Law and Norms:

A Machine Learning Approach to Predicting Attitudes Towards Abortion

Daniel L. Chen, Kristen Kwan, Luisa E. Quispe Ortiz, Maria Leonor Zamora Maass[1]

July 2016

Abstract

Understanding the predictors of societal attitudes has been widely investigated using individuals' responses to surveys and polls. In this paper, we use U.S. Courts of Appeals cases, sociological and attitudinal indicators, criminal statistics, and a variety of survey data to predict societal attitudes towards abortion. We create two classification models: pro-abortion attitudes for health reasons related to the mother, fetus, or rape and pro-abortion attitudes for any other personal non-health related reason. To address high dimensionality, we employ factor analysis to group indicators. Logistic regression and random forests performed best among three types of classifiers evaluated on AUC and accuracy.

For pro abortion attitudes related to health related abortions, the most important factor contains sociological and attitudinal indicators about the frequency of contact with family and friends. The most important U.S. Courts of Appeals indicators include the religion, political affiliation, and ABA ratings of the judges in the Circuit pool. The most important crime indicators are the rates of violent and property crimes.

For pro abortion attitudes towards non-health related abortions, the most important factor contains sociological and attitudinal indicators about job satisfaction, religious preference, religion raised in, and beliefs about the Bible. The most important U.S. Courts of Appeals indicators include the religion and political affiliation of the judges in the Circuit panel. These legal indicators are grouped in the same factor as the most important crime indicators, which are rates of crimes against society, crimes against property, and violent crimes.

---

[1] daniel.li.chen@gmail.com: Toulouse Institute for Advanced Study, kristen.kwan@nyu.edu, lqo202@nyu.edu, mzm239@nyu.edu: NYU Courant Institute, Center for Data Science

I. Introduction

U.S. Courts of Appeals often make rulings in areas that are socially relevant to the American public, such as abortion, gender discrimination, or capital punishment. Whether these rulings affect Americans' social and political attitudes remains an open question.[2] Our goal is to create a prediction model of attitudes and include a random component of court rulings---the composition of the panel of judges assigned to the cases---among the feature set. The random composition renders a causal interpretation and the feature weights offer a natural metric to evaluate the importance of court rulings. We will focus on rulings about abortion. Two secondary questions are which judicial characteristics are most relevant in predicting societal attitudes and when, in a time scale of months, is their influence reflected.

We also explore the influence of crime rates to understand whether and which local criminal activities are predictive of attitudes towards abortion.

II. Datasets

First, we used the US General Social Survey (GSS), which is a long running (1972–) survey on social attitudes and behaviors of US American citizens (Smith et al. 2013). In addition to social attitudes, the GSS provides demographic and life-course data about respondents.

Second, we used a database of U.S. Courts of Appeals cases that were decided at the level of Circuit Courts (Sunstein 2007). The cases were separated by issue (e.g., abortion, gender discrimination). The federal court system is divided into 12 regional Circuits, each of which establishes legal precedent for a group of several states, as well as a Federal Circuit Court, which has nationwide jurisdiction over a variety of Federal matters. A Circuit Court case is decided by a randomly assigned panel of three judges chosen from the pool of judges appointed to that circuit. The dataset provides information about the outcome of each case, which is coded as the number of judges who voted in favor the outcome that can be considered to be more "progressive" (for example, pro-abortion, against gender discrimination, etc.).

Another dataset was used to assign judge characteristics to each case (Zuk et al. 1997). This included, for instance, the percent of panel judges who were female, or who were appointed by a Democratic president. It also included the average percent of female judges (and other characteristics) in the pool of judges for that Circuit at the time of the ruling.

[2] See Chen, Levonyan, Yeh (2015) and Chen and Yeh (2015) for a summary of the literature on law and societal attitudes and for details about the U.S. Federal Courts.

A fourth dataset we used is labor market outcomes from the Merged Outgoing Rotation Groups Current Population Survey (CPS). This data is collected by the Bureau of Labor Statistics jointly with the Census Bureau and contains monthly individual employment outcomes, including weekly earnings, amount of time worked, employment status, and management status to understand the national employment situation. The CPS has been conducted since 1962 and currently interviews ~54,000 households monthly to represent the nation as a whole.

Our final dataset is annual data on crime incidents from the FBI's Uniform Crime Reports (UCR). The UCR data contains nationwide monthly and annual crime data by county collected by 17,000 law enforcement agencies from the years 1977 to 2007.

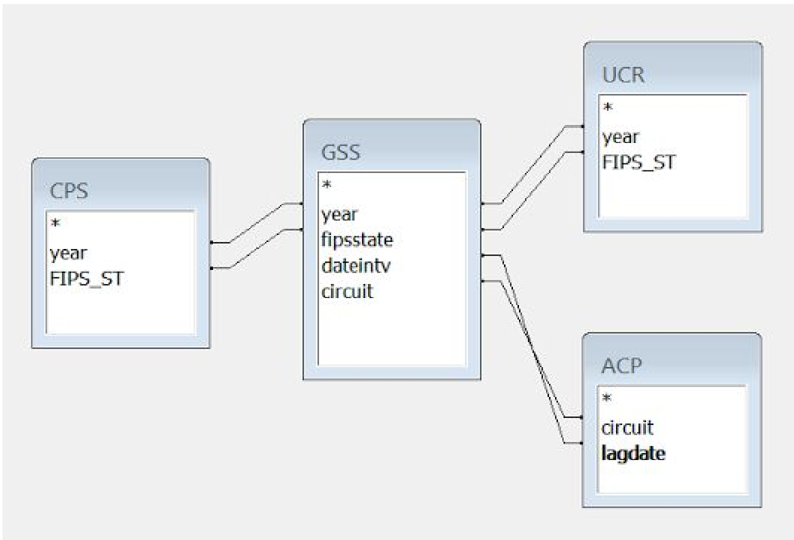
The five different datasets are summarized in Table 1:

Table 1. Summary of datasets.

Dataset		Description	Source
General Social Survey (GSS)		U.S. sociological and attitudinal trends	NORC at the University of Chicago
Abortion Circuit Panel (ACP)	Circuit Court Outcomes	Number of judges who voted in favor for a particular Circuit-level court case outcome	Chicago Judges Project
	Judicial Panel Characteristics	Characteristics of judicial panels for Circuit-level U.S. court rulings related to abortion	Chen, Levonyan, Yeh (2015)
Current Population Survey (CPS)		Monthly labor force and demographics	U.S. Census Bureau, Bureau of Labor Statistics
Uniform Crime Report (UCR)		Nationwide crime statistics	Federal Bureau of Investigation

We engineered features from these datasets and merged them as shown in Figure 1. As the earliest UCR data was from 1977 and the latest GSS data with state identifiers (Federal Information Processing Standard (FIPS) state code) was from 2004, the final dataset represents the time period of 1977-2004. We detail the dataset processing and feature engineering below.

Figure 1. Schema of dataset merges for final dataset used in modeling.



The primary dataset was the GSS, which is a yearly open survey that captures sociological, demographic, behavioral and attitudinal trends. The data spans the years 1972 to 2012 and consists of 57,061 observations and 5,548 features. Each observation represents one response to the survey. Respondents of this survey are adults (18 years and older) living in households in the United States. Respondents must be able to complete the survey in English or Spanish, and cannot be residents of institutions or group quarters. We dropped features with 95% or more missing data, which left 1,116 features. We dropped observations outside 1977-2004, leaving 26,533 observations. For categorical features, we constructed separate features for each category type.

The GSS contains 22 features related to abortion attitudes. However, the majority of these contain more than 93% missing values. After excluding these, six features remained. We used these six features to construct a binary target, where 1 represents a pro-abortion attitude and 0 represents an anti-abortion attitude.

Since abortion attitudes likely differ based on the reason for the abortion, we constructed two distinct targets. For the first target, we grouped the three features related to abortion for health related reasons. For the second target, we grouped the three features related to abortion for non-health reasons. For both targets, we assigned a value of ‘1’ if two out of the three questions were answered as ‘yes’. Blanks, “don’t know” (D.K.), and “NA” responses were excluded. We present summary statistics in Tables 2 and 3.

Table 2. Health related abortion features used to construct Target 1.

VARIABLE	abdefect	abhlth	abrape	Target1
BLANKS / D.K.	31%	31%	32%	31%
NO	14%	7%	12%	11%
YES	55%	62%	56%	58%
In favor of abortion when:	There is strong chance of serious fetal defect	Woman's health seriously endangered	Woman pregnant as result of rape	

CUTOFF	1 out of 3	2 out of 3	3 out of 3
BLANKS / D.K.	29%	29%	29%
NO	6%	11%	18%
YES	64%	60%	52%

Table 3. Non-health related abortion features used to construct Target 2.

VARIABLE	abnomore	abpoor	absingle	Target2
BLANKS / D.K.	31%	32%	32%	32%
NO	38%	36%	38%	37%
YES	30%	32%	30%	31%
In favor of abortion when:	A married couple doesn't want more children	The family has low income and can't afford more children	The woman is not married	

CUTOFF	1 out of 3	2 out of 3	3 out of 3
BLANKS / D.K.	30%	30%	30%
NO	33%	39%	44%
YES	38%	31%	26%

The ACP dataset consists of 142 observations and 218 features. Each observation represents one abortion related court ruling. Each feature provides a characteristic of the court ruling, with 105 of these features relating to the composition of the judge panel and 105 of these features relate to the composition of the judge pool based on all of the judges in the Circuit’s judge pool. The judge pool composition is used to understand the expected judge panel composition. Composition components captured in these features include the race, gender, education, political affiliation, and religion of the judges. The remaining features provide the circuit number, case ruling date, and the number of liberal votes for the case. The dataset contains no missing values. Case ID and case name features were dropped since there was no meaningful way to include them in the model. The case category feature was also dropped since the value was the same for all observations. Observations that occurred during the same month, year, and circuit were averaged together to form one observation. This process resulted in 137 observations and 215 features.

A 'lagdate' feature was added to capture the lag time (by month up to one year) between an abortion case ruling and the interview date of a GSS respondent. For example, if an abortion ruling was made by a circuit panel in July 1997, 12 'lag dates' were created from this ruling date, ranging from August 1997 to July 1998. These lag dates were later used as a merge point between the ACP and GSS datasets, as represented in Figure 1. This process increased the number of observations to 1,644.

To indicate the lag time represented by the 'lagdate', we then constructed 12 additional binary features (lagMonth1, lagMonth2,.…, lagMonth12), each representing a lag time of 1-12 months between an abortion case ruling and a GSS respondent's survey date. Observations with the same 'lagdate' and circuit were then merged into one observation, with the judge panel composition values averaged together and the 12 additional binary features added together. For example, if a GSS respondent was interviewed in January 1975 and one ruling had occurred two months earlier (November 1974) and another ruling had occurred 12 months earlier (January 1974), the ACP features constructed for this GSS respondent involved taking the average of ACP features that may have already been averaged as discussed above. Even though we are taking the average of averages, we preserve the randomization of panel assignments by keeping both the actual and expected composition of the judicial panels as features. Thus, an observation in the final dataset contains case data that represents the average of abortion case ruling characteristics that occurred between 1-12 months before a GSS respondent's survey date. This process resulted in a final dataset of 1,410 observations and 226 features. In other words, this dataset represents case data features for GSS respondents who were surveyed across 1,410 months and circuit regions. Details of this process are further explained in Appendix A.

The CPS data contains ~6.5 million observations. Each observation represents one household member and corresponds to 316 features, including age, sex, race, household, employment status, work experience, and income. After dropping features with more than 95% missing values, 254 features remained. We dropped observations outside of the 1977 - 2004 time frame, resulting in 4.5 million observations. Then, summarization of features was made considering groups formed by year and state. We took group means of continuous-valued features over state and year to merge with the GSS. For categorical features, we constructed separate features for each possible category type, also considering the weight. The value in each of these separate features represents the percentage of respondents who identify with this category type. For example, the feature for marital status contains four possible response types: married(1), single(2), divorced(3), and missing (NA), so we would generate four separate binary features to represent the possible marital status response types, where all four fields for one observation sum to 1. Later, we calculate percentage of respondents that marked married (1), single (2), divorced (3), and missing (NA) by year and state.

After this process, some features still contained missing values, so we added binary flags for these features indicating whether the feature had a missing value for that observation. We also replaced the actual missing values with a known constant, where zero was chosen if 0% of the respondents selected that option. This final dataset consists of 693 features and 1,428 observations.

The UCR data contains ~94,000 observations corresponding to 61 features related to location details and crime categories. Most observations represent the crime statistics reported for one month for a particular year, county, and state. We dropped features related to Inter-University Consortium for Political and Social Research (ICPSR) since there was no meaningful way to include them in the model. We also dropped observations outside of the 1977-2004 timeframe. Since the UCR dataset was merged to the GSS dataset by year and state, we summed the crime statistics together by year and state and dropped features related to counties. Crime statistics were further aggregated into crimes against persons, crimes against property, and crimes against society, resulting in 1,378 observations and 7 features.

After all five of these datasets were merged into one dataset (as depicted in Figure 1), we added binary flags for the features to indicate whether the feature had a missing value for that observation. We also replaced the actual missing values with a known constant. We normalized all features to have mean 0 and standard deviation 1.

We randomized the merged dataset and then split it into 80% training and 20% test datasets. For reproducibility, we set the random seed to 1. Table 4 summarizes the final datasets:

Table 4. Summary of final merged datasets.

	Final Merged Dataset for Health Related Abortions (Target 1)	**Final Merged Dataset for Non-Health Related Abortions (Target 2)**
# observations	8,721	8,664
# features	3,702	3,704
% pro abortion	84.1%	41.3%

III. Methodology

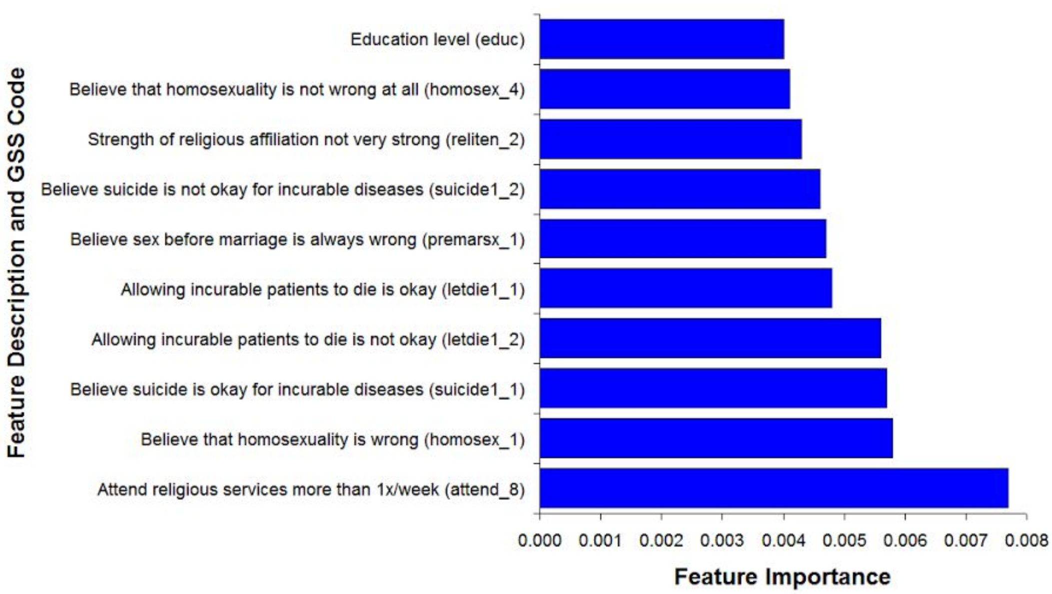
We first constructed a baseline model. To address the high-dimensionality of the datasets, we then performed a factor analysis. Factor analysis proceeded in four steps: a topic analysis, construction of feature dictionaries, factor analysis, and the selection of the optimal number of factors. In the final step, we tried three different classification models and evaluated these models using Area Under the ROC Curve (AUC) and Error (1 - Accuracy) as performance metrics. The model with the best performance was chosen as the final model.

Baseline Model

As a first step, we ran out-of-box (OOB) random forest models on the final datasets to gain some insight and handle on our datasets. Unlike the tuned random forest models used later in our paper, the number of trees used in these random forest models remained at the default value of 10.

From the OOB random forest results, we evaluated features based on their feature weights. The top ten most important features for health related abortion attitudes (model 1) are summarized in Figure 2.

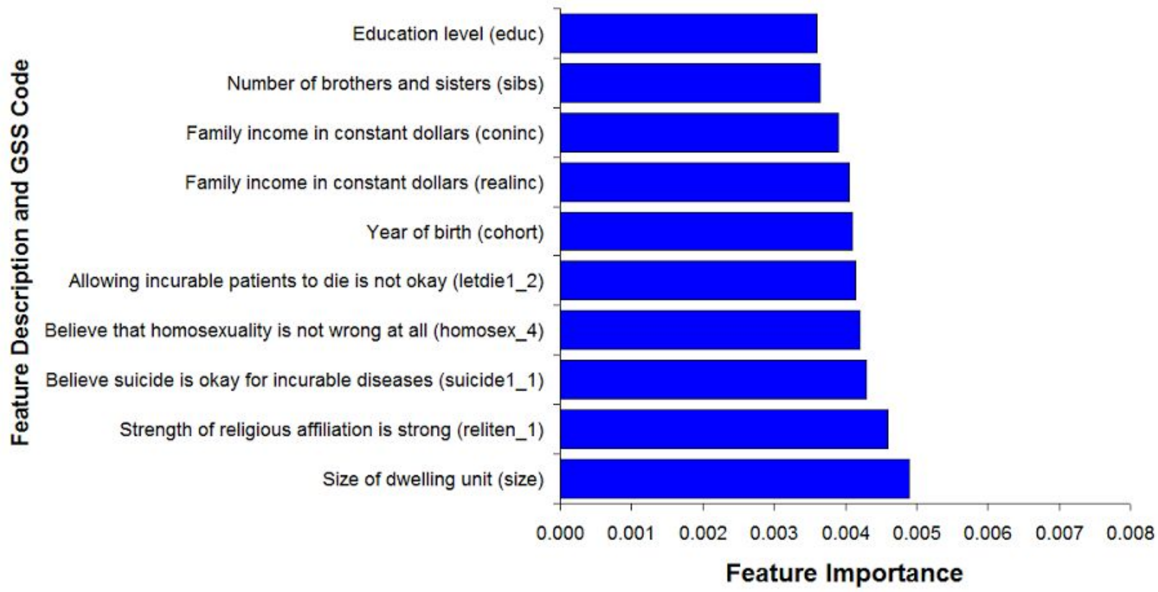
Figure 2. Top 10 important features for health related abortion attitudes (model 1).



As shown in Figure 2, attending religious services more than once per week was the most important feature influencing health related abortion attitudes. Beliefs in support of and against homosexuality, beliefs in support of and against suicide or allowing persons with incurable diseases to die, and beliefs against premarital sex were among the most important. Education level was also important.

Similarly, the top ten most important features for non-health related abortion attitudes (model 2) are summarized in Figure 3.

Figure 3. Top 10 important features for non-health related abortion attitudes (model 2).



For non-health related abortion attitudes, size of dwelling unit and having a strong religious affiliation were the most important predictive features. Beliefs in support of suicide for persons with incurable diseases, beliefs in support of homosexuality, beliefs against allowing incurable patients to die, age, family income, number of siblings, and education level were also important features.

While these initial results provided some insight into the datasets, we were concerned that the data's high dimensionality and the complex correlations between features were not adequately being addressed or captured by the baseline model approach. Two possible approaches were considered to reduce dimensionality and address feature correlation: regularization and vector dimensionality techniques. However, regularization is a type of supervised learning (considering the relationship between the outcome and predictors) and given the survey nature of the feature set that involves highly correlated attitudinal responses, a regularized sparse model like LASSO is less appropriate. The vector dimensionality techniques considered were principal component analysis (PCA) and factor analysis. PCA is used to find the dimensions in the data that can explain the most (or least) amount of variance in the data, while factor analysis is used to describe variability among observed, correlated features in terms of a potentially lower number of unobserved features called factors. Factor analysis is commonly used in social science studies where surveys, such as the GSS that we used, are the main data source. Factor analysis assumes there are latent factors that drive the correlation in observed variables. Thus, since factor analysis performs both dimensionality reduction while considering complex feature correlations, factor analysis was chosen as the most appropriate technique for our paper.

Topic Analysis

To gain a better understanding of the potential underlying factors that may exist in a factory analysis, we first conducted a topic analysis of the features. Based on the description of categories provided in the GSS and CPS codebooks, we classified the features into 19 topics, which are listed in Table 5.

Table 5. Results of topic analysis: List of 19 topics.

INTERVIEW, WORK, MARITAL AND SEXUAL RELATIONS, FAMILY, RESPONDENT STATISTICS, EDUCATION, ETHNICITY, RESIDENCE, LOCATION, INCOME, POLITICAL VOTES, ATTITUDES, RELIGION, BEHAVIOR, HEALTH, REASONING, CRIME, TARGET, ABORTION CIRCUIT CASES

Feature Dictionaries

To facilitate the organization of the ~3,700 features, we then created dictionaries to store each feature's code name, description, dataset origin, and topic category. The descriptions of the features from the GSS and CPS datasets were obtained from their original websites using the Python library "Beautiful Soup". This library allowed us to get the HTML code of those websites and extract the complete questions of each feature from the GSS and CPS surveys. The descriptions of the features from the ACP and UCR datasets were manually entered. These dictionaries were later used extensively to evaluate the features within each factor.

Factor Analysis

We then used factor analysis to detect the underlying latent grouping of the features. Factor analysis is a dimensionality reduction technique that assumes the existence of latent factors (Fricker et al. 2012). Theoretically, the model is:

$$X = \Lambda F + \psi E,$$

where Λ is the matrix of the loadings for the common factors of dimension ($p \times r$) and Ψ is a matrix of dimension $p \times p$ with $\psi 1$ ... ψp on the diagonal and all diagonal entries as zero.

Factor analysis requires a specific number of latent factors, K, so we ran decision tree, logistic regression, and random forests classifiers to find the optimal K. These classifiers and the hyperparameters used are summarized in Table 6.

Table 6. Classifiers and their hyperparameters used to find the optimal K.

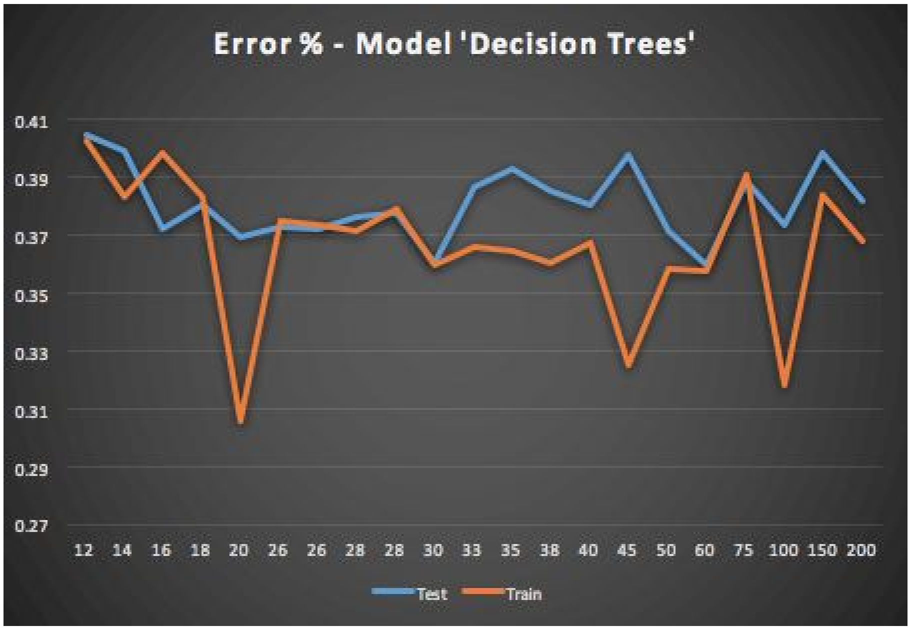
Classifier	Hyperparameter
Decision Tree	min_samples_leaf : 50, 100, 500, 1000
Logistic Regression	L1 and L2 regularizations, with parameters [10^i for i in range(-2, 2)]
Random Forests	n_estimators: 100, 200, 500, 1000

The optimal K was defined as the K corresponding to the highest AUC and lowest error performance, while also maintaining a meaningful grouping, as explored in the topic analysis.

Selection of the Optimal Number of Factors

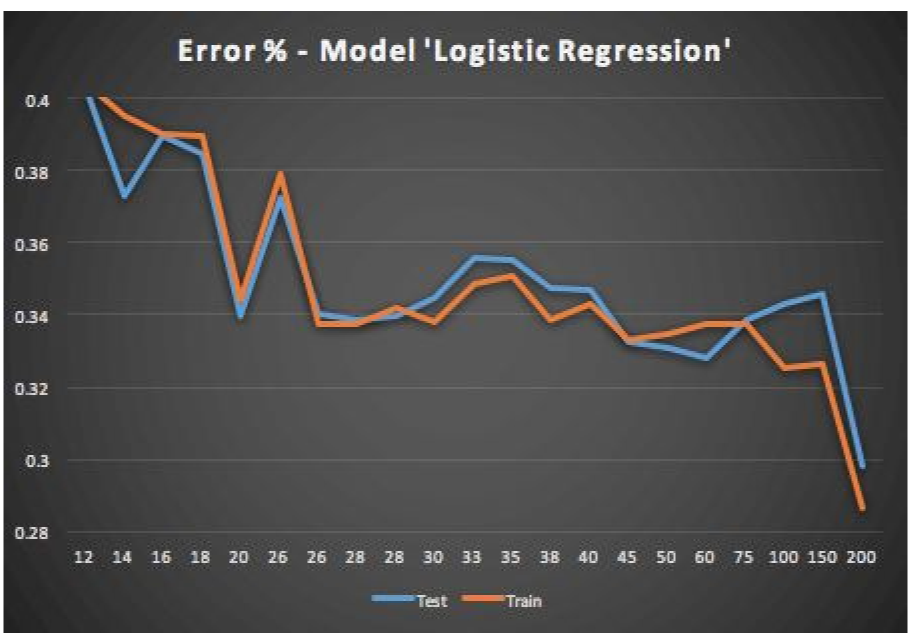
We ran the three classifiers from Table 6 for a range of K = [10, 500]. We selected the best model configuration for each K and each classifier, based on cross-validation and AUC as the main performance indicator. We then evaluated the classifiers' AUC and error rates on the test dataset for each K. The results of these performance indicators for the non-health related abortion model (model 2) are shown in Figures 4-9. Note that the performance indicators for the health related model (model 1) are not shown since they were similar.

Figure 4. Latent factor K versus percent error with decision trees model.



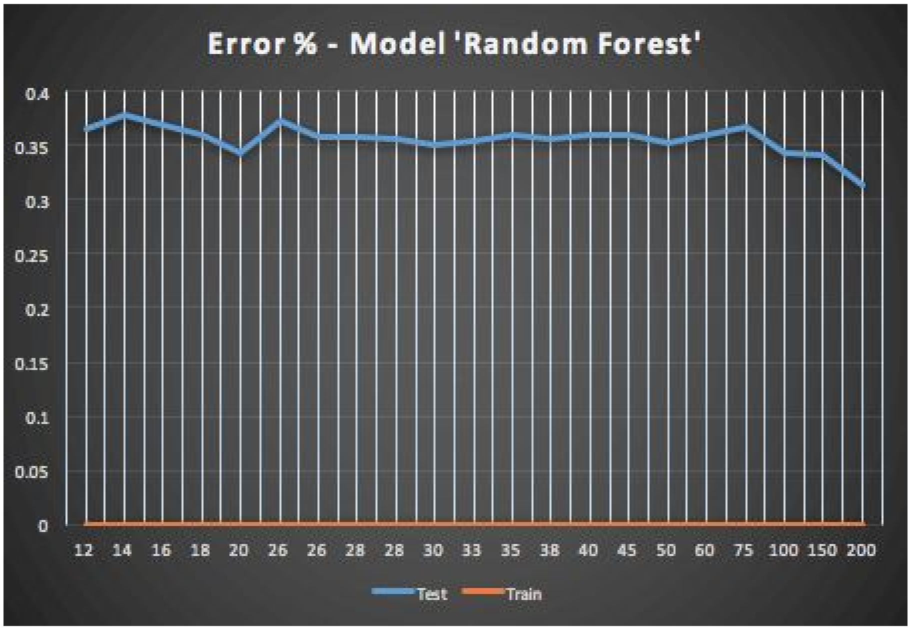
As shown in Figure 4, the lowest test dataset error rates from the decision trees classifier range from 36-37%, corresponding to K values of approximately 65, 30, and 20.

Figure 5. Latent factor K versus percent error with logistic regression model.



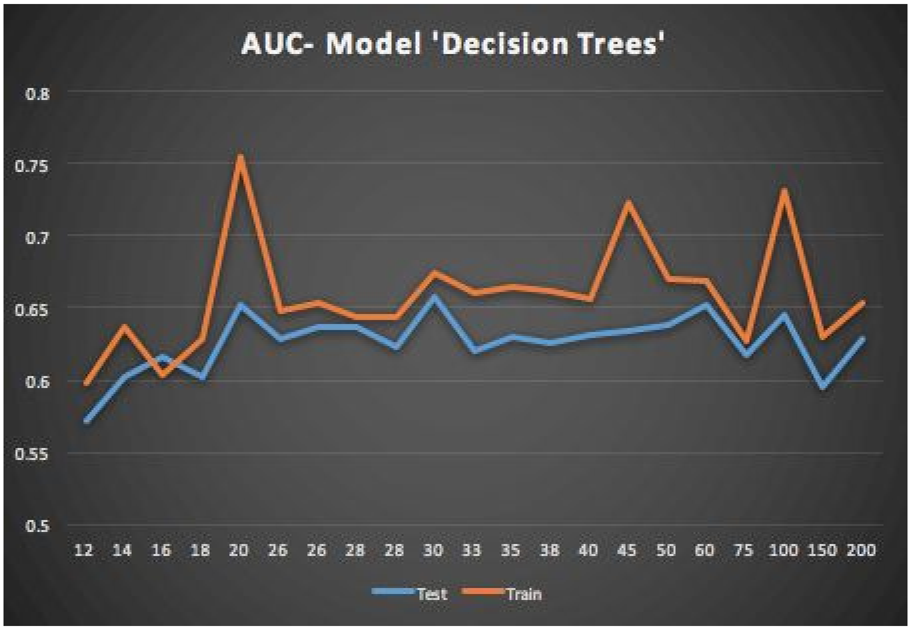
As shown in Figure 5, the lowest test dataset error rate from the logistic regression classifier is below 30%, corresponding to a K value of 200. The next lowest test dataset error rates range from 33-34%, corresponding to K values of approximately 60, 28, and 20.

Figure 6. Latent factor K versus percent error with random forests model.



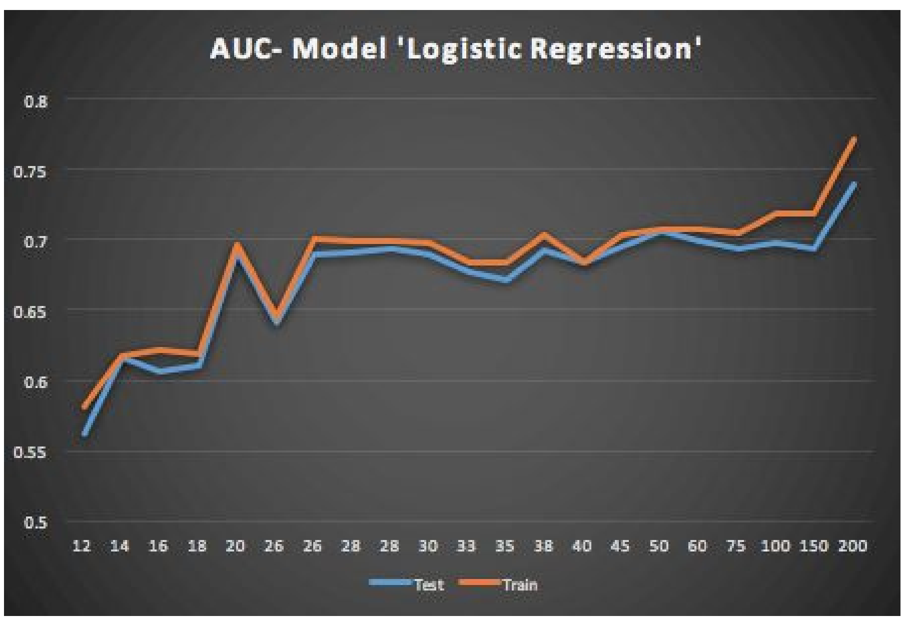
As shown in Figure 6, the lowest test dataset error rates from the random forest classifier range from 32-35%, corresponding to K values of 200 and 20.

Figure 7. Latent factor K versus AUC with decision trees model.



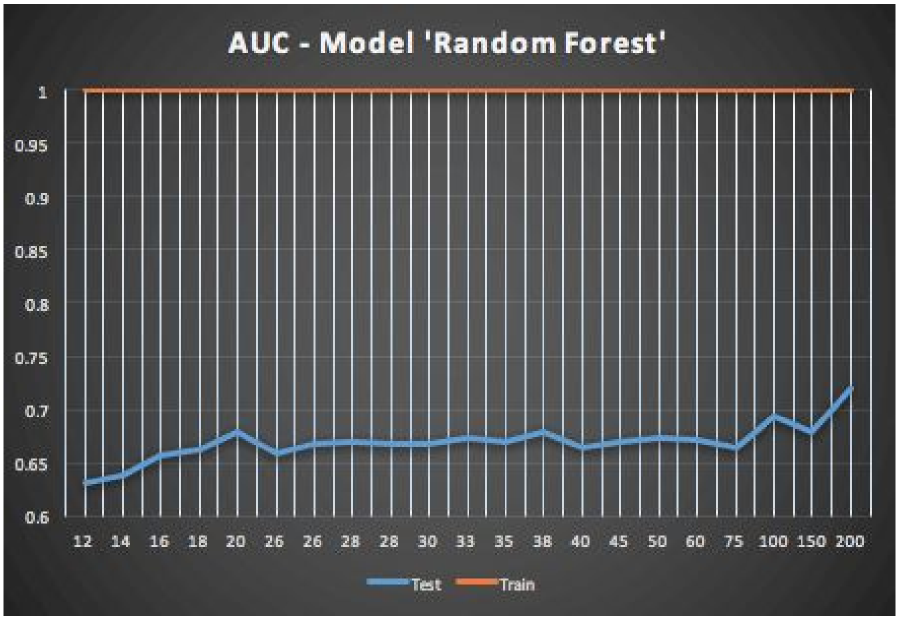
As shown in Figure 7, the highest test dataset AUC from the decision trees classifier is around 0.65, corresponding to K values of 20, 30, and 65.

Figure 8. Latent factor K versus AUC with logistic regression model.



As shown in Figure 8, the highest test dataset AUC from the logistic regression classifier is 0.73, corresponding to a K value of 200. The next highest test dataset AUC values are around 0.7, corresponding to K values of 20 and 50.

Figure 9. Latent factor K versus AUC with random forest model.



As shown in Figure 9, the highest test dataset AUC from the random forest classifier is 0.73, corresponding to a K value of 200. The next highest test dataset AUC values range from 0.68-0.70, corresponding to K values of 20, 38, and 100.

Since K values greater than 30 were mostly empty factors, we narrowed the optimal K range to K = [10, 35]. Combining this with the error rate and AUC results from figures 4-9, this further narrowed the optimal K range to around 20. We then examined the factor loadings of the features in each factor to see how the features' factor loadings changed for the K values around 20. Through this evaluation, we found that the features' factor loadings became insignificantly distributed to factors greater than K=18. Thus, K=18 was chosen as the optimal K for both models.

Final Models

The random forests, logistic regression, and decision trees classifiers were run with the optimal K of 18 to produce final models.

As shown in Table 7, logistic regression is the best classifier for health related abortion attitudes (model 1), based on AUC and accuracy on the test dataset.

Table 7. Health related abortion attitudes: Model 1 comparisons for K=18 factors.

Model 1	AUC Test	Accuracy Test	Setting
Random Forests	61.8%	82.9%	N_estimators = 1000
Logistic Regression	63.1%	83.3%	L2, C= 1
Decision Trees	61.5%	83.3%	Min_leaf = 500

As shown in Table 8, random forests is the best classifier for non-health abortion attitudes (model 2), based on AUC and accuracy on the test dataset.

Table 8. Non-health related abortion attitudes: Model 2 comparisons for K=18 factors.

Model 2	AUC Test	Accuracy Test	Setting
Random Forests	66.2%	64%	min_leaf = 1000
Logistic Regression	61.1%	61.5%	L2, C= 1
Decision Trees	60.2%	61.9%	N_estimators = 500

IV. Discussion

Predicting Attitudes Towards Health Related Abortions (Model 1)

As an interpretive diagnostic, we report the general positive or negative correlation between each of the 18 factors and health related abortion attitudes in Figure 4. We then evaluate the features with the largest factor loadings within each factor. These results are summarized in Table 9.

Figure 4. Factor importance and its correlation with health related abortion attitudes (model 1).

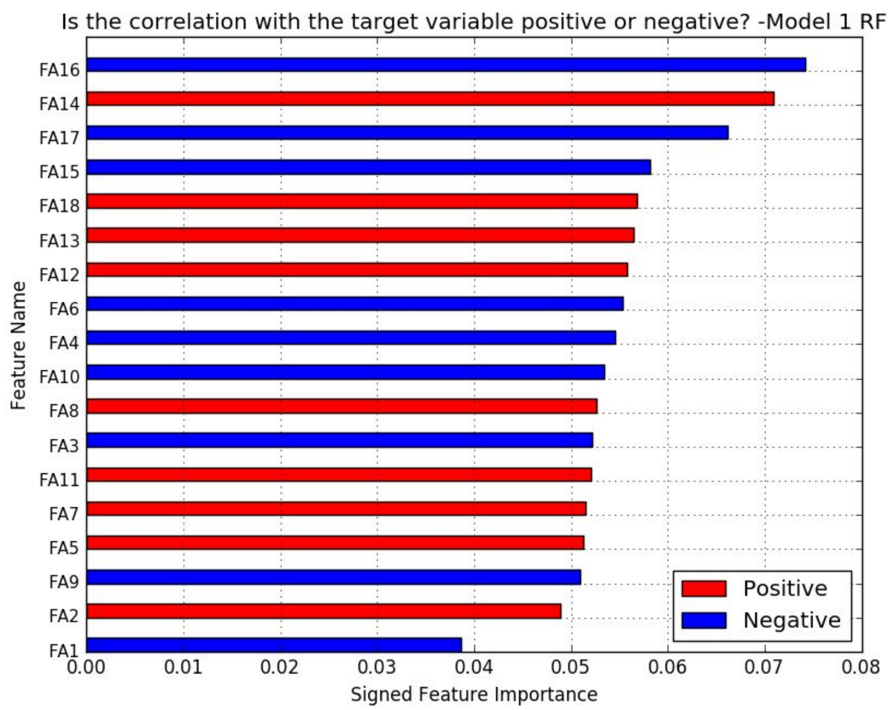


Table 9. Factor analysis: Summary of features correlated to health related abortion attitudes (model 1).

FACTORS			ASPECTS IN FAVOR OF HEALTH RELATED ABORTION			ASPECTS AGAINST HEALTH RELATED ABORTION		
NUMBER	MOST LOADS >= 0.6	MOST LOADS >= 0.4	GSS, CPS [a]	UCR	ACP [b]	GSS, CPS [a]	UCR	ACP [b]
1	Yes	Yes	Job satisfaction, perception of job security, amount of discrimination and harassment experienced in the workplace	-	-	Work life balance, physical health status, workplace environment	-	-
2	Yes	Yes	Occupation, occupational prestige, personal experience as a victim of a property crime, socioeconomic index, number of sex partners, parents' occupation and socioeconomic index	-	-	Occupation and occupational prestige, spouse's occupation and occupational prestige, presence of family members during the interview	-	Pool: female, Republican and female
3	Yes	Yes	Level of hostility at work, volunteer actions, knowledge of genetic testing (Ethnicity, gender)	-	Panel: female, not religous, Democrat and not religious, female and secular, Republican and female	Mental health status in past month (Age, number of years participating as a GSS interviewer)	-	-
4	No	Yes	Marital status, spouse's employment status and situation	-	-	Spouse's education and work status	-	-
5	Yes	Yes	Membership in different types of groups/organizations, attitude towards using physical violence to control a situation, divorce status, attitudes towards Russia, Canada, Japan, China, and Israel	-	-	Attitude towards suicide, group/organization memberships, attitude towards sex education in public schools, willingness to vote for a female president or a Black president	-	-
6	Yes	Yes	-	-	Pool: Protestant, mainline Protestant, mainline Protestant and received BA in same state as current practice, Protestant and received BA in same state as current practice, ABA-rated and Evangelist, ABA-rated and Protestant, Republican and mainline Protestant, Democrat and Protestant	-	-	Pool: Jewish, Jewish and secular, secular, Republican and Jewish, Republican and secular, ABA-rated and Jewish, Catholic, Republican and Catholic, Jewish and received BA degree in same state as current practice, secular and received BA degree in same state as current practice
7	Yes	Yes	Beliefs about heaven and God, attitude towards extramarital sex, number of friends/family/acquaintances with AIDS	-	-	Prayer frequency, involvement in religious activities, parents' attendance to religious services	-	-
8	No	Yes	General health condition, opinions on desirable qualities in children, number of recently divorced relatives, attitude towards birth control information	-	Panel: Republican and non-white Pool: Evangelist and non-white, Evangelist and black, Evangelist and female	Number of recent traumatic events experienced by self or relatives, time spent listening to radio, opinions on desirable qualities in children, number of recently divorced relatives	-	-
9	No	Yes	Attitude towards whether children should be obedient or think for themselves, frequency of refusing to eat meat for moral or environmental reasons	Crimes rates in area, particularly violent crimes and property crimes	Pool: Republican and ABA-rated	Ideal number of children in a family	-	-
10	No	Yes	-	-	-	Attitude towards and involvement in environmental issues	-	-
11	No	No	Frequency of recent poor physical health, happiness/unhappiness towards life, frequency of face-to-face contact with friends and family, attitudes towards interracial marriage	-	-	-	-	-
12	No	No	Religious preference	-	-	Satisfaction with work/occupation, belief in life after death, attitude towards interracial marriage	-	-
13	No	No	Job satisfaction, workplace atmosphere, management's treatment of worker safety	-	-	Length of employment at current workplace	-	-
14	No	No	Work schedule, work arrangement, work productivity, presidential candidate voted for in year 2000 election, attitude towards interracial marriage	-	-	-	-	-
15	No	No	Level of energy at end of day, attitude towards government provision to the elderly, income, ethnicity	-	-	Work schedule, pride towards working for employer, beliefs about the Bible	-	-
16	No	No	Feelings towards worker safety, number of workplace injuries experienced in past year, work life balance, attitude towards being a Good Samaritan	-	-	Frequency of contact with family and friends	-	-
17	No	No	Understanding of job expectations, perception of co-workers' work ethic	-	-	Work arrangement, presidential candidate voted for in year 2000 election	-	-
18	No	No	Ability to work from home, ethnicity, frequency of contact with friends and family via letter writing	-	-	Frequency of email activity, internet usage	-	-

Notes:

[a] The majority of aspects listed in the GSS, CPS column refer to respondent responses to questions asked in the GSS or CPS survey. Aspects in parenthesis indicate an aspect pertaining to the GSS or CPS interviewer.

[b] Abortion circuit panel aspects represent the judge circuit panel composition (indicated by 'panel') and the judge circuit pool composition (indicated by 'pool').

Combining the results of Figure 4 and Table 9, the highest weighted factor was factor 16, which consists of GSS and CPS features. Within factor 16, the features with the largest factor loadings correlated in favor of health related abortions are the frequency of contact with family and friends. The features with the largest factor loadings correlated against health related abortions are respondents' feelings towards worker safety, number of workplace injuries experienced in past year, work life balance, and attitude towards being a Good Samaritan.

The second highest weighted factor was factor 14, which also consisted of GSS and CPS features. Within factor 14, features correlated in favor of health related abortions are respondents' work schedule, work arrangement, work productivity, presidential candidate voted for in year 2000 election, and attitude towards interracial marriage.

The highest weighted factor with ACP features is factor 6. Of the 226 ACP features, 71 of them are associated with factor 6, composing 28% of all the features in factor 6. Within factor 6, the ACP features with the largest factor loadings correlated in favor of health related abortions consist of several judge circuit pool composition aspects, including: Protestant, mainline Protestant, mainline Protestant and received BA in same state as current practice, Protestant and received BA in same state as current practice, ABA-rated and Evangelist, ABA-rated and Protestant, Republican and mainline Protestant, Democrat and Protestant. Within factor 6, the ACP features with the largest factor loadings correlated against health related abortions consist of several judge circuit pool composition aspects, including: Jewish, Jewish and secular, secular, Republican and Jewish, Republican and secular, ABA-rated and Jewish, Catholic, Republican and Catholic, Jewish and received BA degree in same state as current practice, secular and received BA degree in same state as current practice.

The second highest weighted factor with ACP features is factor 8. Of the 226 ACP features, 26 of them are associated with factor 8, composing 6.7% of all the features in factor 8. Within factor 8, the ACP features with the largest factor loadings correlated in favor of health related abortions consist of the actual judge circuit panel composition of Republican and non-white judges, and several judge circuit pool composition aspects, including: Evangelist and non-white, Evangelist and black, Evangelist and female. These ACP features are associated with GSS and CPS features pertaining to respondents' general health condition, opinions on desirable qualities in children, number of recently divorced relatives, and attitudes towards birth control information. There are no ACP features within factor 8 correlated against health related abortions.

The highest weighted factor with UCR features is factor 9. All 7 of the UCR features are associated with factor 9, composing 4% of all the features in factor 9. Within factor 9, the

UCR features with the largest factor loadings correlated in favor of health related abortions mainly consist of the rates of violent crimes and property crimes. These UCR features are associated with ACP features pertaining to the judge circuit pool composition of Republican and ABA-rated judges. These UCR features were also associated with GSS and CPS features pertaining to respondents' attitudes towards whether children should be obedient or think for themselves and their frequency of refusing to eat meat for moral or environmental reasons. There were no UCR features within any of the factors correlated against health related abortions.

The features capturing the lag time between an abortion related court ruling and a GSS respondent's interview date did not have significant factor loadings within any of the factors. This suggests that the time frame of abortion related circuit court case rulings influencing the public is not as impactful judge circuit panel and judge circuit pool compositions in affecting public attitudes towards health related abortions. Table 10 summarizes the lag times and their factor loadings, with highest weighted factors presented first.

Table 10. Summary of results for lag times between abortion related court ruling and respondent interview date for health related abortion attitudes (model 1).

Factor	Lag Time Variables (in Months)	Variable Factor Loading
6	4	- 0.175
	5	0.109
8	7	- 0.184
	6	- 0.114
	8	- 0.093
	1	0.156
7	9	- 0.170
	10	- 0.108
5	12	- 0.127
	3	0.111
	2	0.186
1	11	0.117

Predicting Attitudes Towards Non-Health Related Abortions (Model 2)

As an interpretive diagnostic, we report the general positive or negative correlation between each of the 18 factors and non-health related abortion attitudes in Figure 5. We then evaluate the features with the largest factor loadings within each factor. These results are summarized in Table 11.

Figure 5. Factor importance and its correlation with non-health related abortion attitudes (model 2).

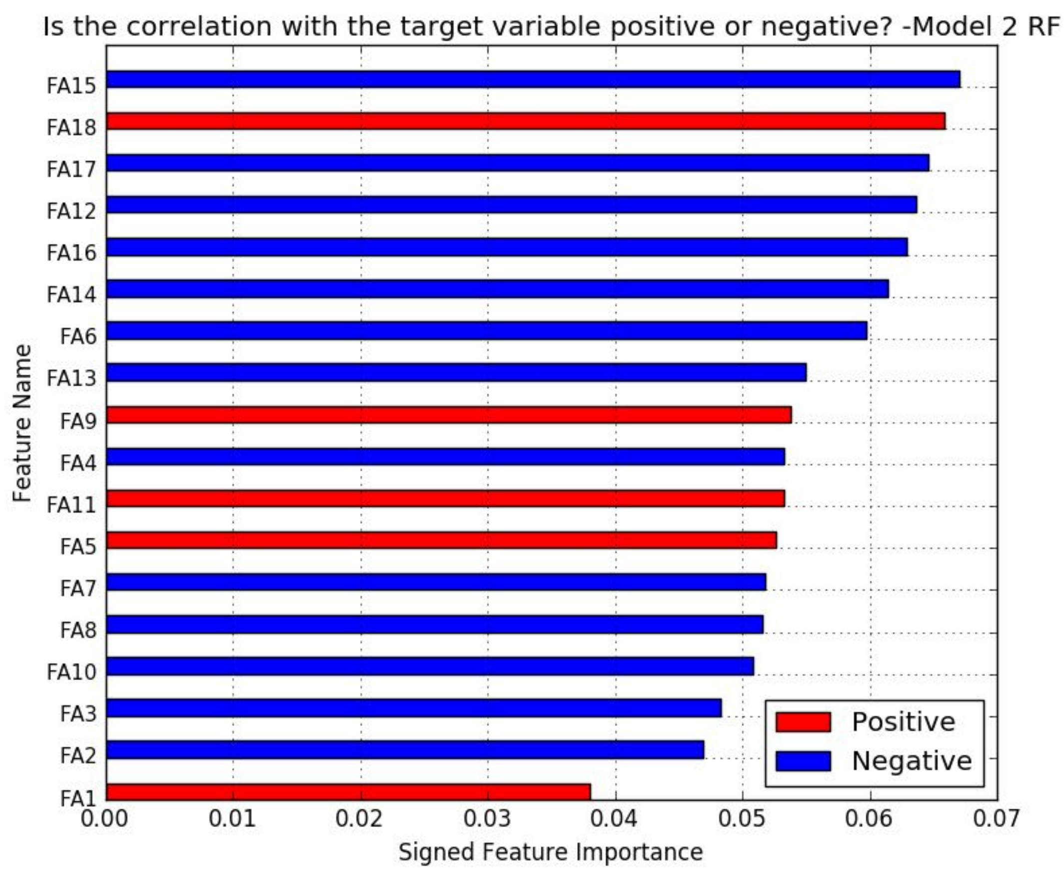


Table 11. Factor analysis: Summary of features correlated to non-health related abortion attitudes (model 2).

FACTORS			ASPECTS IN FAVOR OF NON-HEALTH RELATED ABORTION			ASPECTS AGAINST NON-HEALTH RELATED ABORTION		
NUMBER	MOST LOADS >= 0.6	MOST LOADS >= 0.4	GSS, CPS [a]	UCR	ACP [b]	GSS, CPS [a]	UCR	ACP [b]
1	**Yes**	**Yes**	Amount of discrimination and harassment experienced at work, laid off or not in past year, work arrangement	-	-	Length of employment at current workplace, frequency of recent poor physical health, number of workplace injuries experienced in past year, amount of overtime work, work life balance	-	-
2	**Yes**	**Yes**	Occupation, occupational prestige (Race, gender)	-	Pool: female, Republican and female, not religious, and not religious and secular	Number of weeks worked in past year, ethnicity, number of sex partners in past year, poverty status, experience as a victim of a property crime, frequency of recent poor mental health, age when first child born (Age, years of experience as an interviewer)	-	-
3	**No**	**Yes**	Spouse's religious preference, family income when 16 years old (Observed ethnicity of respondent)	-	-	Spouse's religious preference, presidential candidates voted for in 1992 and 1996 elections, attitude towards government investigating potential government employees for secret or top secret clearance, feelings towards Protestants, Catholics, and Jews, attitude towards interracial marriage, number of children older than 5 years old (Certainty of observed race of respondent)	-	-
4	**No**	**No**	Current and previous marital status, spouse's employment situation, happiness of marriage, spouse's highest degree of education, number of persons in household, total family income	-	-	Spouse's highest level of education, spouse's occupation and occupational prestige, spouse's socioeconomic index, marital status, poverty status	-	-
5	**Yes**	**Yes**	Membership in different types of groups/organizations, attitude towards using physical violence to control a situation, number of divorces in past 5 years	-	-	Number of relatives who have passed away recently and since R was 16 years old, attitude towards suicide, attitude towards sex education in public schools, willingness to vote for a female president or a Black president	-	-
6	**Yes**	**Yes**	Location of residence during interview and when age 16	-	Panel: Republican and mainline Protestant, mainline Protestant, Protestant, Republican and Protestant Pool: mainline Protestant, Protestant, received BA in same state as current practice and mainline Protestant, received BA in same state as current practice and Protestant, ABA-rated and Evangelist, ABA-rated and Protestant, Republican and mainline Protestant, Democrat and Protestant, Democrat and mainline Protestant	Location of residence during interview and when age 16	Crimes against society, crimes against property, and violent crimes	Pool: Jewish, Jewish and secular, secular, Republican and Jewish, Republican and secular, ABA-rated and Jewish, Catholic, Republic and Catholic, Jewish and received BA degree in same state as current practice, secular and received BA degree in same state as current practice
7	**No**	**Yes**	Involvement in religious activities, prayer frequency, parents' attendance to religious services, ethnicity, personal experience being picked up or charged by the police	-	Panel: Protestant and received BA degree in same state as currently practicing in, mainline Protestant and received BA degree in same state as currently practicing in Pool: Democrat	Beliefs about heaven and God, religion raised in, mother's religious preference, attitude towards extramarital sex, outlook on the purpose of life, number of friends/family/acquaintances with AIDS	-	Panel: ABA-rated and Jewish, ABA-rated and secular Pool: Republican
8	**No**	**No**	Standard of living compared to parents' standard of living, outlook on ability to improve standard of living	-	-	General condition of health, opinions on desirable qualities in children, attitude towards whether Americans are divided or united, frequency of refusing to eat meat for moral or environmental reasons	Crimes against persons, crimes against property, crimes against society	Pool: Evangelist and black, Republican and non-white, Evangelist and non-white, Evangelist and female
9	**No**	**No**	Ideal number of children in a family	-	-	Attitude towards women working while raising children, attitude towards a bad marriage versus no marriage, work status of mother while growing up, attitude towards whether children are a joy or an interference	-	Pool: Republican and ABA-rated
10	**No**	**No**	Ability to work from home, attitude towards judicial mistakes, attitude towards government involvement in economy, attitude towards breaking laws, attitude towards government spending on health	-	-	Attitude towards and involvement in environmental issues, confidence in the existence of God, attitude towards modern science	-	-
11	**No**	**No**	Mental health status in past month, frequency of cutting back on driving for environmental reasons, father's occupational prestige	-	-	Attitude towards Blacks overcoming prejudice, belief in life after death, attitude towards living in a half Black neighborhood, attitude towards interracial marriage, attitude towards the intelligence of Blacks and Whites	-	-
12	**No**	**No**	Job satisfaction, work life balance, happiness/unhappiness towards life, participation in blood donations, attitude towards interracial marriage, religion raised in, ethnicity	-	-	Status of stock ownership, job satisfaction and workload, mother's socioeconomic index, living with both mother and father or not when age 16	-	-

13	**No**	**No**	Job satisfaction, political party affiliation, frequency of email activity	-	-	-	-	-
14	**No**	**No**	Workplace safety, job satisfaction, number of friends and relatives at church, income, age	-	-	General condition of health, attitude towards assisting people in trouble, attitude towards the intelligence of Blacks, frequency of email activity		
15	**No**	**No**	Job satisfaction, religious preference, religion raised in, beliefs about the Bible	-	-	Income		
16	**No**	**No**	Internet usage, liberal or conservative political views	-	-	Perception of coworkers' quality of work and work ethic, has lent money to someone or not in past year, ethnicity, frequency of attendance to religious services, frequency of internet activity		
17	**No**	**No**	-	-	-	-	-	-
18	**No**	**No**	Extra hours worked	-	-	Amount of freedom and say in job, attitude towards interracial marriage, consider political views as liberal or conservative	-	-

Notes:

[a] The majority of aspects listed in the GSS, CPS column refer to respondent responses to questions asked in the GSS or CPS survey. Aspects in parenthesis indicate an aspect pertaining to the GSS or CPS interviewer.

[b] Abortion circuit panel aspects represent the judge circuit panel composition (indicated by 'panel') and the judge circuit pool composition (indicated by 'pool').

Combining the results of Figure 5 and Table 11, the highest weighted factor is factor 15, which consists of all GSS and CPS features. Within factor 15, the feature with the largest factor loading correlated against non-health related abortions is respondents' income. The features with the largest factor loading correlated in favor of non-health related abortions are respondents' job satisfaction, religious preference, religion raised in, and beliefs about the Bible.

The second highest weighted factor is factor 18, which consists of all GSS and CPS features. The features with the largest factor loading correlated against non-health related abortions is respondents' amount of freedom and say in job, attitude towards interracial marriage, and whether the respondent considered their own political views to be liberal or conservative. Within factor 15, the predominant feature correlated in favor of non-health related abortions is how many extra hours respondents worked.

The highest weighted factor with ACP features is factor 6. Of the 226 ACP features, 78 of them are associated with factor 6, composing 29% of all the features in factor 6. Within factor 6, the ACP features with the largest factor loadings correlated in favor of non-health related abortions consist of several judge circuit panel composition aspects (Republican and mainline Protestant, mainline Protestant, Protestant, Republican and Protestant) and several judge circuit pool composition aspects (mainline Protestant, Protestant, received BA in same state as current practice and mainline Protestant, received BA in same state as current practice and Protestant, ABA-rated and Evangelist, ABA-rated and Protestant, Republican and mainline Protestant, Democrat and Protestant, Democrat and mainline Protestant). Within factor 6, the ACP features with the largest factor loadings correlated against non-health related abortions consist of various judge circuit pool composition aspects, including: Jewish, Jewish and secular, secular, Republican and Jewish, Republican and secular, ABA-rated and Jewish, Catholic, Republic and Catholic, Jewish and received BA degree in same state as current practice, secular and received BA degree in same state as current practice.

Factor 6 also contains all 7 UCR features correlated against non-health related abortions. These UCR features compose 2.6% of all the features in factor 6 and consist of crime rates for crimes against society, crimes against property, and violent crimes. The GSS and CPS features included in factor 6 are the respondents' location of residence during the interview and at age 16 years.

The second highest weighted factor with ACP features is factor 9. Of the 226 ACP features, 8 of them are associated with factor 9, composing 4.7% of all the features in factor 9. Within factor 9, the ACP features with the largest factor loadings correlated against non-health related abortions consist of the expected judge circuit panel composition of Republican and ABA-rated judges. These ACP features are associated with GSS and CPS features pertaining to respondents' attitudes towards women working while raising children, attitudes towards a bad marriage versus no marriage, work status of mother while growing up, and attitudes towards whether children are a joy or an interference. There are no ACP features within factor 9 correlated in favor of non-health related abortions.

The features capturing the lag time between an abortion related court ruling and a GSS respondent's interview date did not have significant factor loadings within any of the factors. This suggests that the time frame of abortion related circuit court case rulings influencing the public is not as impactful as judge circuit panel and judge circuit pool compositions in affecting public attitudes towards health related abortions. Table 12 summarizes the lag times and their factor loadings, with highest weighted factors presented first.

Table 12. Summary of results for lag times between abortion related court ruling and respondent interview date for non-health related abortion attitudes (model 2).

Factor	Lag Time Variables (in Months)	Variable Factor Loading
6	4	- 0.176
	5	0.106
5	12	- 0.124
	3	0.111
	2	0.187
7	10	0.102
	9	0.164
8	1	- 0.106
	8	0.081
	7	0.116
3	6	- 0.121
1	11	0.094

V. Conclusion

The relationship between courts and societal attitudes is highly debated. From abolition of slavery, to women's liberation, to environmentalism, law is speculated to play a key role in moral revolutions, yet little causal evidence exists to date. We hope this approach illustrates how machine learning can shed light to these debates. We constructed an exogenous, random component of court rulings---the composition of the panel of judges assigned to cases. We then implemented machine-learning algorithms to understand if an exogenous component of court rulings are predictive of societal attitudes. The algorithm also reports a feature weight to benchmark the magnitude of the role of courts as an instrument of social change. Future work can explore the method for other areas of law and society.

VI. References

Chen, D.L., Levonyan, V., and Yeh, S. *Do Policies Affect Preferences? Evidence from Random Variation in Abortion Jurisprudence.* 2015. [Accessed: http://users.nber.org/~dlchen/papers/Policies_Affect_Preferences.pdf]

Chen, D.L. and Yeh, S. *How Do Rights Revolutions Occur? Free Speech and the First Amendment.* 2015. [Accessed: http://users.nber.org/~dlchen/papers/How_Do_Rights_Revolutions_Occur.pdf]

Fricker et al. *From Data to Information: Using Factor Analysis with Survey Data*. Phalanx. Dec 2012.

Muennig et al. *Codebook for General Social Survey 1978 - 2002.* April 2013. [Accessed: http://gss.norc.org/Documents/other/GSS-NDI%20Codebook.pdf].

Smith, T. W., Marsden, P., Hout, M., & Kim, J. *General Social Surveys, 1972–2012.* National Opinion Research Center. Chicago, IL. 2013.

Sunstein, C. R. *Are judges political?: an empirical analysis of the federal judiciary.* Brookings Institution Press. 2007.

Tryfos P. Chapter14: Factor Analysis. York University. 1997.

U.S. Department of Justice and F.B.I. Uniform Crime Reporting Handbook. 2004.

U.S. Census Bureau, Bureau of Labor Statistics, Current Population Survey, 2015 Annual Social and Economic (ASEC) Supplement. 2015.

Zuk Gary; Barrow, G. S., Deborah J.; Gryski. *Multi-User Database on the Attributes of United States Appeals Court Judges, 1801-2000.* ICPSR - Interuniversity Consortium for Political and Social Research. 1997. [Accessed: http://dx.doi.org/10.3886/icpsr06796 doi: 10.3886/icpsr06796].

Appendix A: Circuit Panel Data Processing

This appendix details the data processing and feature engineering process of the circuit panel data. Note that the data presented in this appendix are not observations from the actual dataset and are example observations used to walk through the data processing. Similarly, only a subset of the features are presented here.

The original dataset was 142 rows and 218 columns.

Table A1.

circuit	year	month	case_ID	case_name	panelvote	x_dem	Case category	E_x_femXsecular
8	1974	2	C000165	CD v. DE	2	0.5	2	0
8	1974	2	C000164	EF v. FG	3	0.5	2	0
8	1974	12	C000166	GH v. HI	3	0	2	0

In Table A1, observations with the same circuit, year, and month are highlighted in orange. Such observations were averaged together. Additionally, CaseID, CaseName, CaseCategory features were dropped. This process resulted in 137 rows and 215 columns, as illustrated in Table A2.

Table A2.

circuit	year	month	panelvote	x_dem	E_x_femXsecular
8	1974	2	2.5	0.5	0
8	1974	12	3	0	0

A 'lagdate' feature was added to capture the lag time (by month up to one year) between an abortion decision and the interview date in the GSS. This feature was ultimately merged with the GSS data's interview date feature. A 'numLagMonths' column was also added to denote the lag time (in months) that the lag date represents. This process resulted in 1644 rows and 217 columns, as illustrated in Table A3.

Table A3.

circuit	year	month	numLagMonths	lagdate	panelvote	x_dem	E_x_femXsecular
8	1974	2	1	3/1/1974	2.5	0.5	0
8	1974	2	2	4/1/1974	2.5	0.5	0
8	1974	2	3	5/1/1974	2.5	0.5	0
8	1974	2	4	6/1/1974	2.5	0.5	0
8	1974	2	5	7/1/1974	2.5	0.5	0
8	1974	2	6	8/1/1974	2.5	0.5	0
8	1974	2	7	9/1/1974	2.5	0.5	0
8	1974	2	8	10/1/1974	2.5	0.5	0
8	1974	2	9	11/1/1974	2.5	0.5	0
8	1974	2	10	12/1/1974	2.5	0.5	0
8	1974	2	11	1/1/1975	2.5	0.5	0
8	1974	2	12	2/1/1975	2.5	0.5	0
8	1974	12	1	1/1/1975	3	0	0
8	1974	12	2	2/1/1975	3	0	0
8	1974	12	3	3/1/1975	3	0	0
8	1974	12	4	4/1/1975	3	0	0
8	1974	12	5	5/1/1975	3	0	0
8	1974	12	6	6/1/1975	3	0	0
8	1974	12	7	7/1/1975	3	0	0
8	1974	12	8	8/1/1975	3	0	0
8	1974	12	9	9/1/1975	3	0	0
8	1974	12	10	10/1/1975	3	0	0
8	1974	12	11	11/1/1975	3	0	0
8	1974	12	12	12/1/1975	3	0	0

We then constructed 12 additional binary features, each representing a lag time of 1-12 months between an abortion case decision and a respondent's GSS survey date. The year and month columns were dropped after these binary features were created. This process is illustrated in Table A4 and resulted in 1644 rows and 227 columns.

In Table A4, the two rows highlighted in orange have the same lagdate and circuit. Similarly, the two rows highlighted in yellow also have the same lagdate and circuit. Such rows are merged into one row, with the judge panel composition values averaged together and the lagmonth columns summed together. Finally, the numLagMonths column is dropped since it is no longer needed. Table A5 illustrates these final steps, which result in a final dataset of 1410 rows and 226 columns.

Table A4.

circuit	numLagMonths	lagdate	panelvote	x_dem	E_x_femXsecular	lagMonth1	lagMonth2	lagMonth3	lagMonth4	lagMonth5	lagMonth6	lagMonth7	lagMonth8	lagMonth9	lagMonth10	lagMonth11	lagMonth12
8	1	3/1/1974	2.5	0.5	0	1	0	0	0	0	0	0	0	0	0	0	0
8	2	4/1/1974	2.5	0.5	0	0	1	0	0	0	0	0	0	0	0	0	0
8	3	5/1/1974	2.5	0.5	0	0	0	1	0	0	0	0	0	0	0	0	0
8	4	6/1/1974	2.5	0.5	0	0	0	0	1	0	0	0	0	0	0	0	0
8	5	7/1/1974	2.5	0.5	0	0	0	0	0	1	0	0	0	0	0	0	0
8	6	8/1/1974	2.5	0.5	0	0	0	0	0	0	1	0	0	0	0	0	0
8	7	9/1/1974	2.5	0.5	0	0	0	0	0	0	0	1	0	0	0	0	0
8	8	10/1/1974	2.5	0.5	0	0	0	0	0	0	0	0	1	0	0	0	0
8	9	11/1/1974	2.5	0.5	0	0	0	0	0	0	0	0	0	1	0	0	0
8	10	12/1/1974	2.5	0.5	0	0	0	0	0	0	0	0	0	0	1	0	0
8	11	1/1/1975	2.5	0.5	0	0	0	0	0	0	0	0	0	0	0	1	0
8	1	1/1/1975	3	0	0	1	0	0	0	0	0	0	0	0	0	0	0
8	12	2/1/1975	2.5	0.5	0	0	0	0	0	0	0	0	0	0	0	0	1
8	2	2/1/1975	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0
8	3	3/1/1975	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0
8	4	4/1/1975	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0
8	5	5/1/1975	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0
8	6	6/1/1975	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0
8	7	7/1/1975	3	0	0	0	0	0	0	0	0	1	0	0	0	0	0
8	8	8/1/1975	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0
8	9	9/1/1975	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0
8	10	10/1/1975	3	0	0	0	0	0	0	0	0	0	0	0	1	0	0
8	11	11/1/1975	3	0	0	0	0	0	0	0	0	0	0	0	0	1	0
8	12	12/1/1975	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Table A5.

circuit	lagdate	panelvote	x_dem	E_x_femXsecular	lagMonth1	lagMonth2	lagMonth3	lagMonth4	lagMonth5	lagMonth6	lagMonth7	lagMonth8	lagMonth9	lagMonth10	lagMonth11	lagMonth12
8	3/1/1974	2.5	0.5	0	1	0	0	0	0	0	0	0	0	0	0	0
8	4/1/1974	2.5	0.5	0	0	1	0	0	0	0	0	0	0	0	0	0
8	5/1/1974	2.5	0.5	0	0	0	1	0	0	0	0	0	0	0	0	0
8	6/1/1974	2.5	0.5	0	0	0	0	1	0	0	0	0	0	0	0	0
8	7/1/1974	2.5	0.5	0	0	0	0	0	1	0	0	0	0	0	0	0
8	8/1/1974	2.5	0.5	0	0	0	0	0	0	1	0	0	0	0	0	0
8	9/1/1974	2.5	0.5	0	0	0	0	0	0	0	1	0	0	0	0	0
8	10/1/1974	2.5	0.5	0	0	0	0	0	0	0	0	1	0	0	0	0
8	11/1/1974	2.5	0.5	0	0	0	0	0	0	0	0	0	1	0	0	0
8	12/1/1974	2.5	0.5	0	0	0	0	0	0	0	0	0	0	1	0	0
8	1/1/1975	2.75	0.25	0	1	0	0	0	0	0	0	0	0	0	1	0
8	2/1/1975	2.75	0.25	0	0	1	0	0	0	0	0	0	0	0	0	1
8	3/1/1975	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0
8	4/1/1975	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0
8	5/1/1975	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0
8	6/1/1975	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0
8	7/1/1975	3	0	0	0	0	0	0	0	0	1	0	0	0	0	0
8	8/1/1975	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0
8	9/1/1975	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0
8	10/1/1975	3	0	0	0	0	0	0	0	0	0	0	0	1	0	0
8	11/1/1975	3	0	0	0	0	0	0	0	0	0	0	0	0	1	0
8	12/1/1975	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1