What factors earn a “Survivor” contestant a spot in a future season?

Brianna Beitling

Stat 835: Categorical Analysis

University of Kansas Medical Center

Neal Montgomery, Instructor

15 May, 2019

Abstract

The television show “Survivor” is the subject of interest for many as one of the longest-running reality shows on television, and an ongoing, ever-evolving social experiment that forces people of different walks of life together to form a miniature society for a little over a month. The primary goal of a contestant on the show has shifted over its 38 season run from simply winning the million dollar prize at the end to creating an entertaining character who makes moves and engenders a personality worthy of being asked to compete in subsequent seasons. This paper examines what it takes for the “Survivor” casting crew to ask a contestant to compete again, including personality, strategy in the game, and physical and mental strength.

What factors earn a “Survivor” contestant a spot in a future season?

“Survivor” is a television show with a seemingly simple concept: outlast all but one or two other contestants, remaining in the game long enough to reach the final tribal council. And, on top of that, the ultimate goal is to remain friendly enough or impressive enough to the contestants who have been voted out so that they vote to reward you with the million dollar prize at the end rather than the person (or persons) sitting next to you. For the first seven seasons of “Survivor,” this was the unquestionable goal of most of the contestants participating. Season eight of the show represented something of a turning point: season eight is otherwise known as “Survivor: All-Stars,” and featured all of the “best” returning players from the previous seven seasons.

In all 38 (to date) seasons of “Survivor,” twelve have featured a roster containing at least two returning players. Most “Survivor” players have not won the game, so returning to the show is another chance to win and, perhaps more importantly, another chance to leave a permanent imprint as a legend in the game. Arguably, creating a character entertaining or strategic enough to be asked back by the “Survivor” casting department might now be the primary objective for new contestants, perhaps even moreso than the million dollars.

This study will investigate the question of what it takes to be asked to compete again on “Survivor.” Contestants who have competed on a subsequent season will be segregated from contestants who have not in order to determine whether they differ in age, placement in the game, individual immunity wins, tribal council success, hidden immunity idols found and played, and average confessionals in the season. It is predicted that contestants who are younger, who placed highly in the game, who won more individual immunities, who voted correctly more of the time, who found and successfully played hidden immunity idols, and who are granted

extra airtime via more confessionals are more likely to be asked to compete again by the casting department.

Method

Predictor Variables

Age

Age is defined as the “Survivor” contestant’s age during the season in which they competed. “Survivor” casting seems to favor younger castaways, so it is predicted that age will have a negative relationship with the probability of being asked to return.

Placement

Placement is defined as the “Survivor” contestant’s finish place in the season. This ranges from 1 (the winner) to the maximum number of contestants in the season (usually 16, 18, or 20). In cases where the contestant was voted out but was not officially out of the game yet (such as when Redemption Island has been in play, giving contestants an opportunity to compete in a challenge to earn their way back into the game), their finish after being officially removed from the game was recorded. Higher placing castaways are more memorable, and thus are predicted to be invited back more often.

Individual Immunities Won (IndImmunities)

Individual immunities won is defined as the number of individual immunity challenges (typically occurring after the merge and dissolution of competing tribes to a single tribe) won by the contestant over the course of the season. Individual immunities are won by prevailing in mental and physical challenges, and render the winner safe at the next tribal council (that is, no one can vote for that contestant to be booted from the game). Contestants who are formidable physical threats are memorable, and thus are predicted to be more frequent returnees.

Tribal Council Correct Percentage (TC_Pct)

Tribal council correct percentage is defined as the number of "correct" votes made by the contestant (that is, votes for the person who was actually booted from the game in that tribal council) divided by the number of tribal councils attended by that contestant. Contestants who are in dominant alliances, making moves, and controlling the game are remembered for their strategic game. Therefore, it is predicted that these contestants will return more frequently.

Idol Score (Idol_Score)

Idol score is a sum score calculated based on the contestant's experience finding and playing hidden immunity idols. Hidden immunity idols can be found by contestants in "Survivor" and used at varying points in the game to grant the contestant (or a fellow contestant) immunity after voting occurs but before the votes are read aloud (in most cases). In an attempt to reward contestants for better idol plays (whether played for the contestant him/herself or on a fellow contestant), the idol score is calculated as follows:

$$Idol\_Score = Idols\_Found + Idols\_Played + Votes\_Voided + 10(Boot\_Avoided) + 5(Tie\_Avoided)$$

where Idols_Found is the number of hidden immunity idols found by the contestant on the season, Idols_Played is the number of hidden immunity idols played by the contestant on the season (contestants can be voted off with an idol in his or her pocket which is then useless), Votes_Voided is the number of votes cancelled out by the play(s) of the hidden immunity idol(s), Boot_Avoided is the number of times playing a hidden immunity idol saved that contestant (or the contestant the idol was played for) from being voted out that night, and Tie_Avoided is the number of times playing a hidden immunity idol saved that contestant (or the contestant the idol

was played for) from being deadlocked in a tie vote. Contestants who find and play hidden immunity idols well are predicted to be more likely to return.

Average Confessionals (Average_Confessionals)

Average confessionals is defined as the number of times a one-on-one interview with the contestant aired throughout the season divided by the number of episodes in which the player appeared as a contestant (and not as a jury member). Players who have more personality or who have more strategic game to explain tend to be featured in confessionals; these players are thus predicted to be more likely to be asked to compete again.

Analysis

As these predictor variables are all expected to be linearly related to the probability of being asked to compete again, a multiple logistic regression model will be fit to the data:

$$logit[\pi(InvitedBack)] = \alpha + \beta_1(Age) + \beta_2(Placement) + \beta_3(IndImmunities) + \beta_4(TC\_Pct) + \beta_5(Idol\_Score) + \beta_6(Average\_Confessionals)$$

Although the 38th season of "Survivor" is currently airing, the data will span seasons 1 through 34 since no players have yet returned from seasons 35, 36, and 37. The latter three seasons will be used at the end of this study as a prediction dataset to determine which of those contestants may be likely to return. To check the assumptions of the logistic regression model, univariate models will be fit and graphs produced to confirm each predictor variable's relationship with the response variable; scatterplots and a correlation matrix will also check for the presence of multicollinearity. The rest of the assumptions of the logistic regression model (independence of observations and sufficient sample size) will be assumed to be true, with acknowledgement that due to some returning players competing multiple times, there is not full independence in the data. For contestants who competed on multiple seasons, their data is treated as independent.

Whether the contestant was invited back, in that case, is determined by whether he or she competed in a season subsequent to the current one (so if a contestant competed in seasons 15, 16, and 20, the value of InvitedBack is 1 (true) for seasons 15 and 16, and 0 (false) for season 20).

Results

Assumptions

To assess the assumption of linear relationship between the probability of being invited back and each predictor variable, univariate logistic regression analyses were performed. The test statistic for these analyses is the Wald Chi-Squared test statistic. Age was not found to be a significant predictor at the α = .05 level ($\chi^2 = 2.59$, $p = .11$) and was not included in further analysis. Placement ($\chi^2 = 38.77$, $p < .01$), individual immunities won ($\chi^2 = 32.4865$, $p < .01$), tribal council correct percentage ($\chi^2 = 29.00$, $p < .01$), hidden immunity idol score ($\chi^2 = 6.26$, $p = .01$), and average confessionals ($\chi^2 = 43.48$, $p < .01$) were all significantly linearly related to the response variable of being asked to play “Survivor” again at the α = .05 level. The univariate logistic regression output is given in Tables A through F in the Appendix.

The linear relationship between the probability of being invited back and each predictor variable was also visually assessed via fit plots; this also confirmed the expected direction of the relationship. Graphs A through E in the Appendix show linear relationships for each predictor variable with InvitedBack, as well as the expected direction of the relationship for each.

Another assumption of the multiple logistic regression model is no multicollinearity between predictor variables. This was assessed via both a scatterplot matrix and correlation matrix (Figure A and Table G in the Appendix respectively). Several variables were fairly highly correlated, as could be expected given that hidden immunity idol plays, individual immunity

wins, and voting correctly at tribal council are all associated with a contestant's placement in the game. The variable with the most offensive multicollinearity and least to uniquely add to the model, Idol_Score, was removed at this stage.

Fitting the Model

Tables H through K in the Appendix display output associated with the fitting of the full multivariate logistic regression model. The fitted model is:

$$logit(\hat{\pi}) = -2.9978 - 0.0606(Placement) + 0.2075(IndImmunities) + 0.0112(TC\_Pct) + 0.4352(Average\_Confessionals)$$

The overall model fits well ($\chi^2 = 72.24$, $p < .01$), with each of the predictor variables being either significant or close to significant; placement ($\chi^2 = 3.78$, $p = .05$) and individual immunities won ($\chi^2 = 3.09$, $p = .08$) were both close to significant, while tribal council correct percentage ($\chi^2 = 4.50$, $p = .03$) and average confessionals ($\chi^2 = 31.78$, $p < .01$) were both significant at the $\alpha = .05$ level. Some of the predictors' significance decreased due to multicollinearity; however, the predictors were significant in the multivariate model at the $\alpha = .10$ level and thus remained.

Using the Model for Prediction

Using the fitted model and the prediction dataset of recent seasons' (35, 36, and 37) contestants, predicted probabilities of returning were calculated for the remaining castaways. Two players from the set of 58 were predicted to return with probability greater than 50.0%. These players and their relevant variables are listed in the Appendix in Table L.

Discussion

Higher placement of a contestant on a "Survivor" season, winning individual immunities, finding and playing hidden immunity idols correctly, voting correctly at tribal council, and having many confessionals in a season are all associated with a greater chance of being asked to

play “Survivor” again in a subsequent season. Finding and playing hidden immunity idols correctly does not provide much information about a contestant’s likelihood of being invited back, however, with all of the other variables present. Therefore, if a new player is entering “Survivor” hoping to make a career out of participating again and again in the reality television show, he or she should strive to stay in the game (which is always the goal), to show mental and physical prowess by winning individual immunity challenges, to search for and correctly play hidden immunity idols strategically to avoid boots and ties, and to display a personality and/or strategic game interesting enough that editors use his or her interviews frequently in the final cut. Two contestants in recent seasons have done well enough in all of these facets of the game that they are predicted to return: Domenick Abbate and Nick Wilson. Future seasons’ rosters in time will tell whether these predictions become reality.

References

Buff. (2019, February 18). Re: Survivor: EDGE of EXTINCTION ~ Confessional Count [Web log comment]. Retrieved May 07, 2019, from https://www.tapatalk.com/groups/survivorsucks/survivor-edge-of-extinction-confessional-count-t134773.html

Pitman, J. (2019). The True Dork Times. Retrieved May 03, 2019, from https://www.truedorktimes.com/index.htm

Survivor and Big Brother By The Numbers [Audio blog interview]. (2017, March 06). Retrieved March 29, 2019, from https://robhasawebsite.com/rhappy-hour-survivor-big-brother-by-the-numbers/

Survivor (U.S.). (2019, April). Retrieved April 05, 2019, from https://survivor.fandom.com/wiki/Survivor_(U.S.)

Appendix

Univariate Logistic Regression Analyses

Table A. Logistic Regression for Variable Age

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.8605	0.3554	5.8613	0.0155
Age	1	-0.0169	0.0105	2.5857	0.1078

Table B. Logistic Regression for Variable Placement

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.2688	0.1929	1.9415	0.1635
Placement	1	-0.1356	0.0218	38.7705	<.0001

Table C. Logistic Regression for Variable IndImmunities

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.7457	0.1241	197.9857	<.0001
IndImmunities	1	0.5430	0.0953	32.4865	<.0001

Table D. Logistic Regression for Variable TC_Pct

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.7698	0.2932	89.2531	<.0001
TC_Pct	1	0.0226	0.00419	28.9993	<.0001

Table E. Logistic Regression for Variable Idol_Score

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4762	0.1058	194.7379	<.0001
Idol_Score	1	0.0567	0.0227	6.2565	0.0124

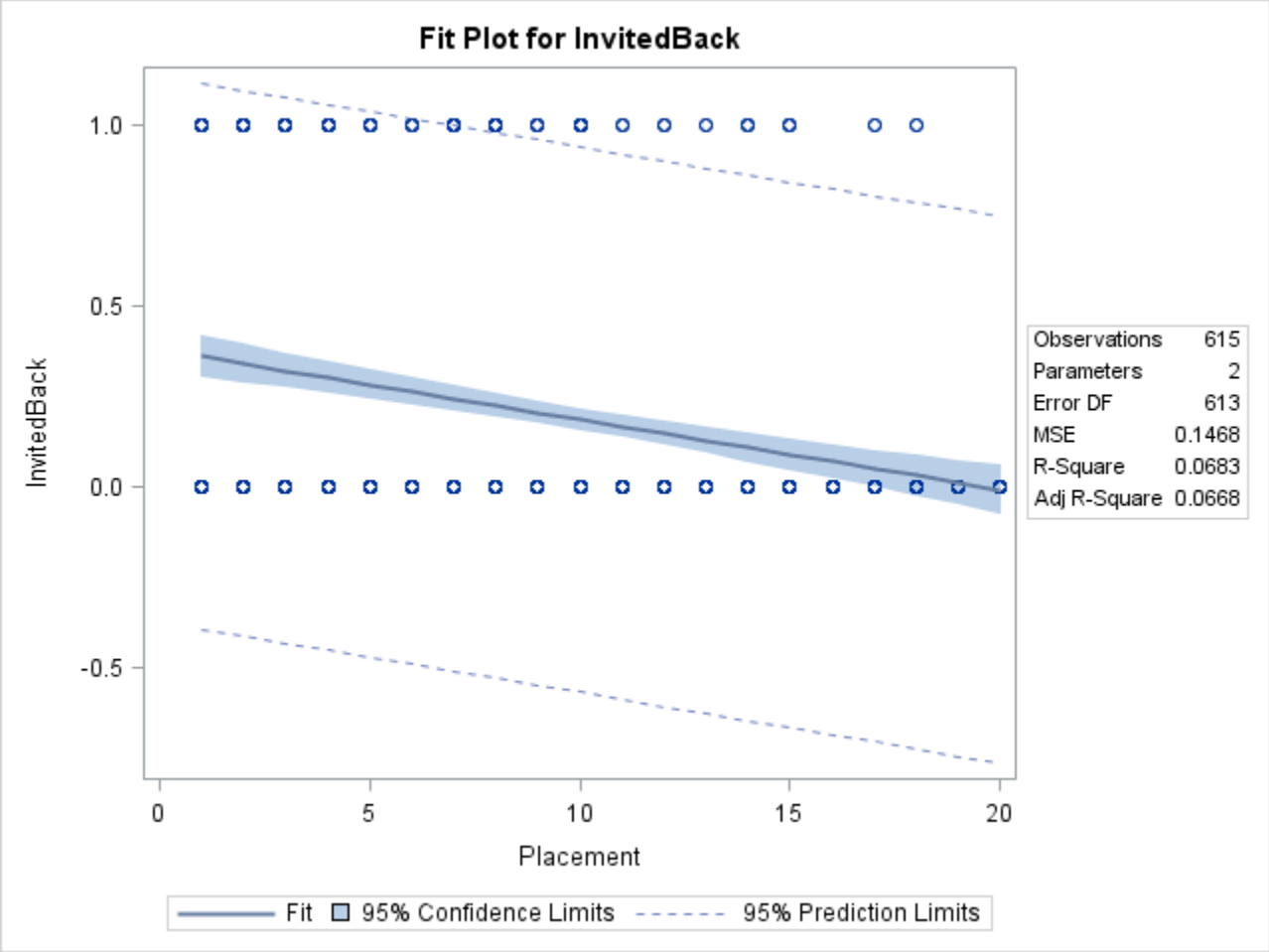
Table F. Logistic Regression for Variable Average_Confessionals

Analysis of Maximum Likelihood Estimates

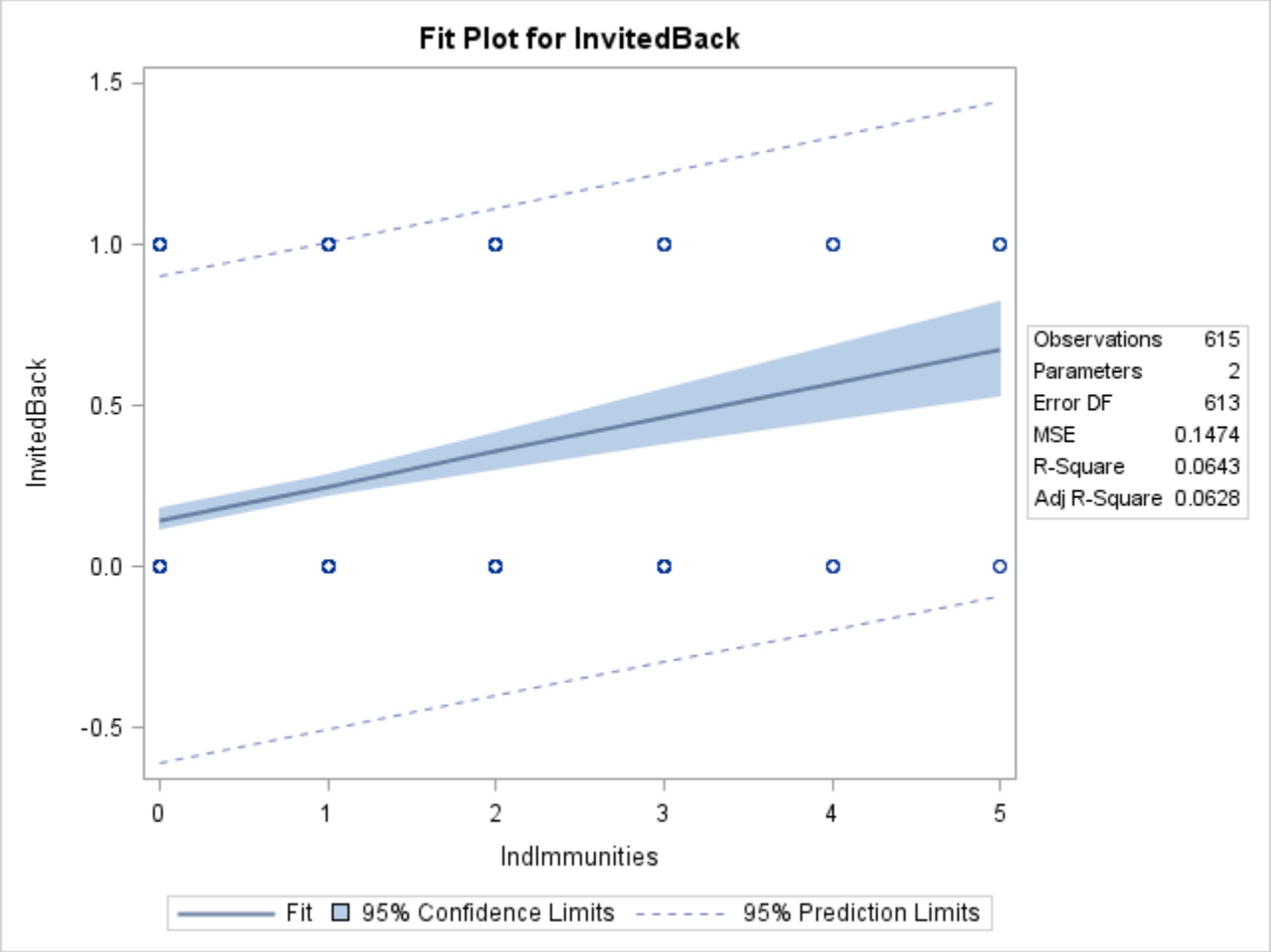
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.8658	0.2580	123.3922	<.0001
Average_Confessional	1	0.4823	0.0731	43.4784	<.0001

Univariate Graphical Relationship of Predictor Variables with InvitedBack

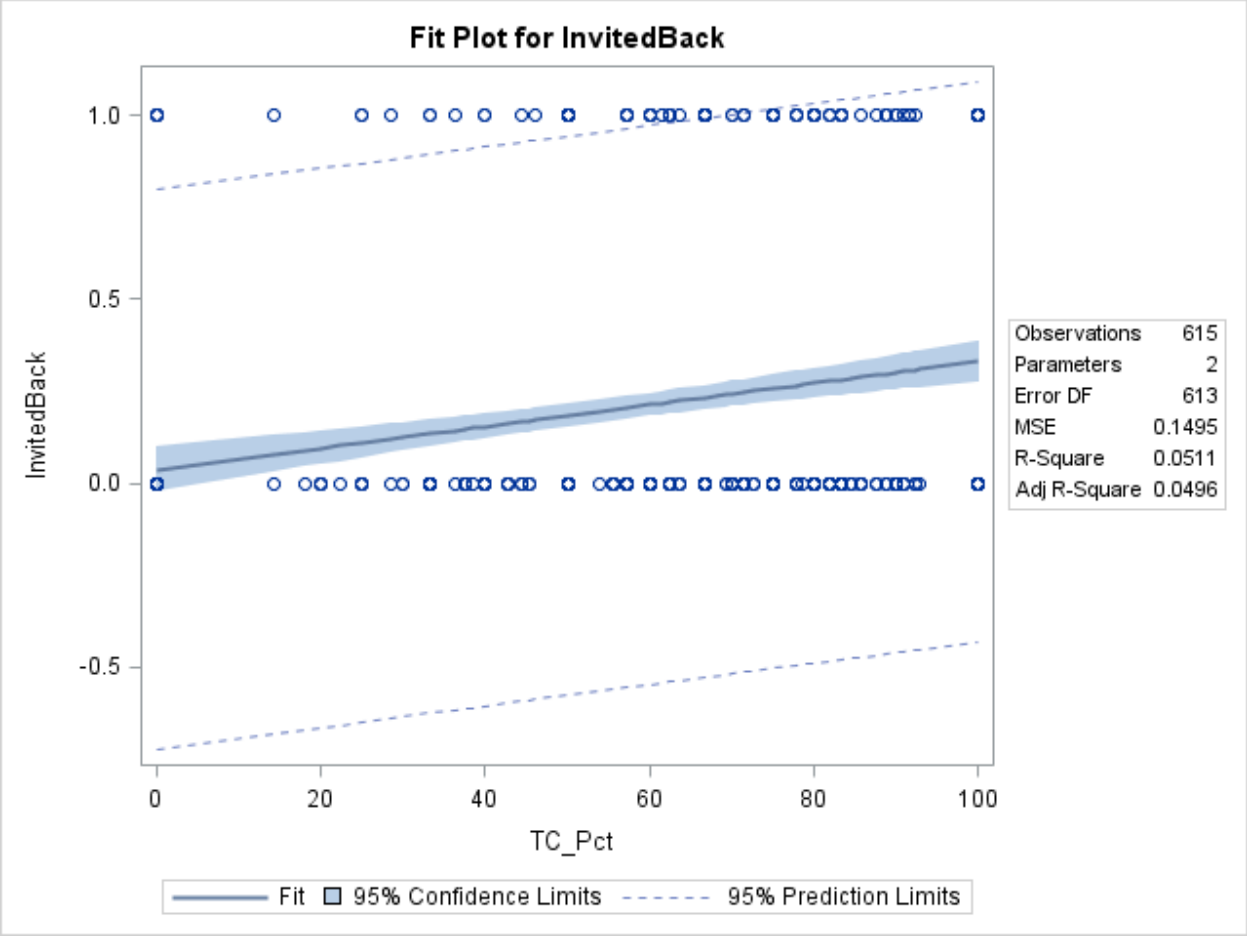
Graph A. Fit Plot for InvitedBack by Placement



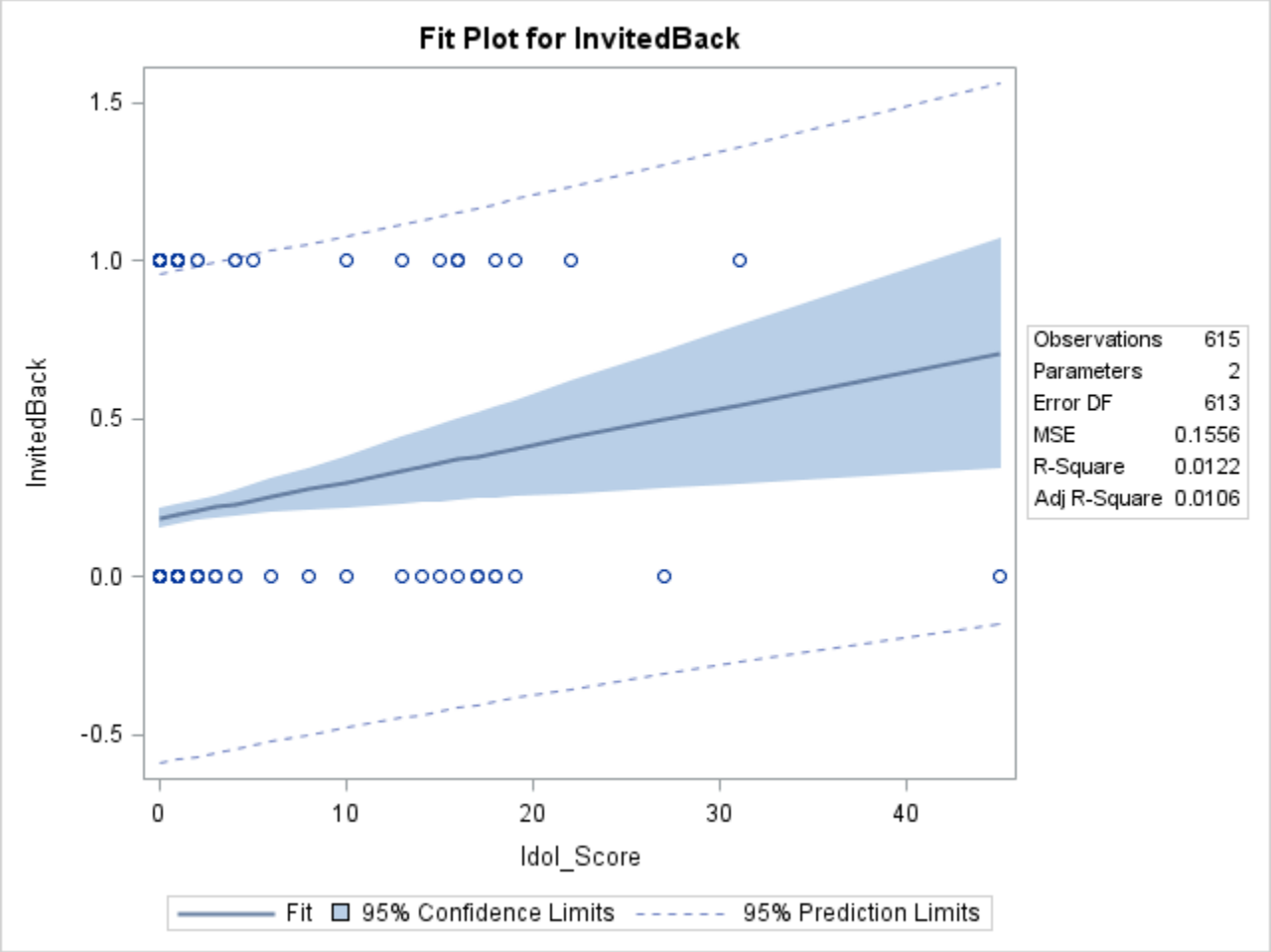
Graph B. Fit Plot for InvitedBack by IndImmunities



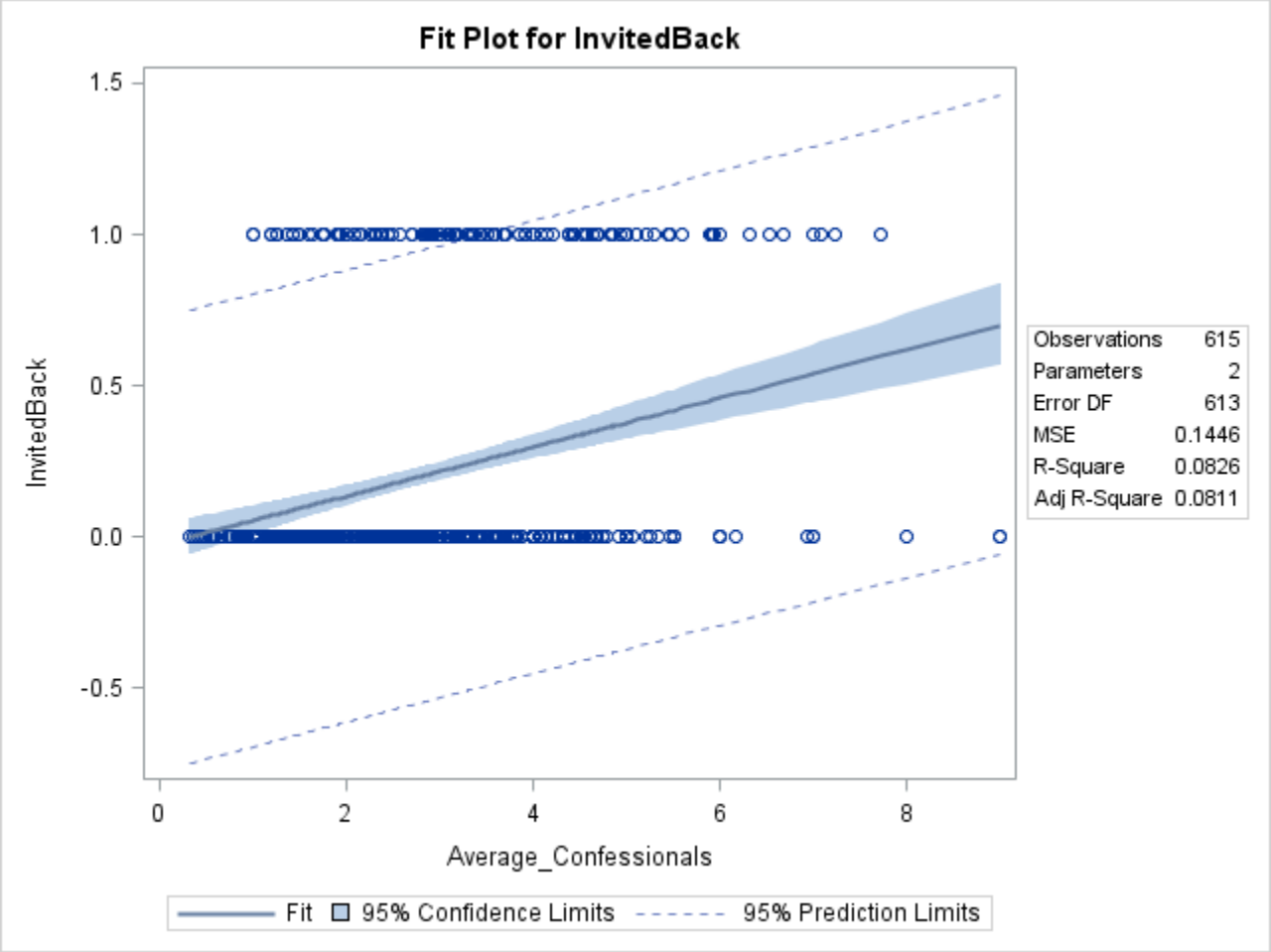
Graph C. Fit Plot for Invited Back by TC_Pct



Graph D. Fit Plot for InvitedBack by Idol_Score



Graph E. Fit Plot for InvitedBack by Average_Confessionals



Figures Examining Multicollinearity

Figure A. Scatterplot Matrix of InvitedBack and the Predictor Variables

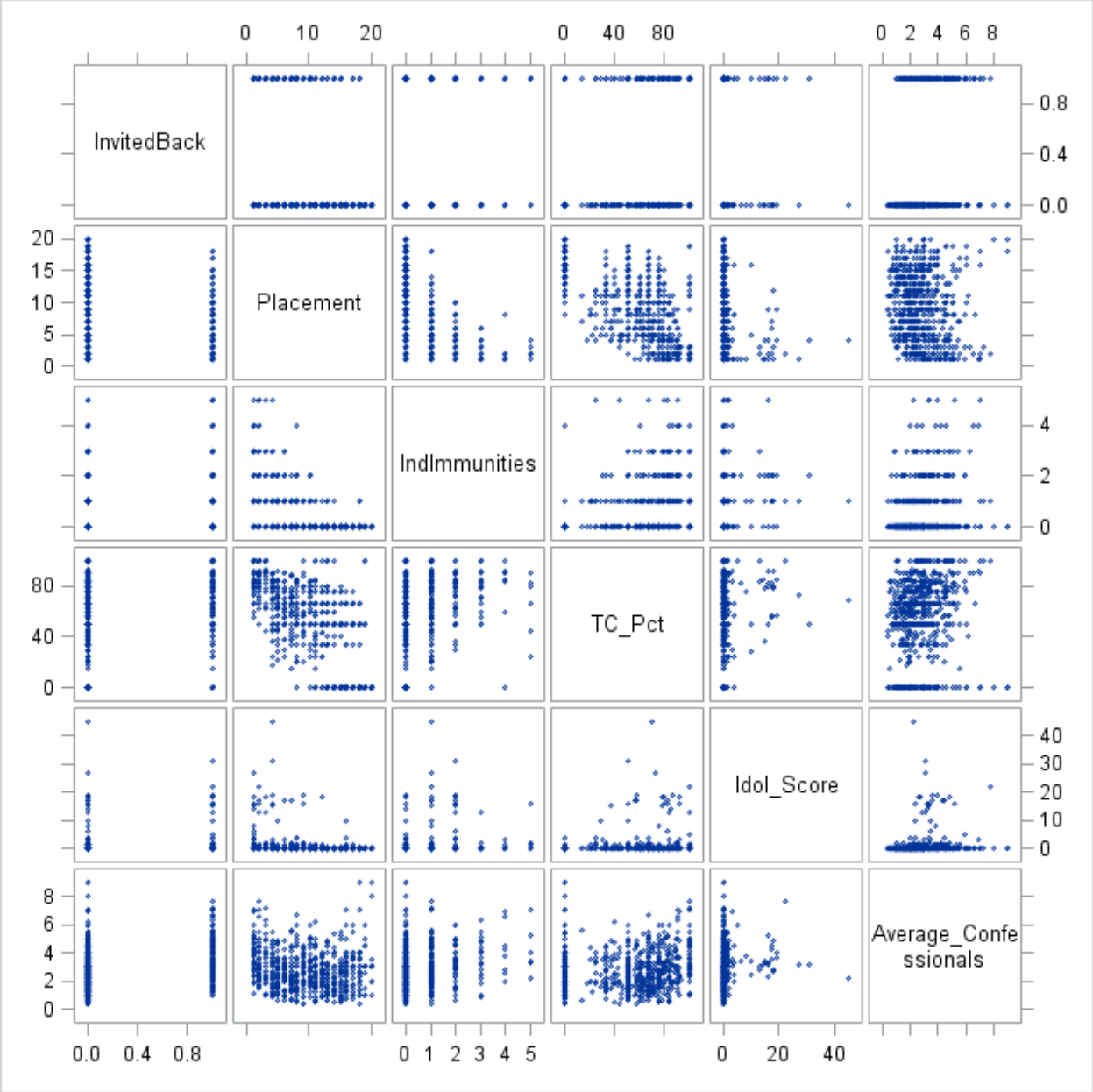


Table G. Correlation Matrix of InvitedBack and the Predictor Variables

Pearson Correlation Coefficients, N = 615 Prob > |r| under H0: Rho=0						
	InvitedBack	Placement	IndImmunities	TC_Pct	Idol_Score	Average_Confessionals
InvitedBack	1.00000	-0.26136 <.0001	0.25360 <.0001	0.22616 <.0001	0.11062 0.0060	0.28734 <.0001
Placement	-0.26136 <.0001	1.00000	-0.52007 <.0001	-0.65254 <.0001	-0.21397 <.0001	-0.12465 0.0020
IndImmunities	0.25360 <.0001	-0.52007 <.0001	1.00000	0.28723 <.0001	0.19292 <.0001	0.21859 <.0001
TC_Pct	0.22616 <.0001	-0.65254 <.0001	0.28723 <.0001	1.00000	0.11824 0.0033	0.10176 0.0116
Idol_Score	0.11062 0.0060	-0.21397 <.0001	0.19292 <.0001	0.11824 0.0033	1.00000	0.12516 0.0019
Average_Confessionals	0.28734 <.0001	-0.12465 0.0020	0.21859 <.0001	0.10176 0.0116	0.12516 0.0019	1.00000

Multivariate Logistic Regression Analysis

Table H. Model Fit Statistics

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	609.085	527.646
SC	613.507	549.754
-2 Log L	607.085	517.646

Table I. Model Overall Fit

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	89.4388	4	<.0001
Score	91.2836	4	<.0001
Wald	72.2399	4	<.0001

Table J. Analysis of Maximum Likelihood Estimates

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.9978	0.5985	25.0845	<.0001
Placement	1	-0.0606	0.0312	3.7766	0.0520
IndImmunities	1	0.2075	0.1179	3.0946	0.0786
TC_Pct	1	0.0112	0.00526	4.5012	0.0339
Average_Confessional	1	0.4352	0.0772	31.7824	<.0001

Table K. Odds Ratio Estimates

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
Placement	0.941	0.885	1.001
IndImmunities	1.231	0.977	1.551
TC_Pct	1.011	1.001	1.022
Average_Confessional	1.545	1.328	1.798

Prediction of Returning Players

Table L. Contestants Predicted to Return

Obs	Name	Placement	IndImmunities	TC_Pct	Average_Confessionals	predictval
1	Domenick Abbate	2	3	80	4.79	0.61872
2	Nick Wilson	1	3	76.92307692	3.79	0.51875