Propensity Score Bin Bootstrapping

<< SAS code and macro >>

Bob Obenchain, PhD, FASA

Table of Contents

1. Introduction

The sample PSBB code in file PCI_PSBB.sas should be somewhat self-explanatory. After all, this code contains extensive comments.

The primary thing that does need to be explained here is how and why the SAS code provided here is somewhat different from that published in Chapter 14: “Costs and Cost-Effectiveness Analysis Using Propensity Score Bin Bootstrapping” of the book **Analysis of Observational Health Care Data Using SAS**. First, the code is now set up to analyse a (non-proprietarty) simulated dataset, PCI15K.sas7bdat, so a few minor changes were mandatory. This dataset contains 15,487 patients who received a PCI procedure with or whithout a hypothetical blood thinner; the first 10,325 patients (two-thirds of the data) were used as the primary numerical example in Chapter 7 on “Local Control.”

On the other hand, three types of key changes to the the PSBB macro have also been made.

1.1 Averages of Within PS Bin Differences

The essence of the PSBB approach is to use bootstrapping to explore the distributions of within-bin averages of outcomes (cost and/or effectiveness) for two treatment groups (treated=1 or control=0 patients.) These distributions could vary from bin-to-bin, signaling detection of patient differential response to treatment. In fact, when Propensity Scores (PSs) are estimated from patient baseline X-characteristics and patients are sorted on these PS estimates, the patients within each of the 5 bins (with adjacent PS estimates) should be much more similar to each other than they are to patients in different bins (with rather different PS estimates.) Thus, it is essential to compute differences in average treatment outcomes (treated minus control) within PS bins. Only then is it appropriate to average these within-bin outcome differences across the 5 PS bins.

Unfortunately, the published PSBB macro did not do this.

1.2 Orientation of the Cost and Effectiveness axes on the ICE plane.

In Incremental Cost-Effectiveness (ICE) analysis, the accepted convention is that each treatment difference in cost (treated minus control) is plotted vertically while the corresponding treatment difference in effectiveness (treated minus control) is plotted horizontally. Thus, movement in the downwards direction on the ICE plane is favorable to treatment (over control) because treatment is thereby becoming “less costly” than control. Similarly, moving to the right on the ICE plane is also favorable to treatment (over control) because treatment is thereby becoming “more effective” than control.

When these accepted conventions (on how treatment differences in outcomes are plotted) are not followed, published “rules” [Gold et al.(1996), Obenchain, Robinson and Swindle(2005), Willan and Briggs(2006)] on how to interpret the meaning of statistics derived from an ICE bootstrap uncertainty scatter are misleading or just plain WRONG!

Unfortunately, the published PSBB macro plotted the cost difference along the horizontal axis.

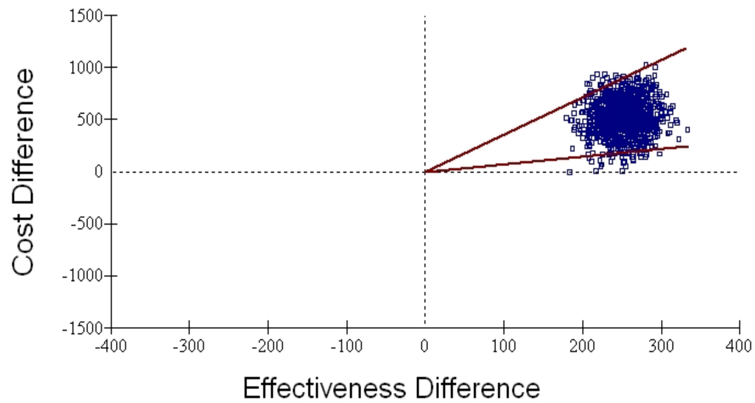
1.3 BCa Confidence Interval Calculations

The published SAS code for implementing the DiCiccio and Efron(1996) approach to bias corrected and accelerated (BCa) confidence intervals is both extensive and complicated. Much more simple and straight-forward sorts of bias corrections have been proposed and evaluated in the cost-effectiveness literature [Stinnett(1996), Briggs and Fenn(1998), Briggs, Mooney and Wonderling(1999).] While these alternatives are being compared, the PSBB code provided here focusses exclusively on traditional, “percentile” methods using bootstrap order statistics to form classical non-parametric confidence (or tolerance) intervals.

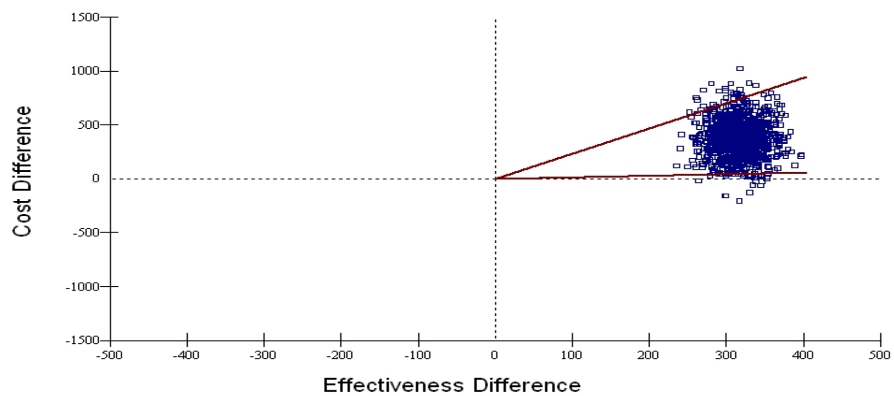
2. Example: PSBB analysis of the PCI15K dataset

The analyses of the PCI15K.sas7bdat dataset presented in Chapter 7 show that considerable treatment selection bias (confounding) is present in this dataset. Patients who were more severely diseased (e.g. "stent" deployed, lower "ejfract" and/or higher "ves1proc") were much more likely to be assigned to trtm=1 (usual PCI care augmented with a hypothetical blood thinner) than in the "control" trtm=0 (usual PCI care alone.) The two figures below illustrate subtle differences between the ICE analyses that result from these data when using the MS Windows applications of Obenchain(2001, 2003). The top figure analyzes the data using the "ICEplane" application assuming the two treatment cohorts are comparable (unbiased relative to each other.) The bottom figure analyzes the data using the "ICEpsbbs" application that adjusts for bias in the same way as the PSBB SAS macro provided here. Both Figures assume that the **Shadow Price of Health** has been set at **Lambda** = $10,000/surv6mo, so that the horizontal (effectiveness) axis can also be scaled in $ cost units.

"ICEplane" application: Data Assumed Unbiased

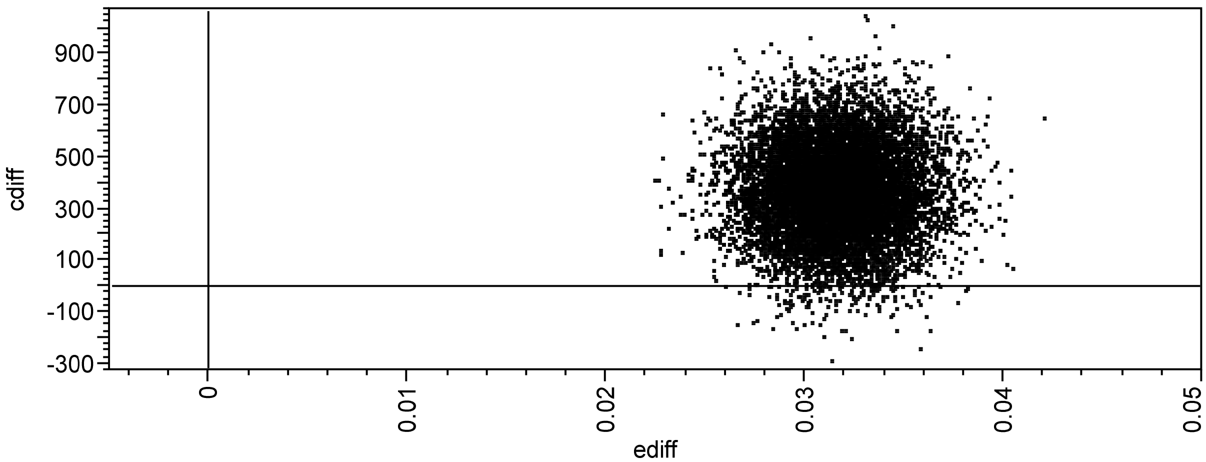


ICEpsbbs application: Data Assumed Biased



While the differences bewteen the two analyses above are not large, it is clear that PSBB adjustment is more favorable to trtm=1 on both dimensions, cost and effectiveness. Specifically, the bootstrap scatter moves both to the right (even more effective) and downwards (towards less costly …but not there yet!) On the other hand, both scatters above do have the same basic, **overall interpretation**: Trtm=1 is **more effective** but **more costly** than trtm=0.

Note also that bootstrap distribution of joint cost and effectiveness differences (trtm=1 minus trtm=0) appears to be very close to a bivariate normal distribution with uncorrelated components:



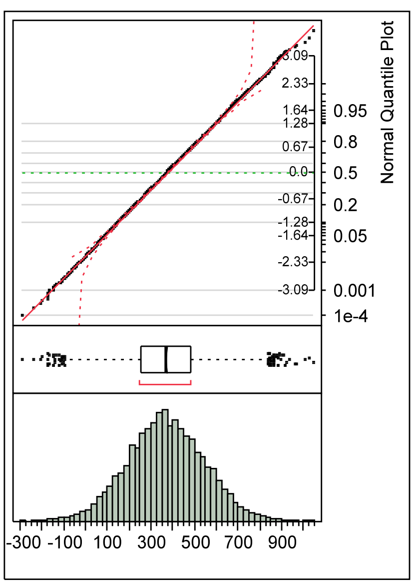
Covariances

	cdiff	ediff
cdiff	29774.887	-0.00023
ediff	-0.00023	0.00001

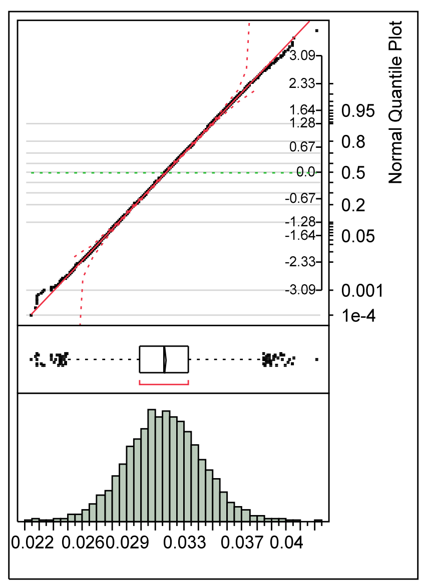
Correlations

	cdiff	ediff
cdiff	1.0000	-0.0005
ediff	-0.0005	1.0000

cdiff marginal distibution



ediff marginal distribution



Cdiff Mean	$369	Ediff Mean	0.0316
Std Dev	$173	Std Dev	0.00250
N	10000	N	10000

When the **Shadow Price of Health** is taken to be **Lambda** = $10,000/surv6mo, as suggested on page 3, the above PSBB marginal statistics show that trtm=1 achieves an average, incremental increased effectiveness worth $316 at an average, incremental increased cost of $369. In other words, on overall average, trtm=1 almost (but not quite) pays for itself.

While the PSBB SAS macro does not compute the wedge-shaped 95% ICE confidence regions displayed in the figures on page 3, it does make basic PSBB calculations available to health care professionals who typically prefer to perform analyses using SAS.

4. References

Briggs AH, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. ***Health Economics*** 1998; **7**(8): 723-740.

Briggs AH, Mooney CZ, Wonderling DE. Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using Monte Carlo simulation. *Stat.Med.* 1999; 18(23): 3245-62.

DiCiccio TJ, Efron B. Bootstrap Confidence Intervals. ***Statistical Science*** 1996; 11: 189-228.

Gold MR, Siegal JE, Russell LB, Weinstein MC. ***Cost-Effectiveness in Health and Medicine***. New York: Oxford University Press, 1996.

Obenchain RL. ICEplane: a Windows application for Incremental Cost-Effectiveness inference via Wedge-Shaped Bootstrap Confidence Regions (two unbiased samples.) 2001-2010. http://members.iquest.net/~softrx/ICErefs.htm

Obenchain RL. ICEpsbbs: a Windows application for Incremental Cost-Effectiveness inference via Propensity Score Bin Bootstrapping (two biased samples.)" 2003-2010. http://members.iquest.net/~softrx/ICErefs.htm

Obenchain RL, Robinson RL, Swindle RW. "Cost-effectiveness inferences from bootstrap quadrant confidence levels: three degrees of dominance." ***Journal of Biopharmaceutical Statistics*** 2005; 15(3): 419-436.

Stinnett A. Adjusting for Bias in C/E Ratio Estimates. ***Health Economics*** 1996; 5: 470-472.

Willan AR, Briggs AH. ***Statistical Analysis of Cost-Effectiveness Data***. John Wiley & Sons. Chichester, West Sussex, England, 2006.

The UNIVARIATE Procedure
Variable: pstrtm1

trtm=0

Moments			
N	8476	**Sum Weights**	8476
Mean	0.42370891	**Sum Observations**	3591.35673
Std Deviation	0.11589255	**Variance**	0.01343108
Skewness	0.56759539	**Kurtosis**	0.33596534
Uncorrected SS	1635.51828	**Corrected SS**	113.828438
Coeff Variation	27.3519276	**Std Error Mean**	0.00125881

Basic Statistical Measures			
Location		**Variability**	
Mean	0.423709	**Std Deviation**	0.11589
Median	0.411263	**Variance**	0.01343
Mode	0.407945	**Range**	0.68806
		Interquartile Range	0.14694

Note: The mode displayed is the smallest of 2 modes with a count of 13.

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	336.5949	**Pr > |t|**	<.0001
Sign	**M**	4238	**Pr >= |M|**	<.0001
Signed Rank	**S**	17962763	**Pr >= |S|**	<.0001

Quantiles (Definition 5)	
Quantile	**Estimate**
100% Max	0.816970
99%	0.749585
95%	0.644210
90%	0.574848
75% Q3	0.491191
50% Median	0.411263
25% Q1	0.344251
10%	0.283428
5%	0.249232
1%	0.203659
0% Min	0.128913

The UNIVARIATE Procedure
Variable: pstrtm1

trtm=0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0.128913	1571	0.810913	1907
0.132833	1419	0.811152	5691
0.134831	1979	0.813541	1250
0.135478	2125	0.814380	3339
0.135478	1320	0.816970	2177

The UNIVARIATE Procedure
Variable: pstrtm1

trtm=1

Moments			
N	7011	**Sum Weights**	7011
Mean	0.487754	**Sum Observations**	3419.64328
Std Deviation	0.12792397	**Variance**	0.01636454
Skewness	0.273255	**Kurtosis**	-0.4074879
Uncorrected SS	1782.66011	**Corrected SS**	114.715436
Coeff Variation	26.227149	**Std Error Mean**	0.00152778

Basic Statistical Measures			
Location		**Variability**	
Mean	0.487754	**Std Deviation**	0.12792
Median	0.479662	**Variance**	0.01636
Mode	0.468464	**Range**	0.67441
		Interquartile Range	0.17119

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	319.2559	**Pr > |t|**	<.0001
Sign	**M**	3505.5	**Pr >= |M|**	<.0001
Signed Rank	**S**	12290283	**Pr >= |S|**	<.0001

Quantiles (Definition 5)	
Quantile	**Estimate**
100% Max	0.819533
99%	0.782392
95%	0.718481
90%	0.675520
75% Q3	0.565431
50% Median	0.479662
25% Q1	0.394237
10%	0.330789
5%	0.294043
1%	0.224029
0% Min	0.145124

The UNIVARIATE Procedure
Variable: pstrtm1

trtm=1

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0.145124	9932	0.808498	14295
0.145124	9177	0.810063	10525
0.145812	10330	0.811152	14493
0.155559	11050	0.814380	9433
0.160140	12369	0.819533	11516

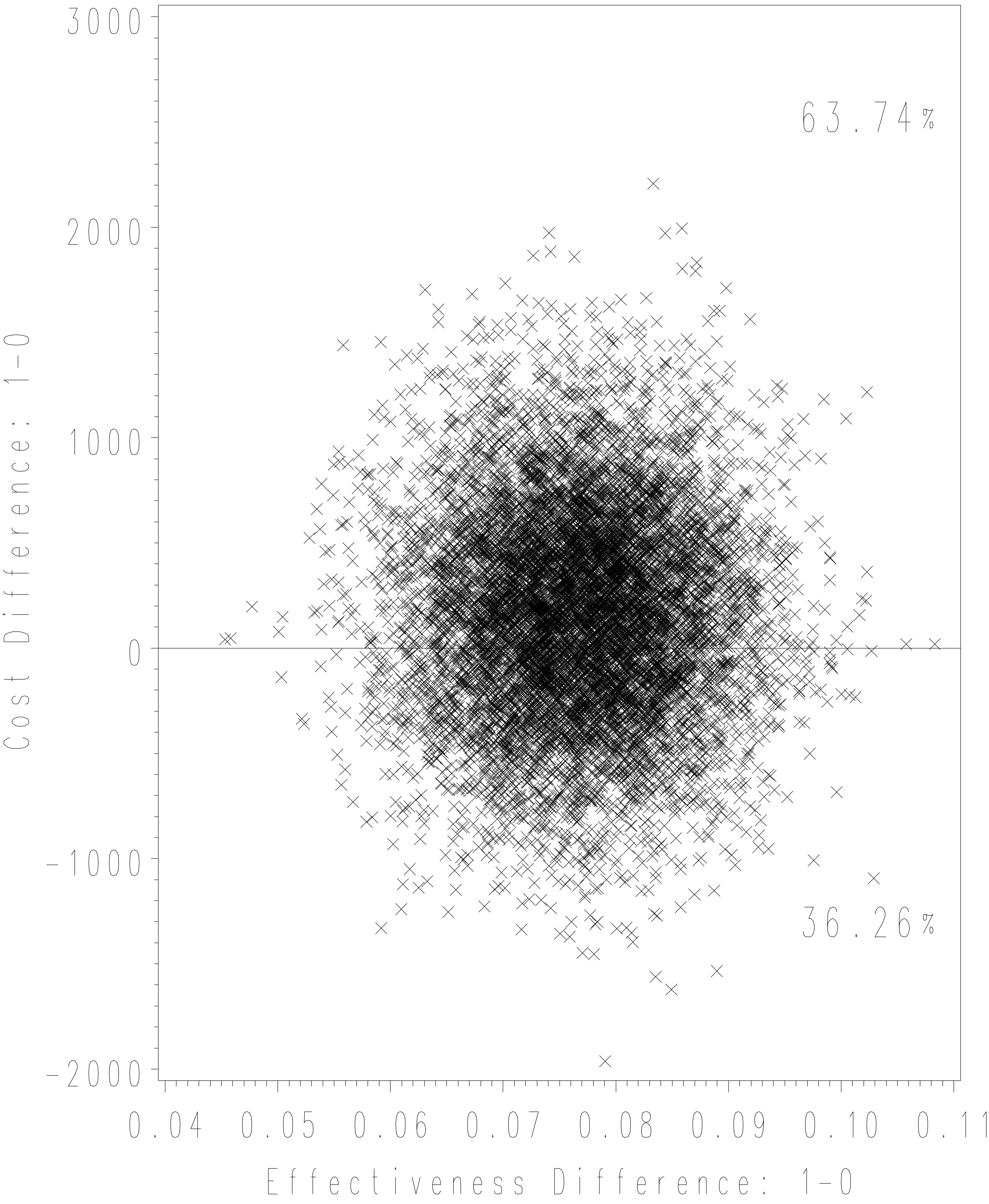
The FREQ Procedure

Frequency Percent Row Pct Col Pct

Table of bin_ps by trtm			
	trtm(trtm)		
bin_ps	0	1	Total
1	2182 14.09 70.46 25.74	915 5.91 29.54 13.05	3097 20.00
2	1995 12.88 64.42 23.54	1102 7.12 35.58 15.72	3097 20.00
3	1747 11.28 56.41 20.61	1350 8.72 43.59 19.26	3097 20.00
4	1453 9.38 46.93 17.14	1643 10.61 53.07 23.43	3096 19.99
5	1099 7.10 35.45 12.97	2001 12.92 64.55 28.54	3100 20.02
Total	8476 54.73	7011 45.27	15487 100.00

		Total Cost			Survive >= 6 Months		
		N	Mean	Std	N	Mean	Std
bin_ps	trtm						
1	0	2182	18734.7271	9589.9619	2182	0.9753	0.1554
	1	915	19657.6710	13863.8763	915	0.9934	0.0808
2	0	1995	17530.1082	7865.9116	1995	0.9789	0.1436
	1	1102	18251.5658	8238.7031	1102	0.9900	0.0995
3	0	1747	17376.0986	7510.8909	1747	0.9691	0.1731
	1	1350	17260.0194	8082.2969	1350	0.9874	0.1115
4	0	1453	18422.2666	10581.3895	1453	0.9553	0.2068
	1	1643	18579.3291	10675.7178	1643	0.9890	0.1041
5	0	1099	19521.8082	12139.3486	1099	0.9063	0.2916
	1	2001	19694.2991	14455.1767	2001	0.9830	0.1293

ICE Quadrant Confidence Levels

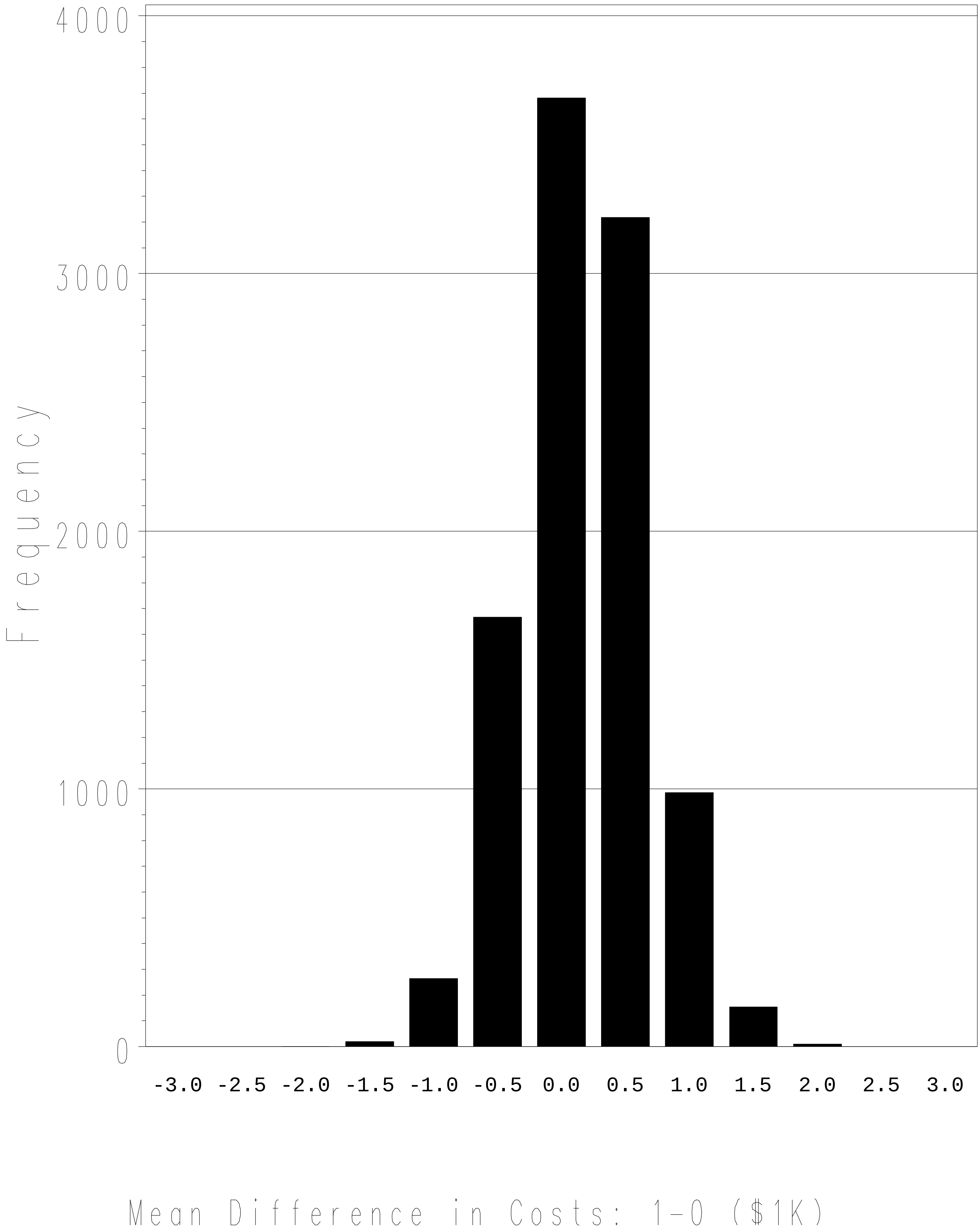


The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
cardcost_diff	Average for cardcost Diff: Trt2-Trt1	10000	173.1638625	490.3879419	-1962.83	2206.82
surv6mo_diff	Average for surv6mo Diff: Trt2-Trt1	10000	0.0767559	0.0076970	0.0453389	0.1083252

Bootstrap 95% Pct. Conf. Limits for cardcost and surv6mo

Obs	cardcost_lcl	cardcost_ucl	surv6mo_lcl	surv6mo_ucl
1	-772.507	1160.17	0.061678	0.091806



4000
3000
2000
1000
0
Frequency
-3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0
Mean Difference in Costs: 1-0 ($1K)