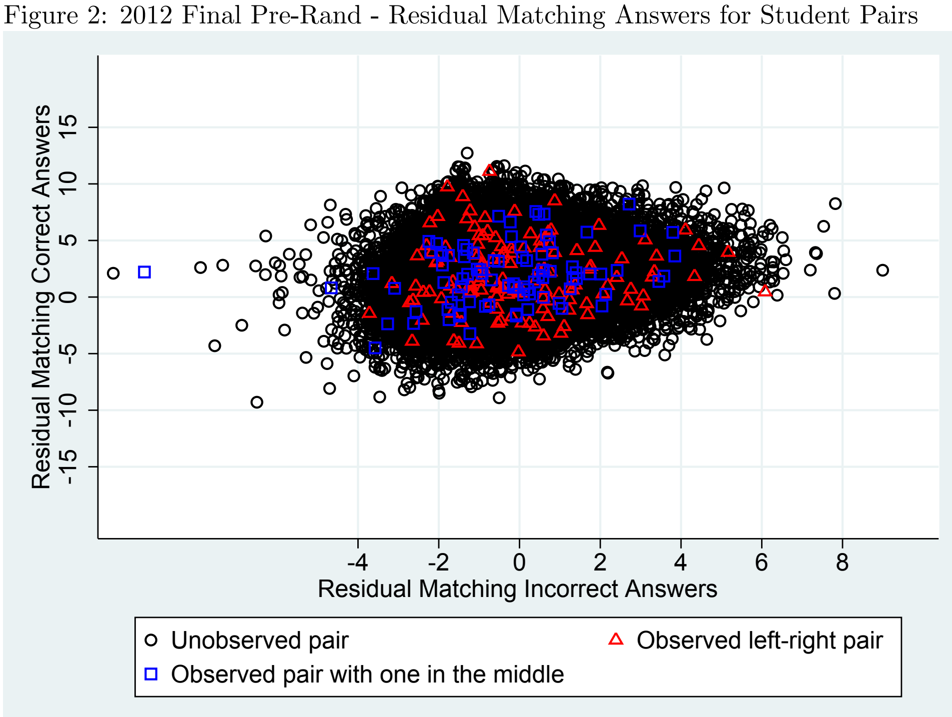

Figure 3: 2012 Final Post-Rand - Residual Matching Answers for Student Pairs

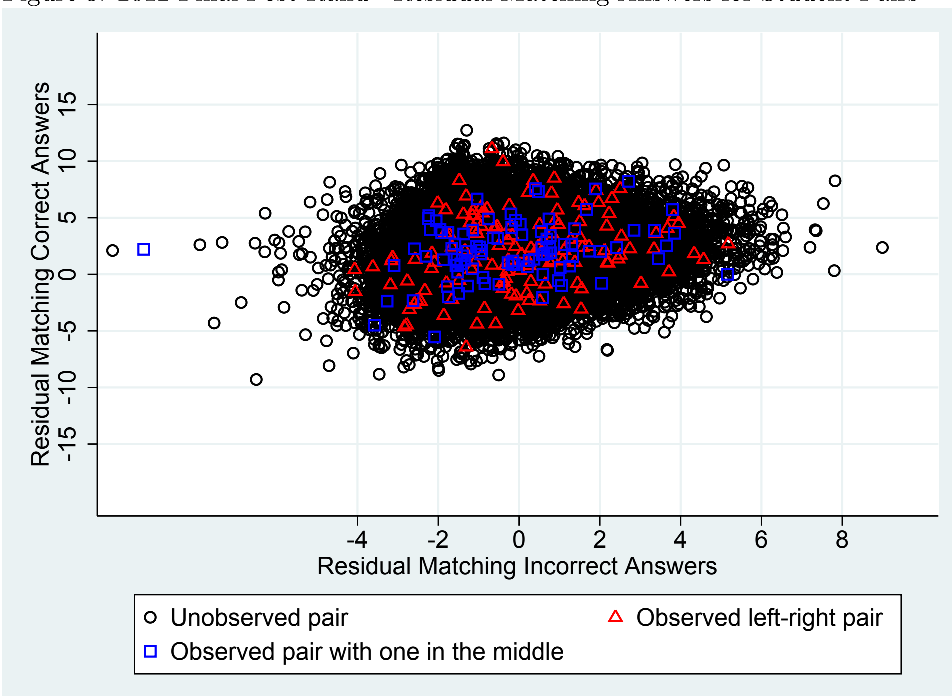

Figure 4: Left-Right Pair - 3rd Midterm

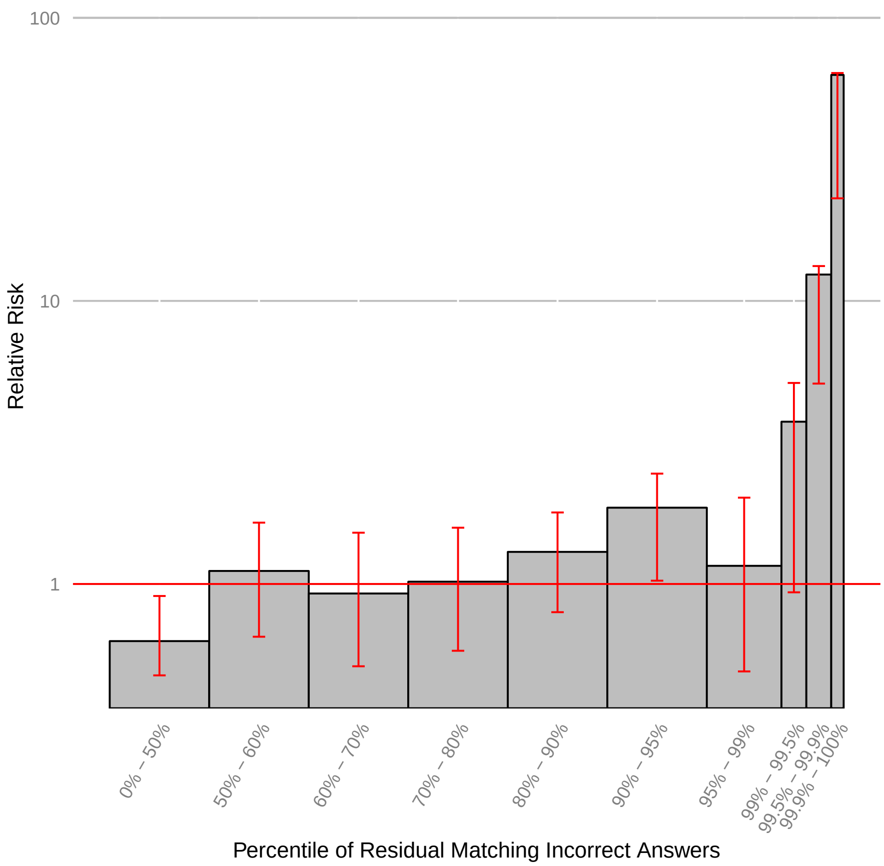

Figure 5: Left-Right Pair - 2012 Final Exam Pre-Randomization

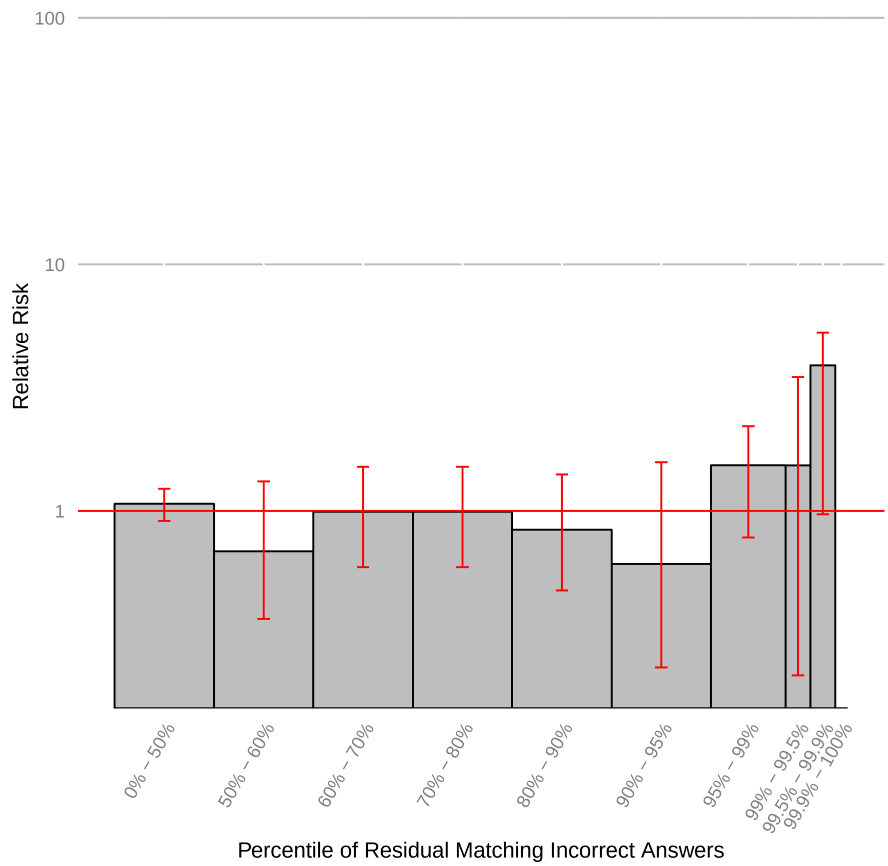

Figure 6: Left-Right Pair - 2012 Final Exam Post-Randomization

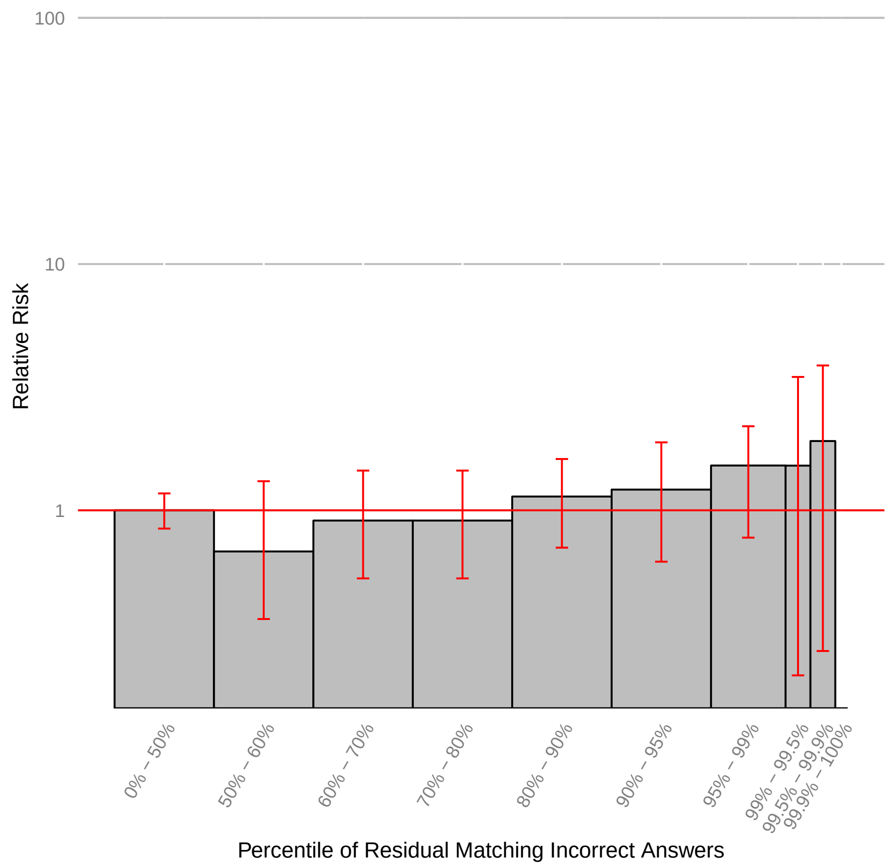

Figure 7: Left-Right Pairs - 3rd Midterm

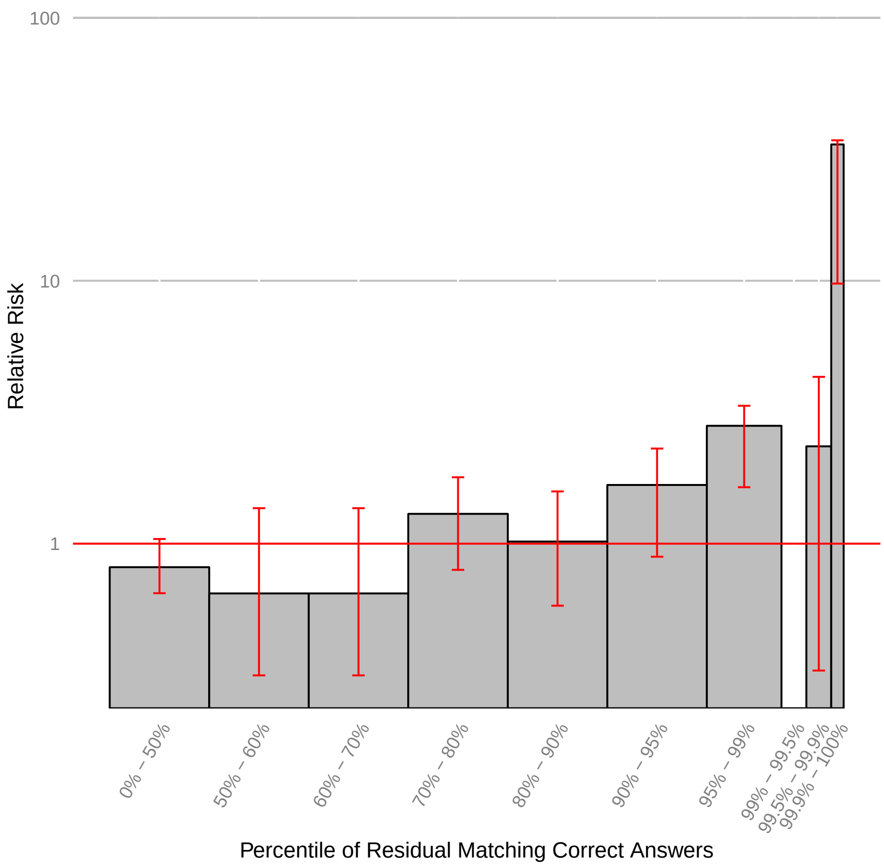

Figure 8: Left-Right Pair - 2012 Final Exam Pre-Randomization

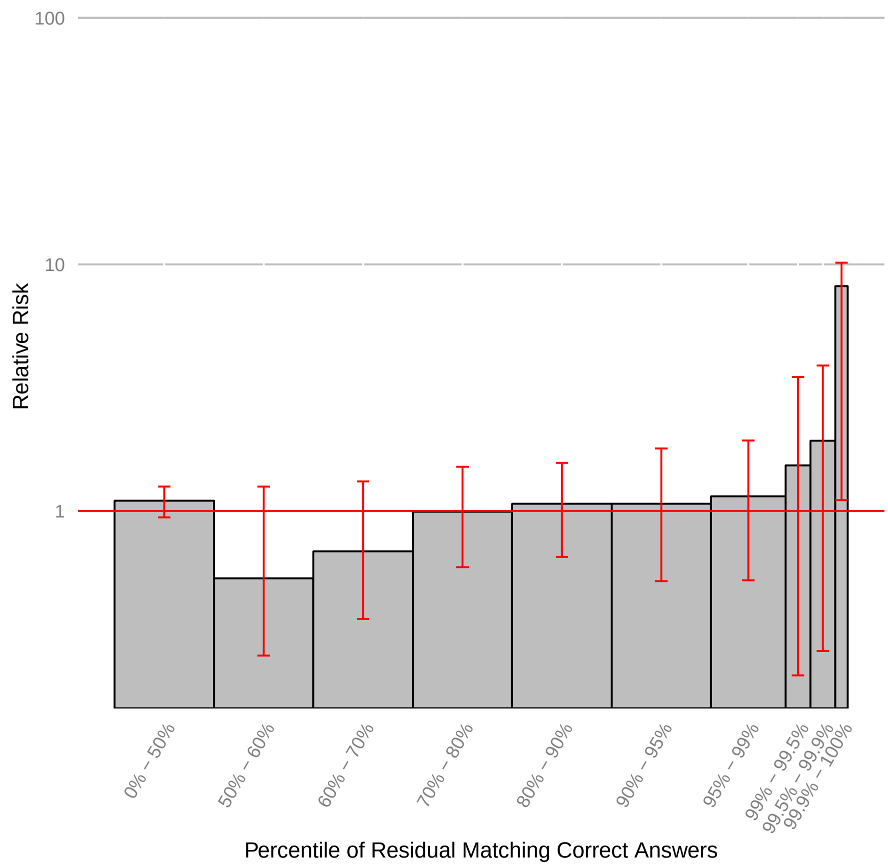

Figure 9: Left-Right Pair - 2012 Final Exam Post-Randomization

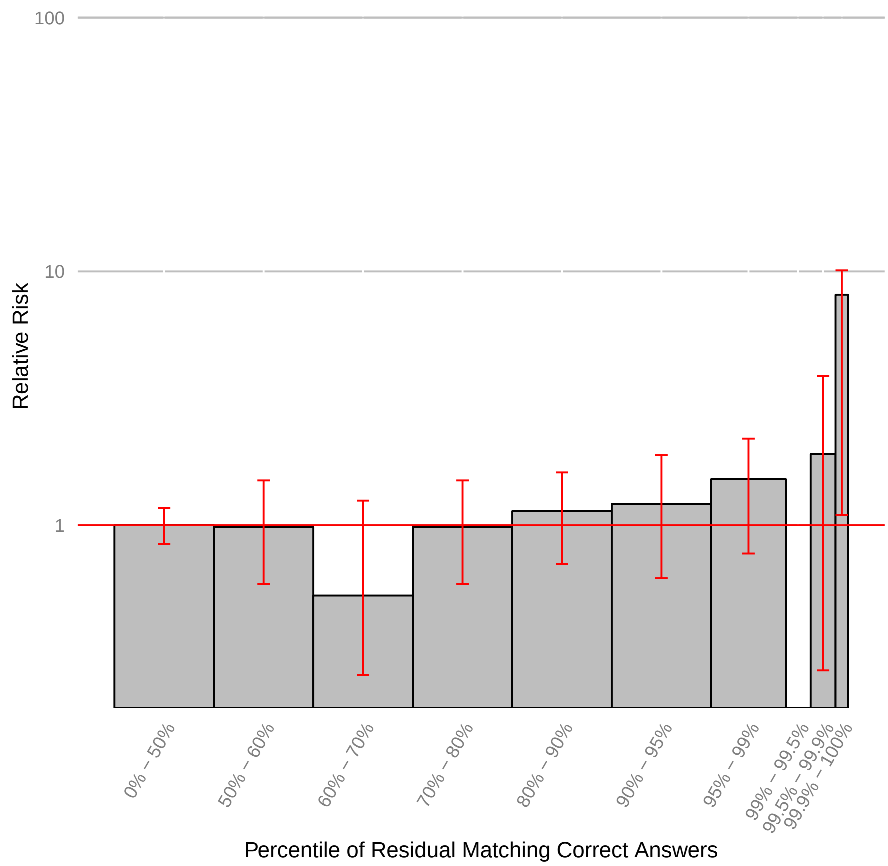

Figure 10: Pair with a student inbetween - 3rd Midterm

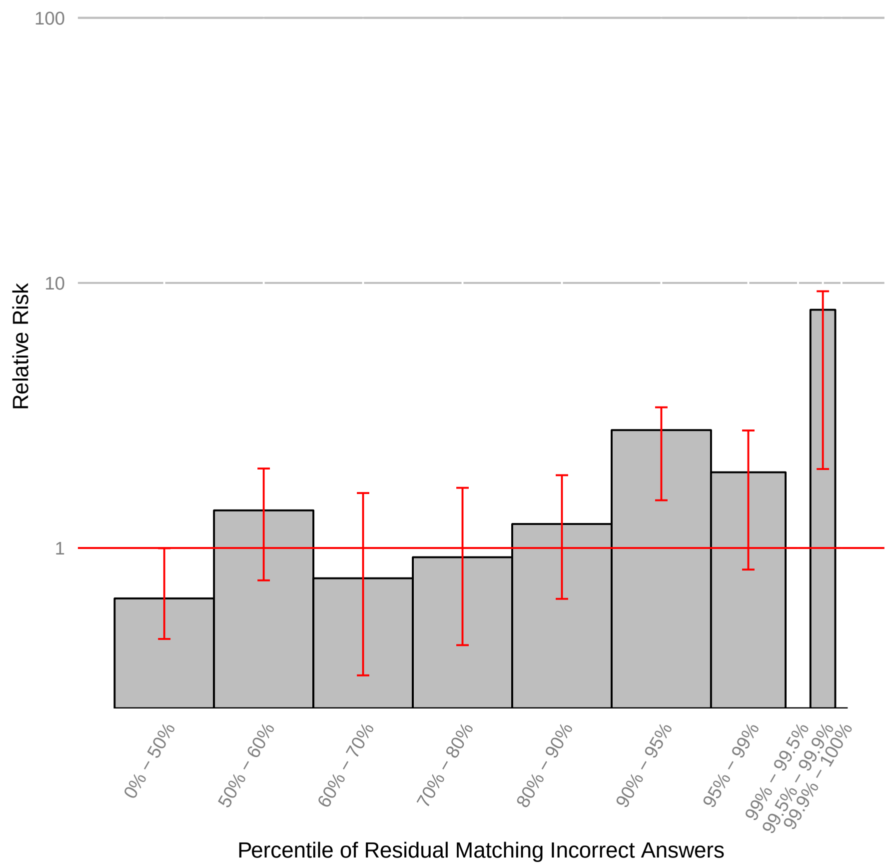

Figure 11: Pair with a student inbetween - 2012 Final Exam Pre-Randomization

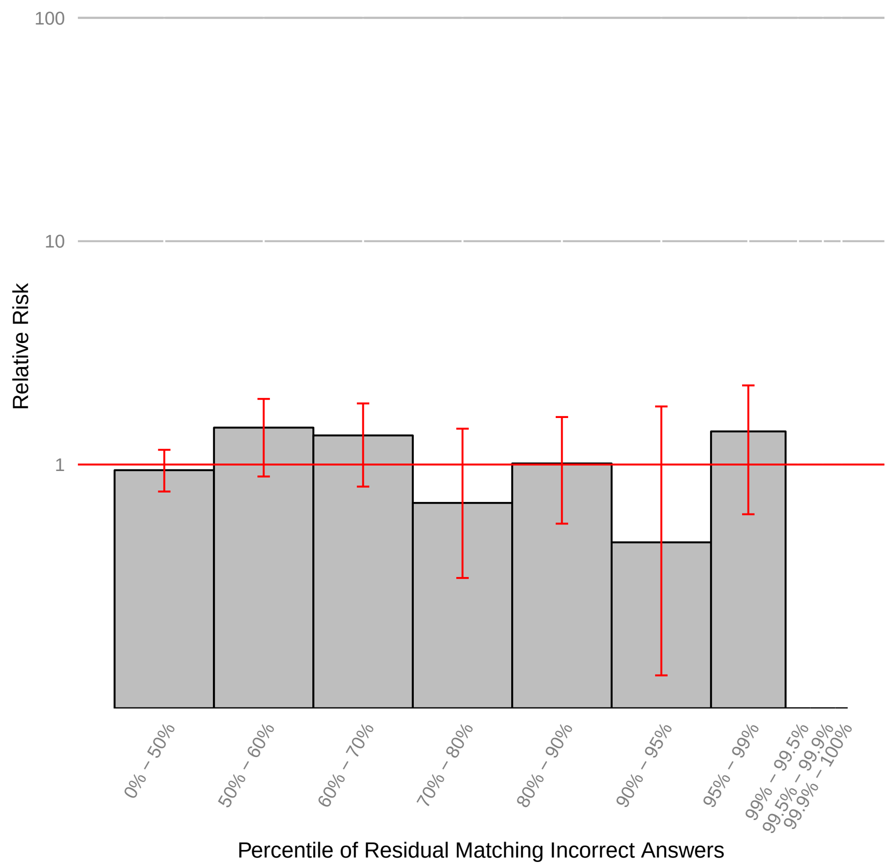

Figure 12: Pair with a student inbetween - 2012 Final Exam Post-Randomization

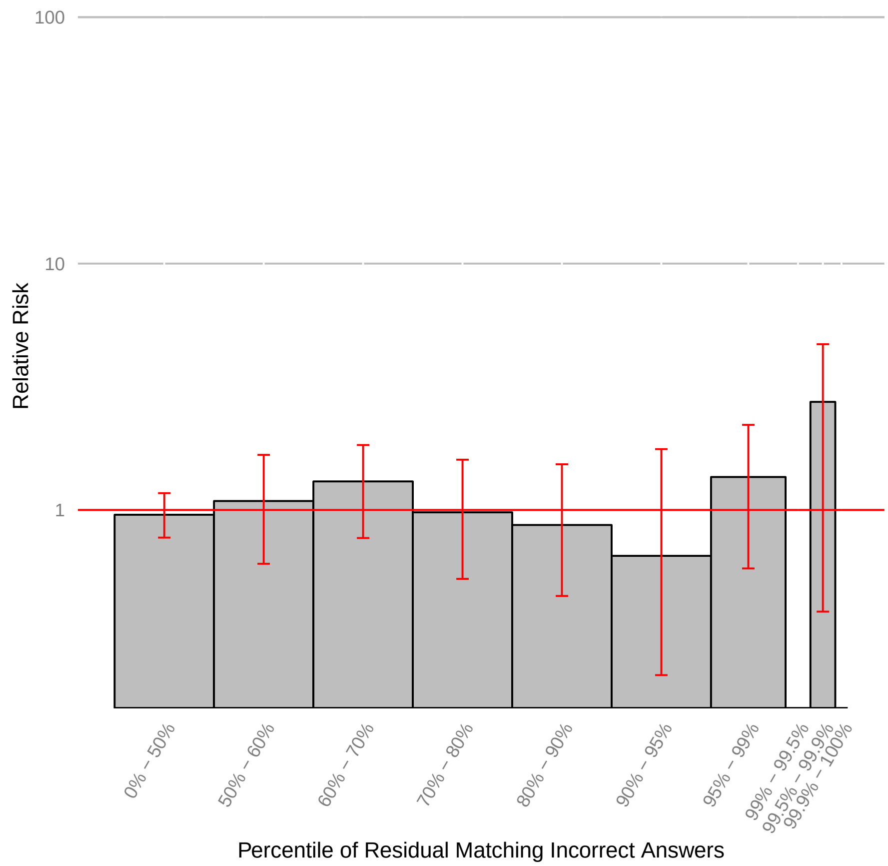

Figure 13: Pair with a student inbetween - 3rd Midterm

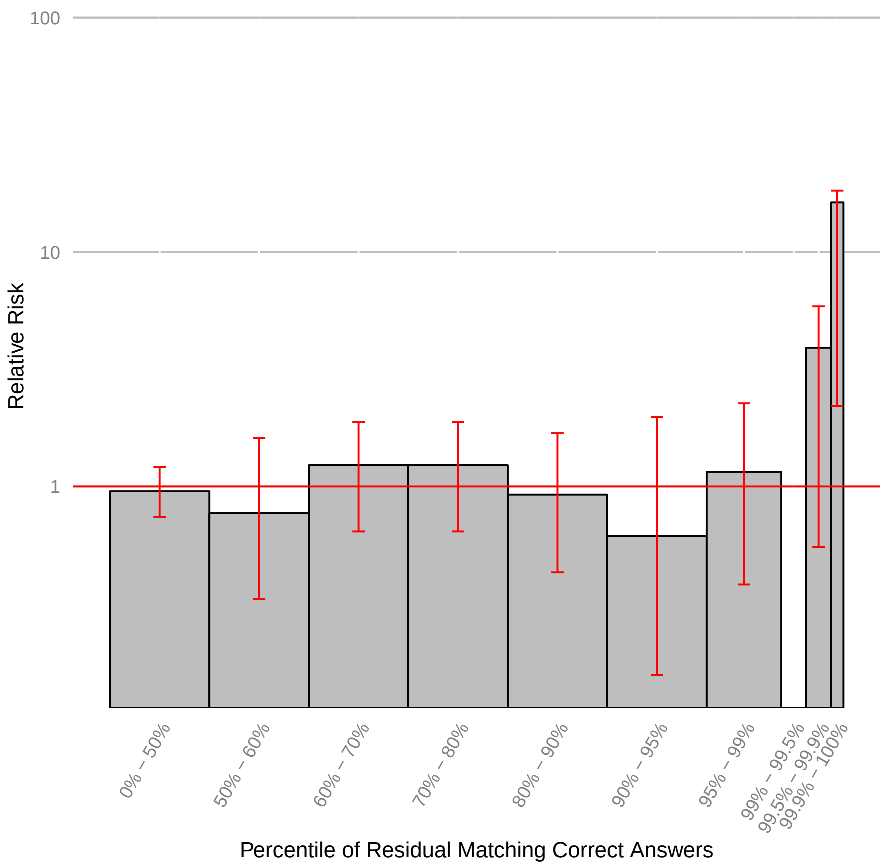

Figure 14: Pair with a student inbetween - 2012 Final Exam Pre-Randomization

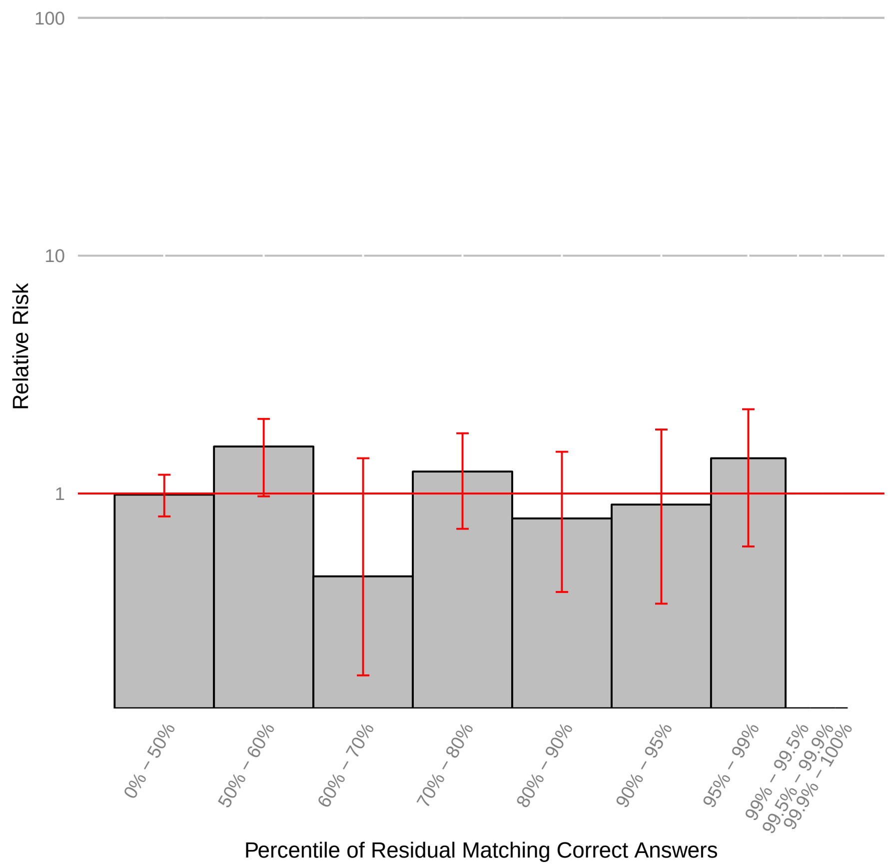

Figure 15: Pair with a student inbetween - 2012 Final Exam Post-Randomization