Additional Commentary on Birth Order and Attempted Base Stealing among Major League Brothers in Baseball

Frank J. Sulloway
University of California, Berkeley, CA, USA

Richard L. Zweigenhaft
Guilford College, Greensboro, NC, USA

November 2010

It has come to our attention that various commentators on a website that is devoted to statistical discussions of baseball performance have had difficulty understanding, and replicating, one of the research findings previously reported in our study of birth order and risk taking in athletics (Sulloway and Zweigenhaft, 2010).[1] For this reason we are presenting here an explanation of the methods that we used, as well as the relevant statistical output from PASW Version 17 (SPSS, Inc.) that was reported in Table 2 of our article. We are pleased that our study has generated such interest, and we are glad to do our best to correct any misunderstandings

A successful replication of statistical research results requires analysis of the same sample, the same dependent and independent variables, and use of the same statistical procedures. Some of those who have attempted replications of our findings, and who have criticized our methods, have not followed these requirements. More particularly,

1. See http://sabermetricresearch.blogspot.com/2010/05/do-younger-brothers-steal-more-bases.html; http://sabermetricresearch.blogspot.com/2010/06/do-younger-brothers-steal-more-bases.html; http://sabermetricresearch.blogspot.com/2010/06/do-younger-brothers-steal-more-bases.html; and http://sabermetricresearch.blogspot.com/2010/07/do-younger-brothers-steal-more-bases.html. We thank Lynn Gale for helpful statistical advice in connection with this commentary.

.

these various failures to replicate our statistical results may be attributed to one or more of the following six problems: (1) a failure to read our published article, as opposed to a *New York Times* (Science Times) news story about our findings (Schwarz, 2010); (2) consequently, a failure to distinguish between results erroneously interpreted in this *New York Times* article and our actual methods, findings, and conclusions; (3) failure to analyze the same dependent variable (*attempted* steals per opportunity rather than successfully completed steals); (4) failure to conduct analyses on the same sample; (5) failure to control call-up sequence (which introduces a substantial confound in the results); and finally (6) a lack of familiarity with Mantel-Haenszel common odds ratios, and hence a failure to calculate such statistics and to understand their meaning.

Inasmuch as our methods and our sample are more fully described in our published article, we report these details here only briefly. Our sample included all brothers who have played in the major leagues, excluding those for whom date of birth was unknown and players still active in 2008. We analyzed two groups of players: 472 position players and 228 pitchers, for a total of 700 brothers. Based on date of birth we coded older brothers as -1 and younger brothers as +1 for the "birth order" independent variable. In a small number of instances where there were three or more brothers (less than 3 percent of the sample), we coded middle brothers with the same contrast value as younger brothers (+1), whose performance statistics they most closely resembled.

Our first dependent variable—stolen base attempts—was normalized for opportunities to steal by dividing the number of attempted steals by the number of times each player reached a base, whether this was by hitting a single, double, or triple, by being walked, or by gaining first base after being hit by a pitch. We refer to this variable as "attempted steals per opportunity." To reduce skew and kurtosis to acceptable levels, we computed $\log_{10}$ values for this normalized measure of attempted steals per opportunity.[2]

For the second dependent or "sibling comparison" variable, we coded a categorical variable based on which brother attempted to steal more bases per opportunity. To each position player we assigned a ranking based on whether the player attempted to steal more bases per opportunity (+1), the same number (0), or fewer bases (-1) compared with his brother. In the small number of cases in which three or more brothers were position players in the major leagues and had differing rates of attempted steals per opportunity, we assigned intermediate rankings (for example, +1, 0, and -1, and, in the case of the five-brother Delahanty family, +1, +0.5, 0, -0.5, and -1).

The Call-up Sequence Confound

As we argue in Sulloway and Zweigenhaft (2010), performance data in major league baseball appear to be strongly biased by call-up sequence. A brother called up first was generally superior in athletic talent to a brother called up later, having a longer playing career (controlling age at first call up) as well as exhibiting superior performance in most of our measures of athletic achievement. Another way of expressing this point is that scouts, coaches, managers, and owners seem to have given preferential treatment to brothers of already successful baseball players, but these brothers generally exhibited a regression toward the mean in overall ability. This particular bias, however, does not express itself equally by birth order, as older brothers—owing to their greater age—were

2. The formula used by us was $\log_{10}$ (1+ [attempted steals per opportunity x 100]). The reason for adding 1 within this transformed measure is that logs are undefined for values of zero.

6.6 times more likely than younger brothers to be discovered first (the relative risk ratio) and hence to receive an earlier call to the major leagues (r=.70, n=682, p<.0001). As a consequence, we are often comparing somewhat more talented older brothers with somewhat less talented younger brothers, who generally appear to have benefited from a "halo effect" in getting a call up to the major leagues. A similar, reverse bias is evident when younger brothers received a call first to the major leagues, as they were generally superior players compared with their own older brothers.

To minimize this call-up sequence bias, we controlled most of the relevant statistical results in our article for this temporal sequence. It is worth emphasizing that if call-up sequence creates no confound in overall performance, its inclusion as a control variable will have no tangible effect on the results of our analyses. It is also worth noting that the actual implementation of this control probably underestimates the extent of any call-up sequence bias, as some older brothers of talented younger brothers were unlikely to be called up owing to the fact that their greater age made them less attractive as career prospects. Hence some superior younger brothers who played major league baseball and whose older brothers played in the minor leagues are almost certainly missing from our sample because the older brothers never received calls to the major leagues. As evidence of this biased trend, although a total of 263 older brothers were called up first to the major leagues, only 41 younger brothers were called up first, suggesting that several hundred pairs of brothers, in which the younger was actually more talented and made it to the major leagues, are absent from our sample. Thus the strong call-up bias that affects our dataset is asymmetrical rather than being evenly balanced by birth order, a circumstance that attenuates, in particular, differences in base-stealing by birth order.

Some of our sabermetrician commentators seem to believe that when older brothers are called up first and younger brothers are called up second that no bias is introduced by call-up sequence. Data on batting ability and career length contradict this supposition. In our within-family sample, older brothers in the called-up-first subgroup played in 88 percent more games than did their younger brothers (n=158, p=.001), hit 47 percent more home runs per at bat (n=158, p=.02), attained 45 percent more RBIs (n=158, p=.01), and had higher career batting averages (.258 versus .225, n=158, p=.02). Older brothers in this same subgroup were 3.8 times more likely to have higher scores on our 8-variable composite measure of batting ability (the odds ratio, corresponding to a correlation of -.25, n=158, p=.001). In addition, older brothers were more likely than their younger brothers to be called up at a younger age (r=.18, n=158, p=.02), and call-up age is in turn a good predictor of attempted steal rates (r= -.18, n=158, p=.02). To confound matters further, superior batters in this subsample attempted more steals per opportunity than did inferior batters (r=.26, n=158, p=.001), and they were also more successful at doing so (r=.38, n=158, p<.001).

It is important to note that a bias by call-up sequence affects the sample as a whole, not just those instances in which older brothers were called up before younger brothers. Batting ability provides a case in point, as younger brothers called up first exhibited superior performance compared with their own older brothers (based on our eight-variable composite measure of this attribute; r=.41, n=44, p<.01). Moreover, the magnitude of the performance edge that these younger brothers had over their older brothers is indistinguishable, statistically, from the magnitude of the performance edge that older brothers had over younger brothers whenever older brothers were called up first (Z=1.40, n=339, p=.16).[3] In short, with the exception of brothers called up during the same year, the bias by call-up sequence operates uniformly on everyone in the sample.

Because we are interested in generalizing about the role of birth order in risk taking, we need to adjust for the various confounding factors that are inherent in our sample of major league baseball players. We can adjust for these confounding factors by controlling playing age and ability or, as we previously did, by controlling call-up sequence (which is an effective proxy for differences in playing age and ability).

Models Controlling Call-up Sequence

Table 1 shows the results of a regression analysis with $\log_{10}$-attempted-steals-per-opportunity as the dependent variable, birth order as the independent variable, and call-up sequence as a control variable. The partial correlation between birth order and attempted steal rates for our within-family data—that is, cases in which all brothers were position players—is .17 (n=185, p=.02). This is the same partial correlation reported in Table 2 of Sulloway and Zweigenhaft (2010). In addition to being controlled for call-up sequence, these within-family results are controlled for social class and sibship size, as all comparisons involve brothers who grew up within the same families. For our sibling comparison variable, the partial correlation between birth order and attempted steals per opportunity, controlling call-up sequence, is .26 (n=185, p<.001). Standard regression

---

3. This test compares the magnitude of the performance edge by birth order without regard to its direction. Because older brothers, when called up first, were superior batters, and because younger brothers were also superior batters whenever called up first, the interaction effect between birth order and call-up sequence is significant when the directionality of the effect is taken into account (t=4.12, df=1/335, pr=.22, p<.001). Neither of the two main effects (birth order or call-up sequence) is a significant predictor in this model. Athletic superiority by birth order is therefore dependent on call-up sequence.

diagnostics revealed no problems with multicollinearity in these and other analyses, and no cases with excessive influence.

In Table 2 we present the results of a Mantel-Haenszel common odds ratio test for our major league data. It may be seen that the common odds ratio is 10.58, as previously reported in Sulloway and Zweigenhaft (2010, Table 2). Within each call-up sequence subgroup, the odds ratio indicates that younger siblings attempted more stolen bases per opportunity, showing consistency across the three subgroups. For brothers called up to the major leagues first, a younger brother was 6.56 times more likely than an older brother to have a higher rate of attempted steals. For brothers called up during the same year (and in 5 instances where a player was called up in between two or more brothers), the odds ratio is 7.00 to 1 in favor of younger brothers attempting more steals per opportunity. Finally, for brothers called up last, the odds ratio is undefined (but obviously large), as no older brother out of 5 attempted more steals per opportunity than his own younger brother, whereas 42 of the 77 younger brothers (55%) exhibited a higher rate of base stealing. When these strata are combined, the Mantel-Haenszel test shows the three subgroups having a common odds ratio of 10.58 to 1 (chi-square=9.87, df=1, n=177, p=.002).[4] The Breslow-Day test for the homogeneity of the odds ratios is not significant (chi-square=0.87, df=2, p=.65), indicating the propriety of pooling odds ratios from the three different strata.[5]

4. The size of this sample is 177 rather than 185, as the analysis excludes 3 ties and 5 cases in sibships of three brothers where one brother's comparison value for attempted steals per opportunity was intermediate (that is, 0 rather than +1 or -1).

5. The Breslow-Day test requires a relatively large sample size within each stratum to be valid. We have therefore conducted a second test in which we have pooled the data for 15 brothers who were called up during the same year (n=10), or in between an older and a

In contrast, pooling the data by eliminating the control for call-up sequence results in an odds ratio of just 2.13 to 1, which is nevertheless statistically significant even without the call-up sequence control (chi-square=6.18, *phi*=.19, *df*=1, *n*=177, *p*=.01). The substantial reduction in effect size, from an odds ratio of 10.58 to 1 to 2.13 to 1, is primarily the result of a phenomenon known as Simpson's paradox, whereby significant relationships present within subgroups are sometimes substantially reduced or lost entirely when these subgroups are indiscriminately pooled together (Simpson, 1951). The lower effect size is also a consequence of failing to control the dataset for the asymmetrical bias associated with call-up sequence.

Heterogeneous odds ratios. Additional evidence, which is related to Simpson's paradox, supports the need for controlling performance data by call-up sequence. If the data for attempted steal rates are stratified by matched sibling pairs—that is, if the data are grouped into (1) older brothers called up first, together with their own younger brothers called up later, (2) brothers called up the same year (or in between an older and a younger brother), and (3) younger brothers called up first, together with their own older brother called up later—the Mantel-Haenszel common odds ratio becomes 2.09 to 1 in favor of younger brothers attempting more steals per opportunity (chi-square=5.41, *df*=1, *n*=177, *p*=.02) This common odds ratio is nearly identical to the odds ratio with no stratification (2.13 to 1), as it should be, since stratifying in this manner is much the same as not stratifying at all. With this form of stratification, however, the Breslow-Day test for the homogeneity of the odds ratio becomes significant (chi-square=6.72, *df*=2, *p*=.03),

---

younger brother (*n*=5), with the data for brothers called up later (*n*=82), whom they closely resemble in their base-stealing proclivities. The Breslow-Day statistic for this test yields a chi-square of 0.01 (*df*=1, *p*=.94), confirming findings using three categories for call-up sequence.

indicating that the three odds ratios are not homogeneous by subgroup, and hence that one should not compute a common odds ratio from such data.[6] Only if one stratifies by call-up sequence, which controls the associated bias and renders the odds ratios homogeneous, is the Breslow-Day test rendered nonsignificant, allowing a common odds ratio to be estimated for these otherwise heterogeneous data.

A Clarification about Odds Ratios

Odds ratios are a useful way of expressing effect sizes and are generally more meaningful to lay readers than are correlations, *t*-values, and *F*-values. For this reason, odds ratios are widely used in newspaper articles reporting therapeutic results in medicine. A good example is provided by the Salk vaccine trials in 1954-55, which yielded a correlation (*phi*) of only .01 between treatment with the vaccine and subsequent rate of affliction by polio, thus explaining only 1/100th of 1 percent of the variance (as denoted by the *R*-squared of .0001). The odds ratio for these same data, however, tells a very different story, and it was this medical story that made the 12 April 1955 announcement of the Salk vaccine trials front-page news around the world. Among study participants who did not receive a vaccination for polio, the odds of contracting this disease were 3.5 times greater than for vaccinated participants (Rosnow and Rosenthal, 2003).

---

6. As a check on the validity of the Breslow-Day test, we have combined the two strata with the smallest sample sizes—namely, brothers called up during the same year (or in between two or more brothers) and cases involving younger brothers called up first, older brothers called up later. (Such pooling is appropriate because these two subgroups have similar birth order trends for attempted steal rates.) The Breslow-Day result for this dichotomous stratification yields a chi-square of 4.58, *df*=1, *n*=177, *p*=.03, indicating significant heterogeneity of the odds ratios and confirming the results of our previous 2-*df* test.

Ironically, although odds ratios are often used in an attempt to clarify complex statistical findings, people who are not familiar with them sometimes misinterpret what odds ratios do and do not mean. In our own data for major league brothers, for example, an odds ratio of 10.58 to 1 in favor of younger brothers attempting to steal more bases per opportunity does not mean that younger brothers attempted 10.58 times the number of steals as did their older brothers. Similarly, this statistic also does not mean, as Schwarz (2010) mistakenly reported in the *New York Times*, that more than 90 percent of younger brothers attempted more steals per opportunity than their own older brothers. Only 59 percent of younger brothers in our sample attempted more steals per opportunity, although this statistic, uncontrolled for call-up sequence, considerably underestimates the overall effect for this measure, just as computing an odds ratio without regard to call-up sequence underestimates the effect. For example, among the 10 brothers in our study who were called up during the same year—where there is no possible bias owing to call-up sequence—80 percent of younger brothers (4 out of 5) attempted more stolen bases per opportunity than their older brothers, yielding a relative risk ratio 4.00 to 1 (80%/20%), and an odds ratio of 16.00 to 1.

Other Potential Confounds: Playing Age and Ability

Mid-career playing age. Some of those who have read our article, or read about our article, have suggested that age differences among these players could be a confounding factor in our results. To test this assertion we previously created a variable for mid-career age (that is, age at first call up + ½ of total number of years played). This variable controls for brothers whose mid-career playing age might have been substantially different than a sibling, which could conceivably affect their speed on the bases. A younger brother called up several years before an older brother, for example, would generally have played at an average age that was younger than his brother, and would perhaps have been faster on the bases. Controlling attempted steals per opportunity for mid-career age as well as call-up sequence produces a partial correlation between birth order and $\log_{10}$-attempted-steals-per-opportunity of .171 instead of .170 (n=185, p=.02). Similarly, after controlling these same two variables, the partial correlation between birth order and our sibling comparison variable for attempted steal rates becomes .257 instead of .256 (n=185, p=.001). In short, playing age differences between the brothers do not appear to be responsible for the significant sibling disparities we have documented in attempted steals. In addition, goodness-of-fit analyses indicate that models that omit mid-career age as a control are superior to models that include this variable, based on Akaike's Information Criterion (AIC).

The reason why differences in mid-career playing age have so little effect on our results is because these differences are modest. Examining the mid-career playing ages of the 185 brothers for whom we have data on attempted steals, the average difference in age by birth order is just 6.0 months, with younger brothers being nonsignificantly younger. Such a small disparity in mid-career playing age is too little to have much effect on either base-stealing proclivities or on other baseball abilities, especially given that the correlation between mid-career age and attempted steals per opportunity is close to zero (r=.02, n=185, p=.74).

Age at first call-up. In addition to controlling mid-career playing age, one might also control age at first call up. This particular control would serve to distinguish between brothers who had similar mid-career playing ages, but who began playing at different times in their careers. Unlike mid-career age, age at first call-up is significantly associated with attempted steals per opportunity (*r*= -.22, *n*=185, *p*=.002). Nevertheless, when age at first call-up is added to call-up status in a multiple regression model, the partial correlation between birth order and our sibling contrast variable for attempted steal rates is little changed (*pr*=.24, *n*=185, *p*=.001; versus *pr*=.26, without the control for age at first call-up). In spite of being significantly related to attempted steal rates, age at first call-up ends up making little difference in the regression model—and is not itself a significant predictor—because it cross-correlates with call-up sequence, which is already included in the model and is a good proxy for it. Once again, goodness-of-fit analyses indicate that models omitting age at first call-up are superior to those that include this covariate, based on AIC values.

Overall athletic ability. Similarly, differences in overall athletic ability as they relate to call-up sequence also do not appear to be responsible for observed differences in base-stealing proclivities. Controlling our eight-variable measure of batting ability as well as call-up sequence, the partial correlation between birth order and $\log_{10}$-attempted-steals-per-opportunity becomes .15 (*n*=185, *p*=.04), as opposed to .17 without this additional control. Controlling these same two covariates in a model with sibling comparisons as the dependent variable, the partial correlation of birth order with attempted steal rates becomes .24 (*n*=185, *p*<.001), as opposed to .26 without including a control for batting ability. A model that includes batting ability is marginally superior to a model omitting this covariate, although most goodness-of-fit statistics favor omitting this predictor.[7] In

7. Using Akaike's Information Criterion, a model that includes birth order, call-up sequence, and batting ability is superior to a two-predictor model that includes just birth order and call-up sequence. Using Consistent AIC and Bayesian Information Criterion,

short, including call-up sequence as a control variable in our models does not create a spurious correlation between birth order and attempted steal rates, either by introducing differences in overall playing ability or by introducing differences in mid-career playing age or age at first call-up. Instead, call-up sequence provides an effective and generally parsimonious control for these kinds of confounding influences.

Estimating Odds Ratios Using Logit Estimators

The 95-percent confidence interval for the odds ratio that we published in our article (10.58 to 1) extends from 2.21 to 1 to 50.73 to 1, so anything within this range may be considered a possible statistical value given expected levels of statistical error. The wide range of these confidence intervals reflects sparseness in some of the cells within the multiway tables. Logit estimators are sometimes used to estimate common odds ratios with data having sparse cells. If a cell in a stratum is zero, logit estimators add 0.5 to each cell within the stratum and also omit any tables having a zero row or column. Logit estimators are not available in SPSS (PASW, Version 17.0), which we used to conduct our published analyses. However, analysis of our data using PROC FREQ in SAS allows us to provide an alternative, and more conservative, measure of the common odds ratio using logit estimators. Based on a model that includes birth order stratified by call-up sequence, we obtain a common odds ratio of 7.88 to 1, with a more limited 95-percent confidence interval of 1.94-31.91 (Cochran-Mantel-Haenszel chi-square=11.62, df=1, n=177, p=.0007). A common odds ratio of 7.88 to 1 is probably a more reasonable estimate of the relationship between birth order and attempted steals given the difficulty of making an accurate estimate from some of the sparse cells in multiway tables.

---

which penalize models more heavily for adding predictors, the two-predictor model without batting ability is superior.

It should be mentioned that all *p*-values in our article were based on correlations and multiple regression analyses, not on Mantel-Haenszel common odds ratio tests. We included common odds ratios in our article as useful expressions of effect sizes, to supplement the statistical information provided by these other measures.

Attempted Steal Rates versus Completed Steals

Some sabermetrician commentators have failed to find a difference in successful base stealing by birth order and have erroneously concluded that this result refutes our own findings. In fact, we also analyzed completed steals per opportunity in our article, although we initially entertained no strong prediction as to which brother would be more successful at base stealing. Still, one might suspect that a higher rate of attempted steals per opportunity might lead, over time, to somewhat greater skill at base stealing. In our article we showed that younger siblings were not more likely to steal more bases per opportunity, but they did exhibit a higher percentage of successfully completed steals in two smaller subsamples for whom these data are available (Sulloway and Zweigenhaft, 2010: Table 2, note d).[8]

Conclusions

Birth order appears to be significantly related to attempted steals per opportunity among brothers who played major league baseball, with younger brothers exhibiting a higher rate than older brothers. This trend is statistically significant whether or not one controls for confounding variables, such as age at first call up, batting ability, career length, and call-up sequence. Inasmuch as brothers called up first tended to be superior

---

8. Computing success rates for stolen bases requires having data on attempted as well as completed steals. Data on attempted steals are not generally available for players who were active before 1920.

athletes compared with brothers called up later, and because older brothers were generally called up first, the trend for younger brothers to attempt more stolen bases per opportunity becomes even stronger when one controls these confounds. When other potentially relevant control variables are added to our models, the statistical findings prove to be almost identical to those using only call-up sequence as a control, indicating that call-up sequence is an effective proxy for these alternative controls. In addition, goodness-of-fit analyses confirm the fact that the best models generally include only two predictors—birth order and call-up sequence. Goodness-of-fit statistics also show that the same two-variable models are always superior to models that include only one of the two predictors. In sum, based on the collective statistical evidence, the hypothesis that controlling call-up sequence introduces a spurious factor into our analyses appears to be ruled out. Rather, controlling call-up sequence produces models that invariably do a superior job of representing the relationship between birth order and base-stealing proclivities.

Table 1. Partial correlation of birth order with attempted steals per opportunity, controlled for call-up sequence (PASW Version 17, SPSS, Inc.).

Model	B	Std. Error	Beta	t-value	Significance	Partial Correlation
	Unstandardized coefficients					
(Constant)	.707	.032		22.103	.000	
Call-up status	-.132	.059	-.290	-2.255	.025	-.165
Birth order	.131	.056	.300	2.328	.021	.170

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.086	2	.543	2.887	.058
	Residual	34.240	182	.188		
	Total	35.326	184			

a. Predictors: (Constant), Birth order, Call-up status. Dependent Variable: Log_{10} (Attempts Stealing per Opportunity). Model R=.175; R-squared=.031. Adding birth order to a model having call-up sequence already entered causes a change in R-squared of .021, F change=5.42, df=1/182, F change p=.021. With birth order included in the model, adding call-up sequence causes a change in R-squared of .027, F change=5.09, df=1/182, F change p=.025.

Table 2. Mantel-Haenszel common odds ratio results for birth order and a sibling comparison variable for attempted steals per opportunity, stratified by call-up sequence (PASW Version 17, SPSS, Inc.).

The Relationship between Attempts Stealing Per Opportunity (Sibling Comparison Variable) and Birth Order, Stratified by Call-up Sequence (*n*=177)[a]

Count

Player's call-up sequence			Birth Order -1.000000 (Older)	Birth Order 1.000000 (Younger)	Total
-1 (First)	ATTEMPTS STEALING per opportunity (Sibling comparisons)	-1 (Less)	42	1	43
		1 (More)	32	5	37
	Total		74	6	80
0 (The same year[b])	ATTEMPTS STEALING per opportunity (Sibling comparisons)	-1 (Less)	4	2	6
		1 (More)	2	7	9
	Total		6	9	15
1 (Last)	ATTEMPTS STEALING per opportunity (Sibling comparisons)	-1 (Less)	5	35	40
		1 (More)	0	42	42
	Total		5	77	82

a. The size of this sample is 177 rather than 185, as the analysis excludes 3 ties and 5 cases in sibships of three brothers where one brother's comparison value for attempted steals per opportunity was intermediate (that is, 0 rather than +1 or -1). Within call-up sequence subgroups, odds ratios may be calculated from the four relevant cells (a, b, c, d) as (a/b)/(c/d).

b. Includes 5 players who were called up before one brother but after another brother. For the 10 brothers who were called up during the same year, 4 of the 5 younger brothers attempted more steals per opportunity.

Mantel-Haenszel Common Odds Ratio Estimate[c]

Odds Ratio Estimate			10.579
ln(Estimate)			2.359
Std. Error of ln(Estimate)			.800
Asymp. Sig. (2-sided)			.003
Asymp. 95% Confidence Interval	Common Odds Ratio	Lower Bound	2.206
		Upper Bound	50.731
	ln(Common Odds Ratio)	Lower Bound	.791
		Upper Bound	3.927

c. The Mantel-Haenszel common odds ratio estimate is asymptotically normally distributed under the common odds ratio of 1.00 assumption. So is the natural log of the estimate. For the Mantel-Haenszel test, chi-square=9.87, n=177, p=.002.

Tests of Homogeneity of the Odds Ratio[d]

	Chi-Squared	df	Asymp. Sig. (2-sided)
Breslow-Day	.873	2	.646
Tarone's	.866	2	.649

d. Because neither the Breslow-Day test nor Tarone's test is statistically significant, we may conclude that the odds ratios for each subgroup are homogeneous, which permits our pooling these ratios together into a common odds ratio by means of the Mantel-Haenszel procedure.

References

Rosnow, R, and Rosenthal, R. 2003. Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, 57, 221-237.

Schwarz, A. 2010. From big leagues: Hints of sibling behavior. *New York Times* (Science Times), 25 May, p. D1**:** http://www.nytimes.com/2010/05/25/science/25birth.html.

Simpson, E. H. 1951. The interpretation of interaction in contingency tables.   , 13, 238-241.

Sulloway, F. J. and Zweigenhaft, R. L. 2010. Birth order and risk taking in athletics: A meta-analysis and study of major league baseball. *Personality and Social Psychology Review*, 14, 402-416. Published online, on 30 April 2010, as DOI: 10.1177/1088868310361241 and available at http://psr.sagepub.com/content/early/2010/04/28/1088868310361241.abstract.

.