Marketing in a Digital World - Implications of Internet Core Trends for TV Advertising

Yingzi Ma

MK711

Prof. Paul D. Berger

December 11, 2018

Introduction

Has the era of TV advertising ended? The booming of AI-driven website ads, social media advertising, location-based customer targeting and augmented reality advertising opens a new chapter for advertising. Traditional advertising such as TV and physical catalog seems to be dying. However, a survey conducted in 2018 January by Pew Research suggested otherwise. The TV is still inseparable for the majority of the population. Most people find it is harder to give up TV as compared to the social media. In order to understand the characteristics of people who enjoy watching TV, this paper focuses on analyzing survey data and examining the implications of internet core trends for TV advertising.

In this paper, various statistical tools are used. Logistic regression with backward selection and classification tree are used. Two models are developed in SAS. Model validation and utility are carefully examined through hypothesis testing to study the significance of explanatory variables.

Literature Review

Is TV Advertising Dead?

"TV advertising offers unique advantages over more transactional forms of online marketing."[1] E-commerce companies such as Wayfair invests heavily in TV advertising. They consider TV as an indispensable part of their advertising portfolio. Although TV advertising is generally expensive, the company can utilize it to attract customer attention for an extended period of time and tell a story about the brand. A good TV commercial can stick in the public's mind for a long period of time which gives a huge return of investment benefit.

[1] Benjamin Schroeder and Nathan Vierling-Claassen. "Optimizing TV Advertising Toward Return on Investment." Wayfair. August 22, 2018. Accessed December 10, 2018. https://tech.wayfair.com/2018/08/optimizing-tv-advertising-toward-return-on-investment/

A study from Adobe which surveyed 1,000 U.S. TV buyers suggested that TV is ranked higher than most digital formats by marketers[2]. However, for individual TV shows, audiences, especially younger viewers continue to decrease. Millennials have gained access to a broader selection of TV series and movies because of online video streaming services such as Netflix and Hulu. But TV isn't dead, because it works. WARC found that there's a greater proportion of TV-led ads in the 2013 to 2016 timeframe than in the prior three years among the 100 global campaigns that has deemed most effective (Wasserman 2018).

Other suggested that TV advertising is not dead, it is evolving. Studies from CMO and Wired suggest that TV advertising is evolving in the interactive age. Advertisers will produce TV advertising in the interactive form in the future just like it is now on social media platform such as YouTube. Audiences can play a game or customize the color on the items displayed in the ads. TV advertising will also become more targeted and selective as the connectivity and AI continue to develop.

A TV commercial has the potential to be watched by more than half of the population. From the survey conducted by Pew Research[3], among 2,002 U.S. adults age 18 & older who responded, around 59% of the population think it is hard to give up television, compared to the cell phone 76%, internet 76% and social network 39%. The numbers suggest that more than half of the population find it is hard to give up television. Although social media is booming, more advertisers returning to TV because they want to ensure maximum reach. So, do advertisers know who they are targeting to when doing TV adverting? What are the characteristics of the TV audience?

---

[2] Todd Wasserman. "TV Advertising Isn't Dead – It's Evolving." CMO. September 26, 2018. Accessed December 10, 2018. https://www.cmo.com/features/articles/2018/9/20/tv-advertising-isnt-deadits-evolving-.html#gs.F_Cz1aE

[3] Pew Research Center. (2018) "Jan. 3-10, 2018 – Core Trends Survey." http://www.pewinternet.org/dataset/jan-3-10-2018-core-trends-survey/

TV's Rivalry Relationship with Smartphone and Social Media

Many papers and studies suggested that smartphone and social media have negative impacts on TV advertising as they draw people's attention. A new study of media and attention by Nielsen Co. confirms that smartphones indeed drew people's attention from television[4].

The data underlying the report shows that TV viewing fell by 10% to 8.4 million people per minutes when there's a 25% increase in TV-connected devices such as streaming boxes or game consoles. The study also investigated the potential effect of the use of social media for TV viewing. Although there are some criticisms of Nielsen's video measurement on social media platform, it is the truth that the number of social media users is growing and it can cause the TV audience to decrease.

One assumption derived in this paper is that people who find it hard to give up on social network and smartphone will find it not hard to give up on TV. The following sections will examine if there's a rivalry relationship between TV and smartphone & social media.

Methodology

Data Collection

The data analyzed in this paper are collected by Pew Research from January 2nd to January 10th 2018[5]. The original dataset is in CSV format and has 70 columns and 2003 rows includes the header. The dependent variable (PIAL5a) in interest is people's attitude towards television. In essence, the question from the survey: "How difficult would it be, if at all, to give up your television in your life?" This variable is collected in the nominal form and has 8 levels

[4] Mathew Ingram. (2015) "The Smartphone Is Eating the Television, Nielsen Admits." http://fortune.com/2015/12/07/smartphone-tv-report/

[5] Pew Research Center. (2018) "Jan. 3-10, 2018 – Core Trends Survey." http://www.pewinternet.org/dataset/jan-3-10-2018-core-trends-survey/

including impossible, very hard, somewhat hard, not too hard, not hard at all, do not use / do not have, don't know and refused. Independent variables that are analyzed including demographic information such as age (age) and gender (sex), and survey questions such as if the respondent has a certain device such as a tablet computer (device1b), laptop computer (device1c) and game console (device1d), and what social media sites the respondent uses (web1). The model also explores people's attitude towards the internet (pial5c), cell phone (pial5b) and social media (pial5d). In addition, the number of books the respondents read during the past 12 months is also examined (books1).

Data Cleaning & Restructure

Since most of the variables from the original dataset are nominal, logistic regression and decision tree are used to further explore the correlation between people's attitude towards TV and other variables. As previously mentioned, the dependent variable (pial5a) has 8 levels. In order to interpret the results easier, the levels are reduced to 2. "Impossible", "very hard", and "somewhat hard" are recoded into "hard". "Not too hard" and "not hard at all" are recoded into "not hard". "Do not use / do not have" and "don't know and refused" from all variables are treated as missing values. The steps described above are not included in the preliminary analysis.

Tools

To serve the purpose of this paper, SAS, SPSS and MS Excel are utilized to analyze the data. SPSS is employed to do preliminary analysis such as descriptive analysis on the means and frequency. SAS is used to further explore the correlations among the variables, build detailed models and perform model utility and validation checking. Since the data cleaning and

restructuring approach is relatively basic for this dataset, all the related data processes are done in MS Excel.

Analysis and Discussion of Results

Descriptive Statistics

Among the 2,002 respondents who responded to the question concerning people's attitude towards TV, 655 found very hard to give up TV, 500 found somewhat hard and 23 found it is impossible to give up TV. In general, 59% of the population found it hard to give up TV.

Table 1. Descriptive Statistics of Frequency

PIAL5a. How difficult would it be, if at all, to give up the following things in your life? If you do not use or have the item, just tell me. How hard would it be for you to give up... Your television ?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Very hard	655	32.7	32.7	32.7
	Somewhat hard	500	25.0	25.0	57.7
	Not too hard	331	16.5	16.5	74.2
	Not hard at all	431	21.5	21.5	95.8
	(VOL) Impossible	23	1.1	1.1	96.9
	(VOL) Do not use / Do not have	57	2.8	2.8	99.8
	(VOL) Don't know	4	.2	.2	100.0
	(VOL) Refused	1	.0	.0	100.0
	Total	2002	100.0	100.0	

Classification Tree

In order to further explore the correlations of people's attitude towards TV (pial5a), a classification tree is built with pial5a as the dependent variable. The initial model fitting is conducted using SAS and the full output can be found in the appendix. The final model is as follow:

$$Y(pial5a) = age + pial5b + party + books1$$

Age: Current age of the respondents

Pial5b: How difficult would it be to give up cell phone

Party: Do you consider yourself a Republican, Democrat, or Independent?

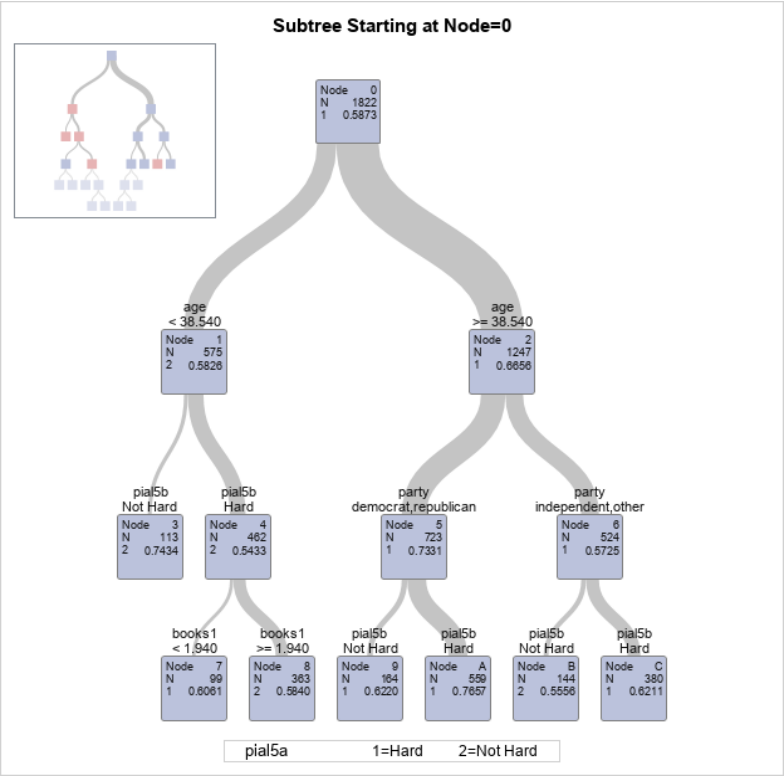
Books1: During the past 12 months, about how many books did you read either all or part of the way through?

The classification tree in Figure 1 suggests the following rules to determine if a respondent find it hard to give up TV:

1. If the respondent's age < 38.5, found it hard to give up cell phone, and read more than 2 books for the past 12 months, he/she is more likely to find it not hard to give up TV.
2. If the respondent's age < 38.5, found it hard to give up cell phone, and read less than 2 books for the past 12 months, he/she is more likely to find it hard to give up TV.
3. If the respondent's age >= 38.5, is a Democrat or Republic, he/she is more likely to find it hard to give up TV.
4. If the respondent's age >= 38.5, is an Independent or other and found it hard to give up cell phone, he/she is more likely to find it hard to give up TV.
5. If the respondent's age >= 38.5, is an Independent or other and found it not hard to give up cell phone, he/she is more likely to find it not hard to give up TV.

From the classification tree, respondents who are most likely to find it hard to give TV are those who are older than 38 years old, are Democrat or Republican and find it hard to give up the cell phone. The probability implied by this model is 77%.

Figure 1. Classification Tree from SAS Output.



The HPSPLIT Procedure

Model-Based Confusion Matrix

Actual	Predicted: Hard	Predicted: Not Hard	Error Rate
Hard	830	240	0.2243
Not Hard	344	408	0.4574

Model-Based Fit Statistics for Selected Tree

N Leaves	ASE	Mis-class	Entropy	Gini	RSS
12	0.2119	0.3205	0.8853	0.4238	772.1

Figure 2. Confusion Matrix from SAS Output

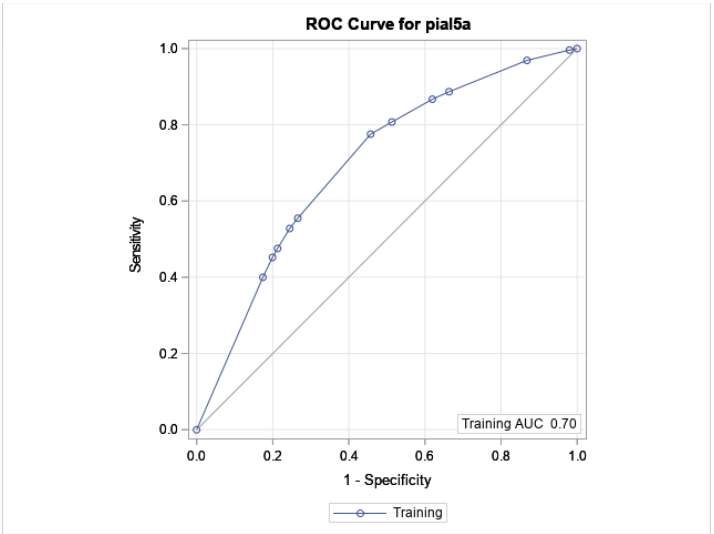


Figure 3. ROC Curve for Classification Tree

The confusion matrix from Figure 2 shows that the model has a specificity of 78% and a sensitivity of 54%. The overall misclassification rate of this model is 32%, which means that 68% of the observations are accurately classified using this model. Since we care more about the characteristics of respondents who find it hard to give up TV, we select the model with the highest specificity. The fitted classification tree model demonstrates a modest predictive performance as indicated by the area under the receiver operating characteristic curve of 0.70 reported in Figure 3. From this model we find that age, whether the respondent finds it hard to give up the cell phone or not, party and number of books read are important predictors of whether the respondent finds it hard to give up TV or not.

Since there are certain limitations to the model derived by classification tree. The data will be further analyzed through logistic regression. The limitations will be discussed in detail in the next section.

Logistics Regression

To systematically study the relationship between whether the respondents find it hard to give up TV and the predictors provided in the survey, a logistics regression analysis is conducted to predict if the respondents find it hard to give up TV or not. Different variables are introduced in this analysis with to a technique called backward selection. The initial model fitting is conducted using SAS and the full output can be found in the appendix. The final model is as following:

$$Y(pial5a) = age + educ2 + marital + pial5b + pial5d + books1$$

Age: Current age of the respondents

Educ2: The highest level of school the respondents have completed or the highest degree they have received

Marital: Marital status of the respondents

Pial5b: How difficult would it be to give up cell phone?

Pial5d: How difficult would it be to give up social media?

Books1: During the past 12 months, about how many books did you read either all or part of the way through?

Table 2. Analysis of Maximum Likelihood Estimates

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.1308	0.2629	18.5032	<.0001
age		1	0.0339	0.00416	66.3448	<.0001
educ2	1	1	-0.0446	0.4245	0.0110	0.9163
educ2	2	1	0.0152	0.2729	0.0031	0.9557
educ2	3	1	0.2859	0.1485	3.7077	0.0542
educ2	4	1	0.0127	0.1589	0.0063	0.9365
educ2	5	1	-0.0558	0.1785	0.0977	0.7546
educ2	6	1	0.3282	0.1350	5.9090	0.0151
educ2	7	1	-0.4048	0.1922	4.4358	0.0352
marital		1	-0.0795	0.0317	6.2852	0.0122
pial5b	Hard	1	0.3723	0.0847	19.3451	<.0001
pial5d	Hard	1	0.3217	0.0663	23.5127	<.0001
books1		1	-0.00872	0.00299	8.5036	0.0035

The fitted logistic model demonstrates a relatively good predictive performance as indicated by the area under the receiver operating characteristic curve of 0.7113 reported in Figure 4. As Table 2 has shown above, most of the selected predictors are statistically significant at the 5% significance level. However, there are some categories in the predictor (educ2) are not statistically significant. From which, only category 6 (four-year college or university

degree/Bachelor's degree) and category 7 (some postgraduate or professional schooling) are found to be statistically significant. Number of books read, marital status and whether the respondents went to graduate school are negatively related to whether the respondents find it hard to give up TV. Table 3 shows that with a null hypothesis that the model generally fits well, the test fails to reject indicating that the model is a reasonable description of the relationship between the predictors and the response.

Though the model appears to provide some insights into respondent's attitude towards TV, it does not appear to have sufficient predictive ability to be utilized to make specific predictions about individual person. The highest correct classification rate that the model achieves is roughly 65.1%.

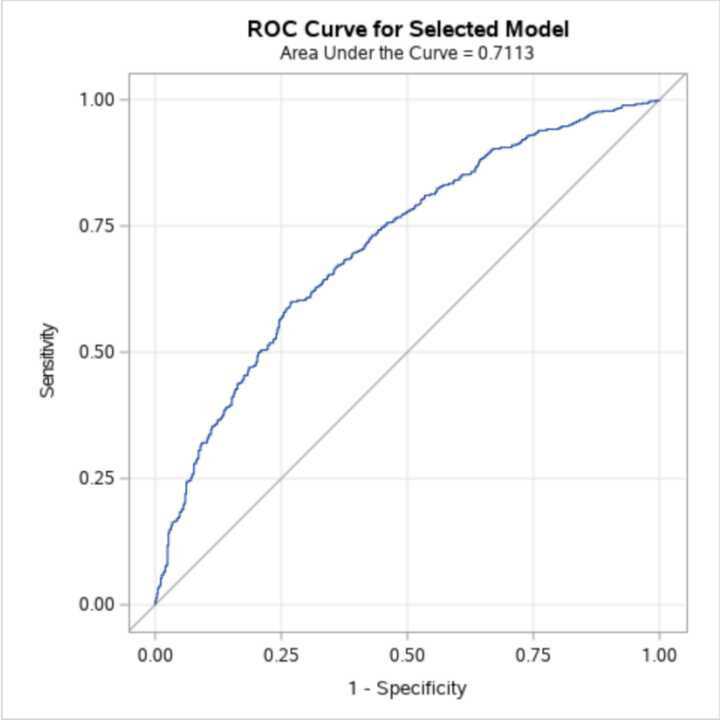


Figure 4. ROC Curve for Final Model.

Table 3. Hosmer and Lemeshow Test

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
4.3957	8	0.8198

Based on the survey data collect by Pew Research, there is certainly evidence that whether people find it difficult to give TV is related to people's age, education level, marital

status, whether they find it difficult to give up cell phone, whether they find it hard to give up social network and the number of books they read during the past 12 months.

Limitations and Directions for Future Research

There are certain limitations for both of the models. Classification tree analysis is very subjective since there is no statistical basis or assumptions. The tree generated in this report is not the "perfect" model for classifying the characteristics of audiences for TV as there is no clear answer what "perfect" classification tree model is. In addition, the sample size of the dataset is very small. Thus, it might not generate a model that can be generalized to a larger sample size. If there are a lot of outliers in this dataset, the classification can be very inaccurate.

As for logistic regression, although the HL test suggests that the model generally fits well, the Pearson Chi-Square Residual vs Index Plot show some patterns indicating there might be time series structure in the dataset. Thus, additional information needs to be introduced into this model in order to use the model to predict external datasets.

Additional information can be gathered based on these initial findings to determine other factors may play a role in whether a person finds it is hard to give up TV. In addition, a possible challenge for this research is that people's behavior changes over time. New technologies such as virtual reality can have an impact on people's attitude towards the TV. The TV can also evolve and change form.

For future studies, surveys like Pew Research conducted can ask more questions related to newer technologies and more specific questions related to TV advertising. Researchers can increase the sample size of the respondents or eliminate the number of null responses as many as possible. The survey can even be extended to another country to understand the global effect.

References

Benjamin Schroeder and Nathan Vierling-Claassen. "Optimizing TV Advertising Toward Return on Investment." Wayfair. August 22, 2018. Accessed December 10, 2018. https://tech.wayfair.com/2018/08/optimizing-tv-advertising-toward-return-on-investment/

Mathew Ingram. (2015) "The Smartphone Is Eating the Television, Nielsen Admits." http://fortune.com/2015/12/07/smartphone-tv-report/

Todd Wasserman. "TV Advertising Isn't Dead – It's Evolving." CMO. September 26, 2018. Accessed December 10, 2018. https://www.cmo.com/features/articles/2018/9/20/tv-advertising-isnt-deadits-evolving-.html#gs.F_Cz1aE

Pew Research Center. (2018) "Jan. 3-10, 2018 – Core Trends Survey." http://www.pewinternet.org/dataset/jan-3-10-2018-core-trends-survey/

Appendix

Descriptive Statistics

PIAL5a. How difficult would it be, if at all, to give up the following things in your life? If you do not use or have the item, just tell me. How hard would it be for you to give up... Your television ?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Very hard	655	32.7	32.7	32.7
	Somewhat hard	500	25.0	25.0	57.7
	Not too hard	331	16.5	16.5	74.2
	Not hard at all	431	21.5	21.5	95.8
	(VOL) Impossible	23	1.1	1.1	96.9
	(VOL) Do not use / Do not have	57	2.8	2.8	99.8
	(VOL) Don't know	4	.2	.2	100.0
	(VOL) Refused	1	.0	.0	100.0
	Total	2002	100.0	100.0	

PIAL5b. How difficult would it be, if at all, to give up the following things in your life? If you do not use or have the item, just tell me. How hard would it be for you to give up... Your cell phone or smartphone ?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Very hard	982	49.1	50.8	50.8
	Somewhat hard	433	21.6	22.4	73.2
	Not too hard	210	10.5	10.9	84.1
	Not hard at all	232	11.6	12.0	96.1
	(VOL) Impossible	53	2.6	2.7	98.8
	(VOL) Do not use / Do not have	12	.6	.6	99.4
	(VOL) Don't know	8	.4	.4	99.8
	(VOL) Refused	3	.1	.2	100.0
	Total	1933	96.6	100.0	
Missing	System	69	3.4		
Total		2002	100.0		

PIAL5c. How difficult would it be, if at all, to give up the following things in your life? If you do not use or have the item, just tell me. How hard would it be for you to give up... The internet ?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Very hard	897	44.8	50.3	50.3
	Somewhat hard	417	20.8	23.4	73.6
	Not too hard	196	9.8	11.0	84.6
	Not hard at all	212	10.6	11.9	96.5
	(VOL) Impossible	48	2.4	2.7	99.2
	(VOL) Do not use / Do not have	12	.6	.7	99.8
	(VOL) Don't know	3	.1	.2	100.0
	Total	1785	89.2	100.0	
Missing	System	217	10.8		
Total		2002	100.0		

PIAL5d. How difficult would it be, if at all, to give up the following things in your life? If you do not use or have the item, just tell me. How hard would it be for you to give up... Social media ?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Very hard	179	8.9	13.3	13.3
	Somewhat hard	341	17.0	25.4	38.7
	Not too hard	404	20.2	30.1	68.8
	Not hard at all	413	20.6	30.8	99.6
	(VOL) Impossible	2	.1	.1	99.7
	(VOL) Do not use / Do not have	2	.1	.1	99.9
	(VOL) Don't know	2	.1	.1	100.0
	Total	1343	67.1	100.0	
Missing	System	659	32.9		
Total		2002	100.0		

SAS Code

```
/*Classification Tree*/
ods graphics on;
proc hpsplit data=trend plots=all;
class pial5a pial5b party educ2;
model pial5a(event='Hard')=age pial5b party educ2 books1;
run;
ods graphics off;

/*Logistic Regression*/
ods graphics on;
proc logistic data=trend plots(unpack)=(roc);
class pial5a pial5b pial5d party educ2 device1d;
model pial5a(event='Hard')=age party educ2 marital pial5b pial5d books1 device1d hh1
/ lackfit ctable selection=backward;
run;
ods graphics off;
```

SAS Output

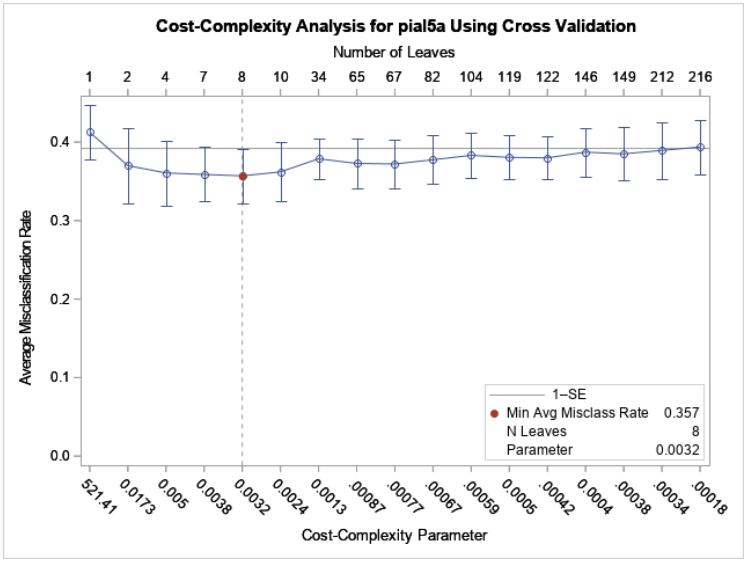
The HPSPLIT Procedure

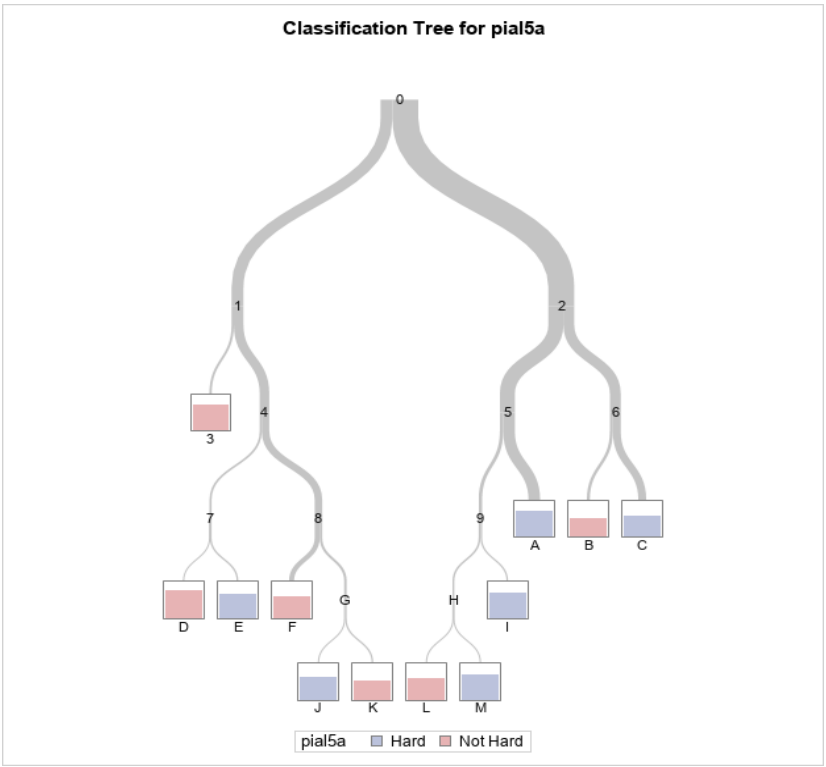
Performance Information	
Execution Mode	Single-Machine
Number of Threads	4

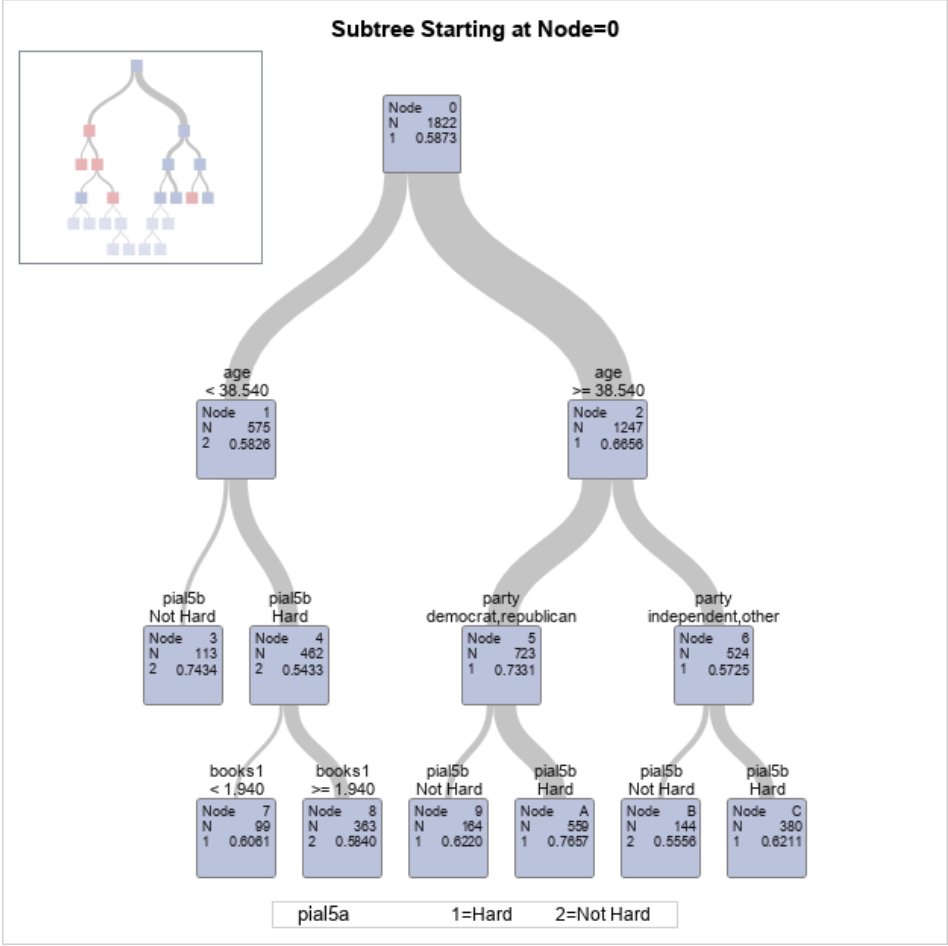
Data Access Information			
Data	Engine	Role	Path
WORK.TREND4	V9	Input	On Client

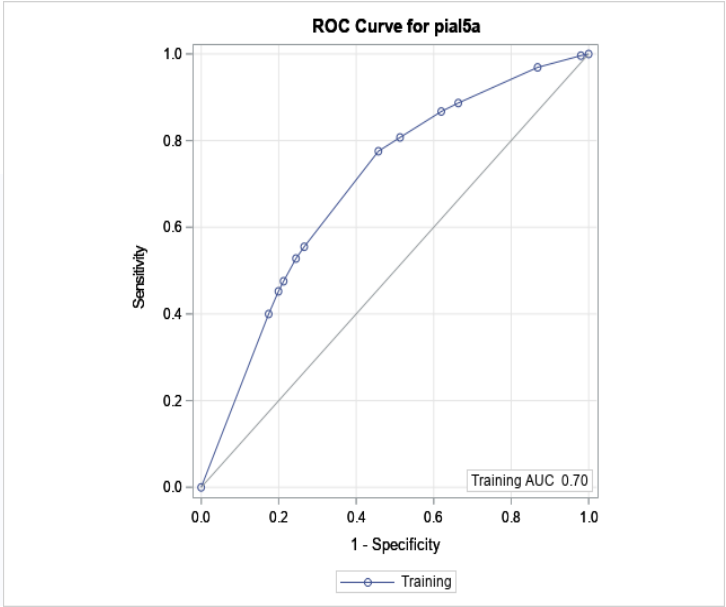
Model Information	
Split Criterion Used	Entropy
Pruning Method	Cost-Complexity
Subtree Evaluation Criterion	Cost-Complexity
Number of Branches	2
Maximum Tree Depth Requested	10
Maximum Tree Depth Achieved	10
Tree Depth	5
Number of Leaves Before Pruning	253
Number of Leaves After Pruning	12
Model Event Level	Hard

Number of Observations Read	2007
Number of Observations Used	1822









The HPSPLIT Procedure

Model-Based Confusion Matrix

Actual	Predicted: Hard	Predicted: Not Hard	Error Rate
Hard	830	240	0.2243
Not Hard	344	408	0.4574

Model-Based Fit Statistics for Selected Tree

N Leaves	ASE	Mis-class	Entropy	Gini	RSS
12	0.2119	0.3205	0.8853	0.4238	772.1

Variable Importance			
	Training		
Variable	Relative	Importance	Count
age	1.0000	7.7681	3
pial5b	0.5614	4.3612	3
party	0.5094	3.9572	1
educ2	0.3704	2.8770	2
books1	0.3599	2.7958	2

The LOGISTIC Procedure

Model Information	
Data Set	WORK.TREND
Response Variable	pial5a
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	2002
Number of Observations Used	1256

Response Profile		
Ordered Value	pial5a	Total Frequency
1	Hard	723
2	Not Hard	533

Probability modeled is pial5a='Hard'.

Note: 746 observations were deleted due to missing values for the response or explanatory variables.

Backward Elimination Procedure

Class Level Information								
Class	Value	Design Variables						
pial5b	Hard	1						
	Not Hard	-1						
pial5d	Hard	1						
	Not	-1						
party	democrat	1	0	0				
	independent	0	1	0				
	other	0	0	1				
	republican	-1	-1	-1				
educ2	1	1	0	0	0	0	0	0
	2	0	1	0	0	0	0	0
	3	0	0	1	0	0	0	0
	4	0	0	0	1	0	0	0
	5	0	0	0	0	1	0	0
	6	0	0	0	0	0	1	0
	7	0	0	0	0	0	0	1
	8	-1	-1	-1	-1	-1	-1	-1
device1d	N	1						
	Y	-1						

Step 0. The following effects were entered:

Intercept age party educ2 marital pial5b pial5d books1 device1d hh1

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1714.333	1563.451
SC	1719.469	1655.893
-2 Log L	1712.333	1527.451

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	184.8823	17	<.0001
Score	174.2844	17	<.0001
Wald	152.4744	17	<.0001

Step 1. Effect hh1 is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1714.333	1563.175
SC	1719.469	1650.481
-2 Log L	1712.333	1529.175

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	183.1584	16	<.0001
Score	172.5614	16	<.0001
Wald	151.0488	16	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
1.7231	1	0.1893

Step 2. Effect device1d is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1714.333	1564.614
SC	1719.469	1646.785
-2 Log L	1712.333	1532.614

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	179.7194	15	<.0001
Score	169.4590	15	<.0001
Wald	148.5559	15	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
5.1339	2	0.0768

Step 3. Effect party is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1714.333	1565.965
SC	1719.469	1632.728
-2 Log L	1712.333	1539.965

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	172.3685	12	<.0001
Score	162.9107	12	<.0001
Wald	143.6644	12	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
12.4851	5	0.0287

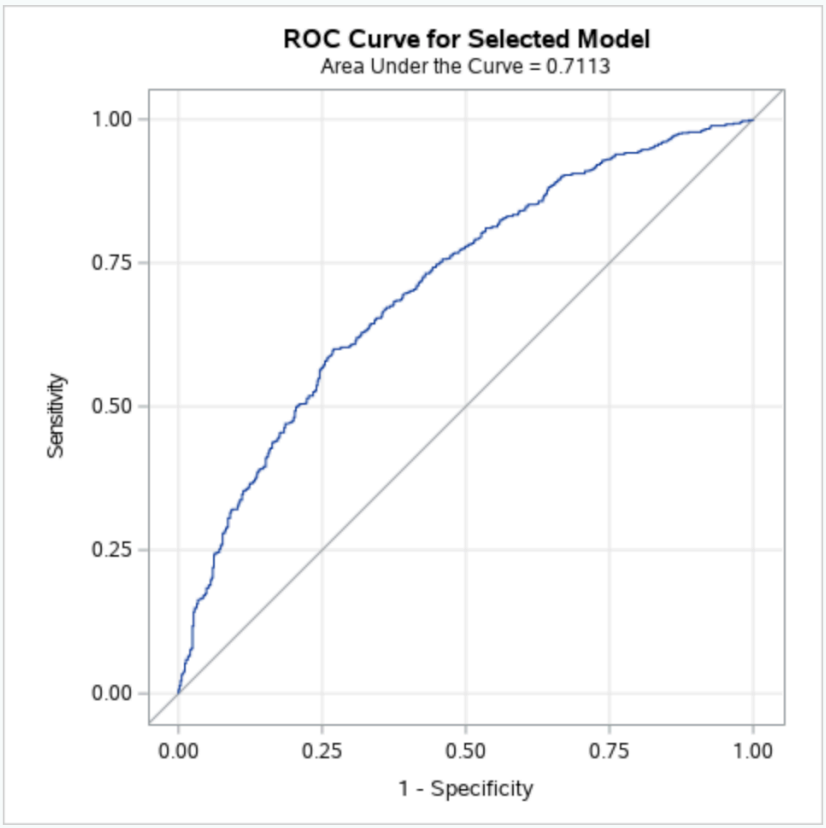
Summary of Backward Elimination					
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq
1	hh1	1	8	1.7190	0.1898
2	device1d	1	7	3.4127	0.0647
3	party	3	6	7.3350	0.0620

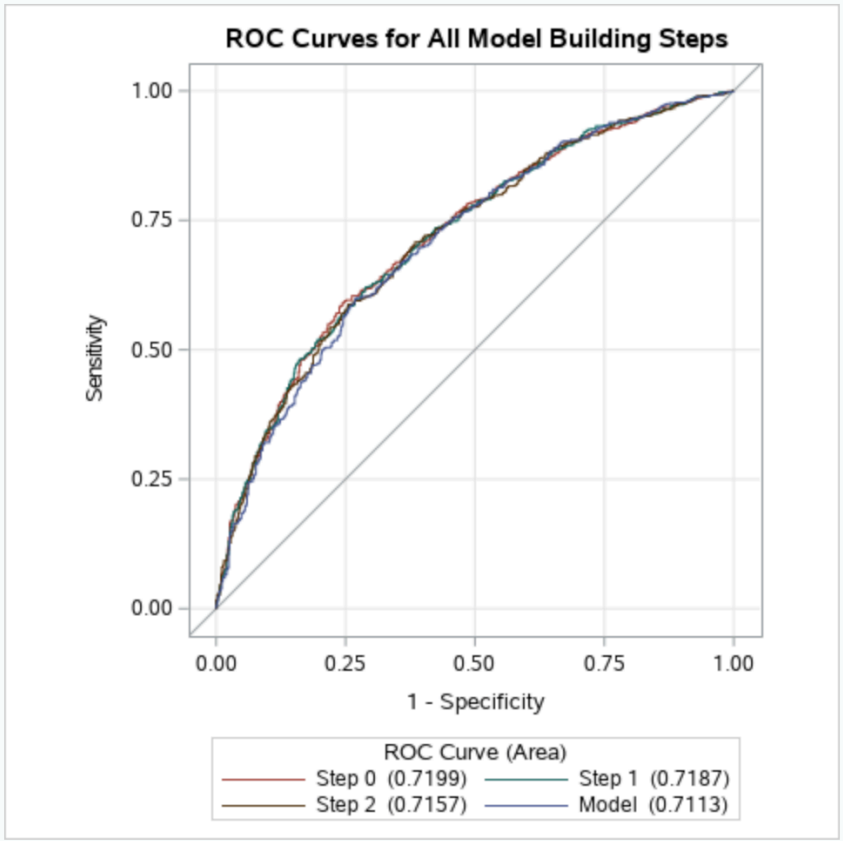
Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
age	1	66.3448	<.0001
educ2	7	14.3768	0.0449
marital	1	6.2852	0.0122
pial5b	1	19.3451	<.0001
pial5d	1	23.5127	<.0001
books1	1	8.5036	0.0035

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.1308	0.2629	18.5032	<.0001
age		1	0.0339	0.00416	66.3448	<.0001
educ2	1	1	-0.0446	0.4245	0.0110	0.9163
educ2	2	1	0.0152	0.2729	0.0031	0.9557
educ2	3	1	0.2859	0.1485	3.7077	0.0542
educ2	4	1	0.0127	0.1589	0.0063	0.9365
educ2	5	1	-0.0558	0.1785	0.0977	0.7546
educ2	6	1	0.3282	0.1350	5.9090	0.0151
educ2	7	1	-0.4048	0.1922	4.4358	0.0352
marital		1	-0.0795	0.0317	6.2852	0.0122
pial5b	Hard	1	0.3723	0.0847	19.3451	<.0001
pial5d	Hard	1	0.3217	0.0663	23.5127	<.0001
books1		1	-0.00872	0.00299	8.5036	0.0035

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
age	1.034	1.026	1.043
educ2 1 vs 8	1.096	0.394	3.053
educ2 2 vs 8	1.164	0.565	2.398
educ2 3 vs 8	1.526	0.924	2.520
educ2 4 vs 8	1.161	0.694	1.943
educ2 5 vs 8	1.084	0.628	1.871
educ2 6 vs 8	1.592	0.993	2.550
educ2 7 vs 8	0.765	0.438	1.336
marital	0.924	0.868	0.983
pial5b Hard vs Not Hard	2.106	1.511	2.934
pial5d Hard vs Not	1.903	1.467	2.468
books1	0.991	0.986	0.997

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	71.1	Somers' D	0.423
Percent Discordant	28.9	Gamma	0.423
Percent Tied	0.0	Tau-a	0.207
Pairs	385359	c	0.711

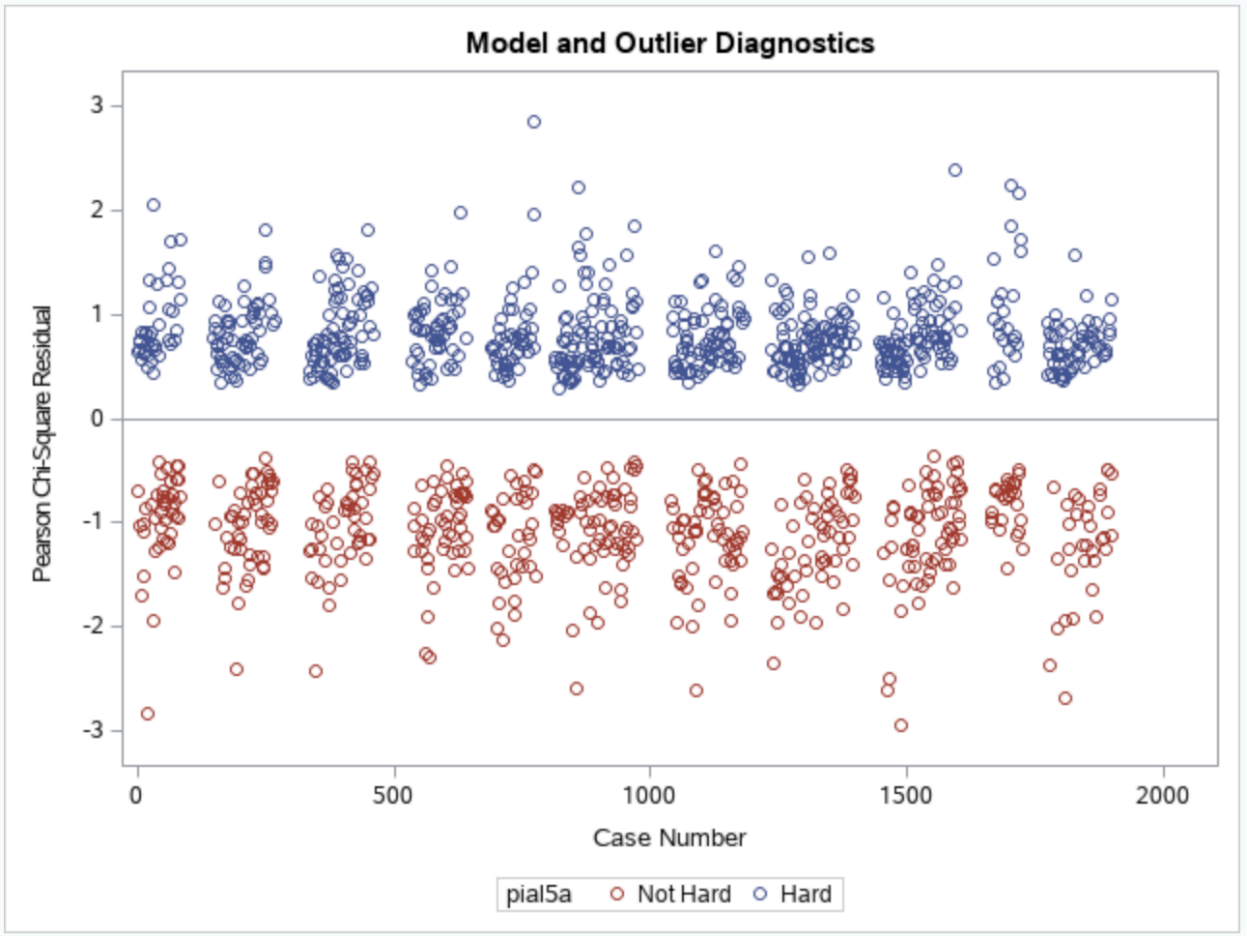

ROC Curve for Selected Model
Area Under the Curve = 0.7113
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00


ROC Curves for All Model Building Steps
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00
ROC Curve (Area)
Step 0 (0.7199)
Step 1 (0.7187)
Step 2 (0.7157)
Model (0.7113)

Partition for the Hosmer and Lemeshow Test					
		pial5a = Hard		pial5a = Not Hard	
Group	Total	Observed	Expected	Observed	Expected
1	126	34	30.76	92	95.24
2	126	40	46.38	86	79.62
3	126	62	56.53	64	69.47
4	126	62	64.70	64	61.30
5	126	70	72.11	56	53.89
6	126	80	78.10	46	47.90
7	126	82	83.71	44	42.29
8	126	88	89.99	38	36.01
9	126	98	97.30	28	28.70
10	122	107	103.41	15	18.59

Hosmer and Lemeshow Goodness-of-Fit Test		
Chi-Square	DF	Pr > ChiSq
4.3957	8	0.8198

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.080	723	0	533	0	57.6	100.0	0.0	42.4	.
0.100	722	0	533	1	57.5	99.9	0.0	42.5	100.0
0.120	722	1	532	1	57.6	99.9	0.2	42.4	50.0
0.140	721	2	531	2	57.6	99.7	0.4	42.4	50.0
0.160	719	8	525	4	57.9	99.4	1.5	42.2	33.3
0.180	718	13	520	5	58.2	99.3	2.4	42.0	27.8
0.200	716	21	512	7	58.7	99.0	3.9	41.7	25.0
0.220	712	32	501	11	59.2	98.5	6.0	41.3	25.6
0.240	709	41	492	14	59.7	98.1	7.7	41.0	25.5
0.260	706	49	484	17	60.1	97.6	9.2	40.7	25.8
0.280	701	65	468	22	61.0	97.0	12.2	40.0	25.3
0.300	695	77	456	28	61.5	96.1	14.4	39.6	26.7
0.320	687	86	447	36	61.5	95.0	16.1	39.4	29.5
0.340	679	104	429	44	62.3	93.9	19.5	38.7	29.7
0.360	668	123	410	55	63.0	92.4	23.1	38.0	30.9
0.380	657	144	389	66	63.8	90.9	27.0	37.2	31.4
0.400	649	161	372	74	64.5	89.8	30.2	36.4	31.5
0.420	638	180	353	85	65.1	88.2	33.8	35.6	32.1
0.440	612	192	341	111	64.0	84.6	36.0	35.8	36.6
0.460	602	212	321	121	64.8	83.3	39.8	34.8	36.3
0.480	583	233	300	140	65.0	80.6	43.7	34.0	37.5
0.500	561	255	278	162	65.0	77.6	47.8	33.1	38.8
0.520	546	272	261	177	65.1	75.5	51.0	32.3	39.4
0.540	524	294	239	199	65.1	72.5	55.2	31.3	40.4
0.560	502	311	222	221	64.7	69.4	58.3	30.7	41.5
0.580	473	334	199	250	64.3	65.4	62.7	29.6	42.8
0.600	448	355	178	275	63.9	62.0	66.6	28.4	43.7
0.620	411	381	152	312	63.1	56.8	71.5	27.0	45.0
0.640	368	403	130	355	61.4	50.9	75.6	26.1	46.8
0.660	340	423	110	383	60.7	47.0	79.4	24.4	47.5
0.680	303	438	95	420	59.0	41.9	82.2	23.9	49.0
0.700	270	455	78	453	57.7	37.3	85.4	22.4	49.9
0.720	236	473	60	487	56.4	32.6	88.7	20.3	50.7
0.740	204	488	45	519	55.1	28.2	91.6	18.1	51.5


Model and Outlier Diagnostics
Pearson Chi-Square Residual
3
2
1
0
-1
-2
-3
0
500
1000
1500
2000
Case Number
pial5a
Not Hard
Hard