Information Transparency and Free Speech in China

Charlie Zhang

May 15, 2021

Summary

Government transparency and free speech supplement each other to secure individual freedom and enforce government accountability. However, China's public opinions display inconsistent preferences that they value information transparency yet deem (indiscriminate) imitation of western-style freedom as dangerous. By employing parametric and non-parametric statistical learning techniques, the study intends to understand how and what could explain the configuration of preferences. Surprisingly, the study finds that well-educated or high-income people are more likely to hold these inconsistent preferences. The potential explanation is that they might understand the flaws of western-style freedom and are particularly worried about "indiscriminate" imitation. After digging into the inconsistent preferences, the study finds that democracy, market, and nationalism play a key role in configuring China's ideological landscape. Lastly, the study discusses several limitations, including failing to understand what drives inconsistencies and the limited dataset that fails to capture the underprivileged's and current situation.

1. Background

Ideology has been a central concept muddled by diverse uses in the study of politics. Converse (1964) perhaps provided a classical operationalization of ideology– a "belief system" that represents the configuration of attitudes and ideas. Since then, spatial approaches which regard ideology as mappable in a low-dimension space have been prevalent. The bipartisan system makes the United States an exemplary single-dimensional ideological mapping. The two-dimensional characterizations that include equality and freedom axes are also common (Leader Maynard & Mildenberger, 2018).

Although the study of ideology has attracted intensive attention in democratic settings, as Pan and Xu (2018, p. 254) pointed out, "it has received minimal consideration in non-democratic contexts." They further explain that inattention could relate to the perception that organized preferences may not exist in an authoritarian regime without elections. However, studying ideology can help understand how China sustains its ruling, especially in the post-Tiananmen era when the Party cannot appeal to Communist ideology as in the Maoist era (Zhao, 2009).

In their study, Pan and Xu (2018) find that some weak constraints bind together public preferences in China, and the *political*, *economic/social*, and *nationalist* dimensions would best explain China's ideological spectrum. Regions with higher economic development levels or individuals with higher income and education tend to favor *liberal*, *pro-market/non-traditional*, and *non-nationalist* end of the spectrum. In a subsequent study, Pan and Xu (2020) confirm the multi-dimensional preferences and find "nationalism is predictive of more positive assessments of China's current political system and higher

levels of trust in central and local governments." Similarly, Wu and Meng (2017) figure out that *political* and *economic* dimensions are pivotal to Chinese urban residents.

Beyond the ideological spectrum, Cantoni et al. (2017) focuses on how mandatory high school politics curricula might shape student's ideology. They find that the new curriculum introduced in 2005 has successfully changed students' attitudes in the Chinese government's direction. Students exposed to the new curriculum generally see China as more democratic and express skepticism towards the unconstrained democracy and free markets.

Undoubtedly, these studies provide a contour of the Chinese ideological spectrum and what key factors shape people's ideology. Specifically, a pair of the controversial and recurring issues in China's public space is government transparency and free speech. One side of the scale are these themes, and the other is stability. Previous studies found that the Chinese government would censor and remove the posts that would potentially bring collective actions (King et al., 2013). Interestingly, it is not merely the authority valuing the stability over transparency and freedom, but also a significant proportion of the public agreeing on such preferences. Nevertheless, when the crisis happens, the same group of people often criticize the government's hiding information and suppress the whistle-blower, which is the exact situation of China's early-stage Covid-19 situations. The domain of China's public opinion makes one confused, and it is pivotal to understand how the Chinese perceive government transparency and free speech in China and what could explain their configuration of (in)consistent preferences.

2. Data

The study intends to use the *zuobiao* survey (Chinese Political Compass) that contains 50 questions about a wide array of critical issues in China (*Chinese Political Compass*, 2018). A group of graduate students and researchers at the Peking University designed the survey, the same source used in Pan and Xu (2018)'s study. The dataset, by default, does not contain any missing values, as it only records respondents who finished the survey. As Pan and Xu (2018) have reweighted the dataset based on China's population distribution, each region's ratio represents the actual population density.

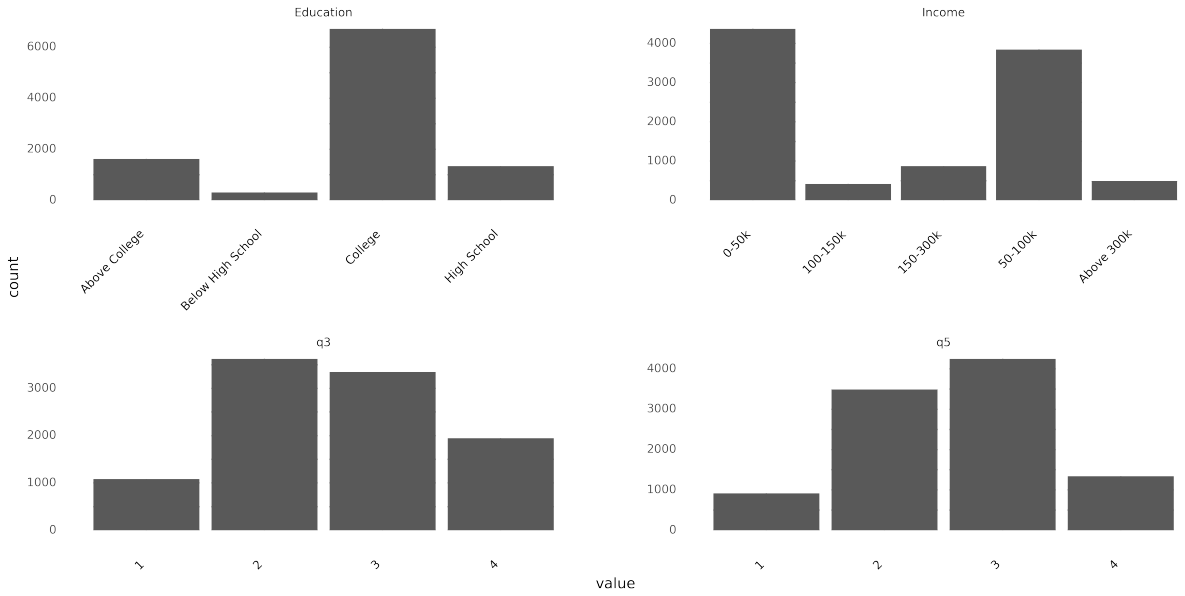
Table 1: Descriptive Statistics

Variables	Mean	Variables	Mean
Age	39.252	Below high school	0.032
Gender		High school	0.134
Male	0.507	College	0.672
Female	0.493	Above college	0.162
Annual Income (in CNY):		Region:	
0- 50k	0.437	North	0.139
50- 100k	0.385	Northeast	0.113
100- 150k	0.042	East	0.315
150- 300k	0.087	South Central	0.294
Above 300k	0.049	Southwest	0.089
Education:		Northwest	0.051

Regarding the demographic data, respondents' ages range from 18 to 60, and the average is roughly 39 years old. As Table 1 shows, 43.7% of respondents' annual income is below 50,000 Yuan, 38.5%

falls within 50,000 to 100,000 Yuan, and the remaining 17.8% is above 100,000 Yuan. Compared to the 8% college population in China, surprisingly, more than 80% of respondents have received a college degree or above. Note that the median of Disposable Income in 2014 was 26,635 Yuan, and the respondents' demography reflects that they are well-educated with higher income than the average Chinese (National Bureau of Statistics of China, 2015).

Figure 1: Histogram of Variables of Interest

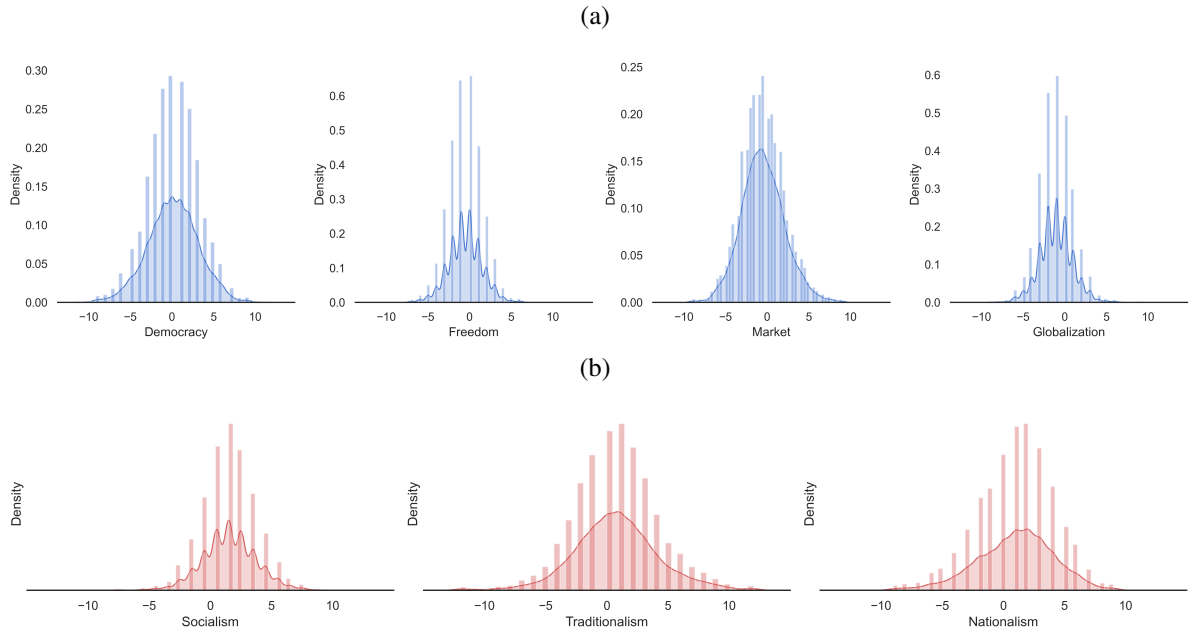


Each question in the survey is on a four-point Likert scale– "strongly disagree (1)," "disagree (1)," "agree (1)," and "strongly agree (1)." *Q5* and *Q4* constitute our outcome of interest. *Q5* measures people's perception of replicating western-style freedom in China, and *Q3* sets up a hypothetical yet usual situation in China to let people make the judgment. As Table 2 displays, around 55% of respondents disagree with the statement that indiscriminately imitating western-style freedom of speech will lead to social disorder in China, while the rest 45% of respondents agree or strongly agree. Despite the increasing risks of social unrest, 47% of respondents support transparency under the crisis, and 53% disagree with the statement. *Q5* and *Q4* reveal that respondents favor information transparency under the crisis, a key component of democratic accountability, yet reject the idea of imitating western-style freedom. Although the term "indiscriminate" in *Q5* leads respondents to disagree with the statement, the underlying preferences should be consistent, suggesting that people who support information transparency should also promote free speech or vice versa, despite the cost of bringing social disorder.

Q9 is an exemplary question measuring nationalism, and around 65% of responses deem national unity as the highest interest of the society. Turning to the perception of the free market, *Q30* reflects that 57% of respondents regard the state as the primary player to improve the lives of low-income people.

The study manually creates synthetic indices based on a group of questions revolving around the same theme. Building on previous studies, the survey questions are classified into seven themes: *Democracy*, *Freedom*, *Market*, *Socialism*, *Globalization*, *Traditionalism*, and *Nationalism*. Each question is labeled as +1 to represent the support and -1 to represent the non-support of one specific theme. Each index is the combination of such values. The mean of each index is 0.128 for *Democracy*, -0.525 for *Freedom*,

Figure 2: Distribution of Each Index



-0.506 for *Market*, 1,556 for *Socialism*, -1.057 for *Globalization*, 0.755 for *Traditionalism*, and 1.007 for *Nationalism*. These metrics portray a fictitious average respondent in China: one slightly favors democracy, looks a little bit conservative, believes socialism and traditional Chinese culture, prioritizes nationalism, and is suspicious of the free market and economic globalization.

3. Methodology

The four-point scale without the neutral option employed by the survey brings more information than dichotomous (agree-disagree) choices. The study transforms the four-point scale to a binary scale that 1 represents the agreement and 0 represents the disagreement about the statements to facilitate further analysis. Besides, the study turns categorical variables measuring the demography into dummy variables. The dataset now includes 50 binary-scale survey variables, 17 dummy variables (two about *gender*, five about *income*, four about *education*, and six about the *region*), and one continuous variable (*age*).

Inference Methods

To understand which and how features matter, the study firstly runs randomized searches over logistic regression with different λ. In other words, we will compare Ridge ($\mathcal{L}$-2 regularization), Lasso ($\mathcal{L}$-1 regularization), and standard logistic regressions and find the one with the best accuracy score. All the model evaluations are under the train-test split. The study will also run the standard logistic regressions on total data to provide a glimpse of the baseline situation.

Ridge and Lasso regression are shrinkage methods to impose a penalty on coefficients. Unlike the Lasso regression that introduces a constant factor, λ, and truncating at 0, Ridge regression does a proportional shrinkage and enables us to see how much a feature matters. In the study of ideological attitudes, Zmigrod et al. (2020) use ridge regression to test whether there exists an overfitting problem for ordinary linear

regression. The rationale of using Ridge or Lasso is to reduce the multicollinearity. Features revolving around the same theme may correlate with each other, e.g., respondents who support public education (under the free market theme) tend to agree that the state should provide financial support to low-income people, or vice versa. Along with the shrinking process, Lasso selects any one feature among the highly correlated ones and reduced the rest coefficients to zero.

Prediction Techniques

The prediction of outcomes of interest relies upon the Decision Tree (DT), Random Forests (RF), and k-Nearest Neighbors (k-NN). Following the same protocol, the study runs a search over grids of hyper-parameters and displays the best model.

DT is the non-parametric algorithm specialized in predicting a qualitative outcome, e.g., to support information transparency and imitate western-style freedom of speech or not. As survey responses have turned to the binary scale, DT will use the binary splitting to grow the classification tree. Ruger et al. (2004) employed the DT to forecast each Supreme Court Justice's decisions of every argued case during the 2002 term. Their model predicted 75% of the Court's affirm/reverse results and provided a tree graph including the splitting criteria at each node. They revealed that the following criteria could influence former Justice O'Conner's decisions: whether the lower court's decision is liberal, whether the case is from the federal circuit, and whether the respondent is the United States.

The rationale of employing the DT is its closely mirroring the human decision-making process. Applying the DT here means we treat ideology as a series of conditional choices and are interested in predicting the next preference given the prior preferences. The other rationale is interpretability, though DT may fall behind other classifiers in prediction accuracy.

The study turns to RT, an ensemble method that uses bootstrapping and randomly selects subsets of features to train the model. Other than DT and RF, the study makes use of k-Nearest Neighbors (k-NN), the non-parametric algorithm that relies on the distance to the k nearest neighbors to classify whether a new data point belongs specific class membership. In other words, k-NN tries to make an individual fit into a class based on the feature similarity. For DT and k-NN, the study will create the hyper-parameter set to find the best parameters to predict.

4. Results

Parametric Models

After randomized searches over outcome variables with different penalties, as Table 2 shows, training and testing accuracy scores are similar and scattered around the interval between 0.6 and 0.8. Results of indices and survey questions are comparable, and the study will mainly focus on the former because of their interpretability. Then, by using the training data, we re-run the models extracted from the randomized search and understand how each feature matters to the outcome variables.

As Figure 3a shows, one that leans towards democracy, freedom, and the market is more likely to support information transparency under the crisis, while the supporter of nationalism or/and traditionalism is less likely to agree. If one inclines to the socialist ideology, he or she may also agree. Among all the indices,

Table 2: SUMMARY OF PARAMETRIC MODELS

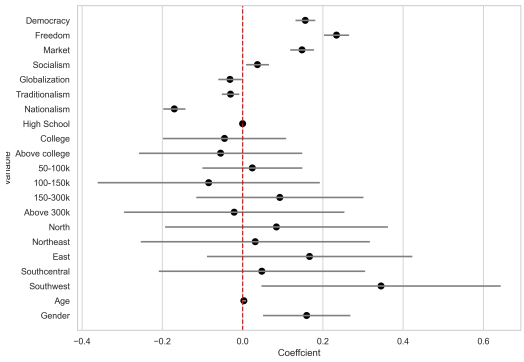
	Outcome variables					
	Q3	Q5	Q3=0 & Q5=0	Q3=0 & Q5=1	Q3=1 & Q5=0	Q3=1 & Q5=1
Index						
Train accuracy	0.720*	0.610*	0.833	0.746	0.750*	0.740
Test accuracy	0.721*	0.610	0.837	0.734	0.757*	0.741
Original						
Train accuracy	0.770*	0.633*	0.833	0.762*	0.767	0.745*
Test accuracy	0.772*	0.654*	0.838	0.758*	0.778	0.742*

Note: *$\mathcal{L}$-1 regularization

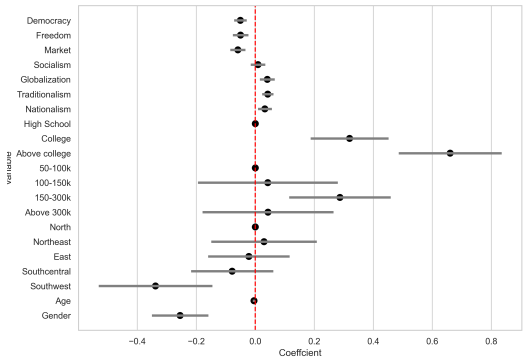
Freedom is the main force driving towards information transparency, and Nationalism is the primary opposite factor. The result is in line with our intuition that freedom and government domination, the usual practice of which is to withhold information, are incompatible. Besides being male, which is positively associated with information transparency, the remaining demographic variables do not display a clear pattern.

Figure 3: Logistic Regression Effects under 95% Confidence Intervals

(a) Q3: *Support Information Transparency under the Crisis=1* (under $\mathcal{L}$-1 Regularization)



(b) Q5: *Not Support Imitating Western-style Freedom=1* (under $\mathcal{L}$-1 Regularization)

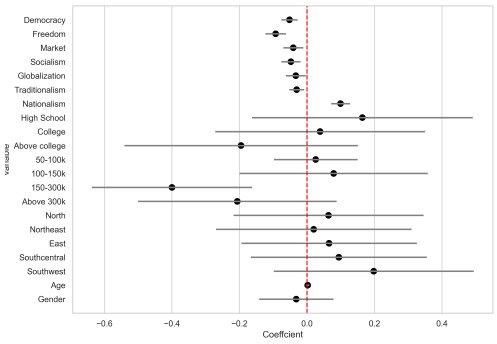


Turning to the statement that (indiscriminately) imitating western-style freedom will lead to social disorder in China, Figure 3b reveals that one supports democracy, freedom, and the market tends to disagree. One who follows traditionalism or nationalism also embraces the idea. Interestingly, respondents with a college education or above compared to those who fail to attend high school are more likely to agree with the statement. A possible explanation for this is that well-educated people in China may have a deeper understanding and see the flaws of freedom than the laypeople, and thus they tend to reject the idea of **indiscriminate** imitation. The same explanation could also apply to individuals whose income is between 150k and 300k. As for living in southwest China and being male, it is hard to figure out the exact reasons. We may attribute that phenomenon to Lasso's arbitrary feature selections.

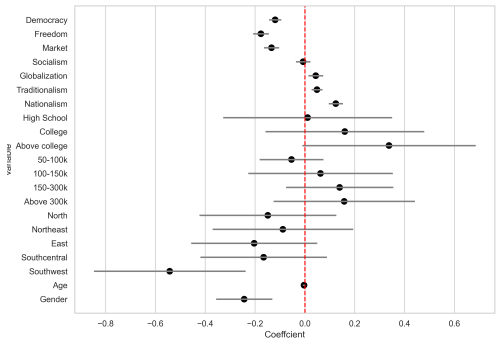
The first set of inconsistent preferences is $\{Q3 = 0, Q5 = 0\}$, representing the group that rejects information transparency yet believes free speech is not harmful to China's stability. As Figure 4a shows,

Figure 4: Multinomial Logistic Regression Coefficients under 95% Confidence Intervals

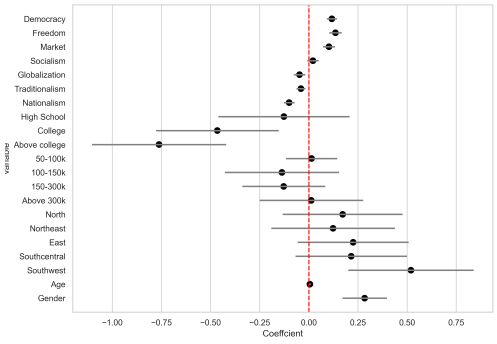
(a) $\{Q3 = 0, Q5 = 0\}$



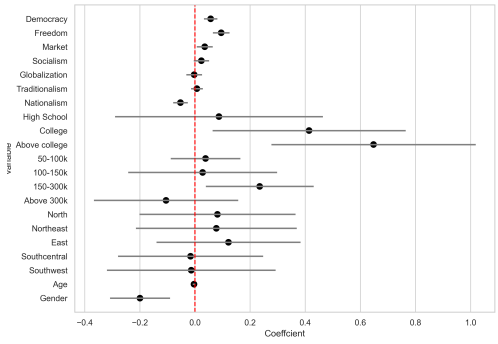
(b) $\{Q3 = 0, Q5 = 1\}$



(c) $\{Q3 = 1, Q5 = 0\}$ (under $\mathcal{L}$-1 Regularization)



(d) $\{Q3 = 1, Q5 = 1\}$



nationalism, being the only feature with a clear direction, is positively associated with such inconsistent preference. Behind the phenomenon, one might suspect that nationalism offsets the other concerns of mimicking western-style free speech and make the group of people over-confident on China's regime resilience. Under the exact opposite situation constituted by the other set of inconsistent preferences, $\{Q3 = 1, Q5 = 1\}$, nationalism is negatively associated with the outcome.

Democracy, freedom, and market align with each other and bring negative effects on $\{Q3 = 0, Q5 = 1\}$ (not support information transparency under the crisis & free speech will lead to social disorder) and positive effects on $\{Q3 = 1, Q5 = 0\}$ (support information transparency under the crisis & free speech will not lead to social disorder). The effects of college education or above in Figure 4c and 4d are opposite, and it appears to confirm the previous speculation that they either can see the flaws of western-style freedom or are cautious about the term "indiscriminately" and then incline to reject the statement.

Similarly, the income group between 150k and 300k is more likely to support information transparency and concerned about social disorder brought by imitating western-style free speech and vice versa. After the re-examination of the composition of the income group, 820 out of 872 (93.9%) respondents have finished college education or above, suggesting education could be the driving force of such preferences.

Non-Parametric Models

Table 3: SUMMARY OF NON-PARAMETRIC MODELS

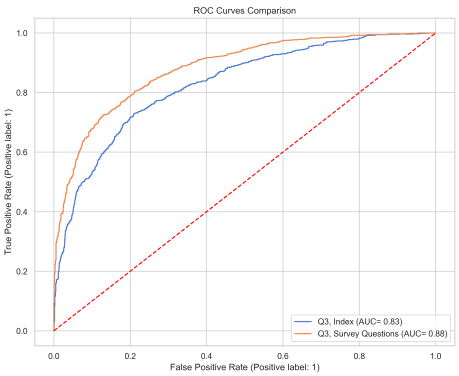
	Outcome variables					
	Q3	Q3 (Index)	Q5	Q5 (Index)	Q3=Q5=0	Q3=Q5=1
Sampling Methods	-	-	-	-	SMOTE	SMOTE
Best Model	RF	RF	RF	RF	RF*	RF*
Max Depth	10	10	10	10	10	10
Number of Trees	750	750	750	750	1000	1000
Training						
Accuracy	0.928	0.852	0.907	0.785	0.932	0.980
AUC	0.978	0.933	0.967	0.900	0.984	0.998
F1	0.931	0.857	0.919	0.830	0.934	0.981
Test						
Accuracy	0.795	0.752	0.696	0.654	0.866	0.852
AUC	0.882	0.828	0.758	0.698	0.940	0.928
F1	0.803	0.758	0.754	0.732	0.867	0.850

Note: * k-NN gets the highest score in AUC. Yet, considering all metrics, RF is a better one.

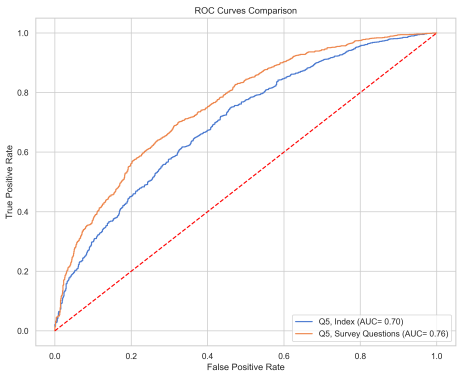
Unsurprisingly, RF, as an ensemble method, outperforms the k-NN and DT. Table 3 displays the accuracy, AUC, and F1 scores for each outcome variable. AUC score measures the true-positive versus false-positive cases, and AUC= 0.5 equals a random guess. F1 is the harmonic mean of recall and precision, and a high score indicates good precision and recall. In terms of accuracy, Q3 and Q5 perform better than the index models and obtain 0.9 levels. However, the discrepancies between training and test accuracy are explicit, suggesting that models confront difficulties in generalizing and predicting. AUC and F1 scores follow similar shrinkage when moving from training to test performance.

Figure 5: ROC Curve Comparisons

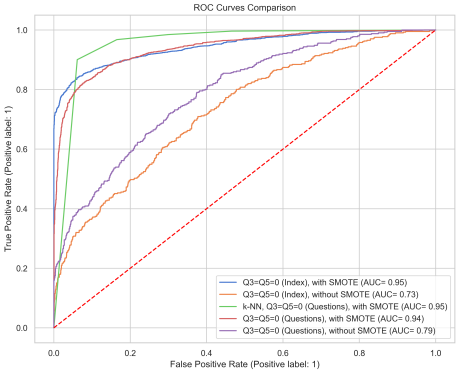
(a) Q3



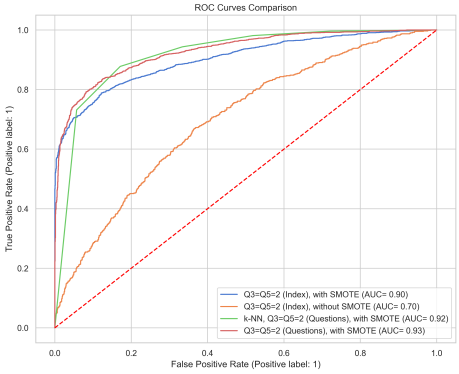
(b) Q5



(c) $\{Q3 = 0, Q5 = 0\}$



(d) $\{Q3 = 1, Q5 = 1\}$



The sets of inconsistent preferences $\{Q3 = 0, Q5 = 0\}$ and $\{Q3 = 1, Q5 = 1\}$, have moderate degree of class imbalances.[1] Class imbalances make accuracy metrics lose efficacy. As displayed in Figure 7e and 7g, the true positives cases are 9 and 19 out of 2500 test data samples, while true negatives are 2094 and 1866. Although we can obtain a high accuracy score, the model does not help predict the outcomes of interest. Therefore, we use the Synthetic Minority Oversampling Technique (SMOTE) to avoid this problem. As Table 3 displays, all training metrics appear to be excellent, and there is no huge gap between training and test metrics. Figure 5c and 5d present the AUC score differences between with and without SMOTE, and the gaps are at 0.2 levels, implying a big leap in true versus false positives performance. Likewise, true positives in Figure 7f and 7h are comparable to false positives.

The examination of feature importance helps us understand which and how much a feature matters. Figure 8a reveals that *Q6*, *Q12*, and *Q2* are the top three critical features for Q3. Furthermore, we employ Individual Conditional Expectation (ICE) plots to illustrate how the features mentioned above influence the target outcomes. Figure 9 tells that an individual who supports universities to recruit students by themselves, accepts besmirching leader's images in artistic or literary work, and prioritizes the universal human rights over sovereignty tends to favor information transparency. Respondents initially incline to information transparency with the increase of age, and then the likelihood starts to drop at the mid-40s. The top features for *Q5* include *Q2*, *Q44*, and Q16. Based on ICE plots, respondents who prioritize universal human rights and value mass art tend to disagree with the statement that free speech will lead to social disorder. Supporter of using force to reunify Taiwan under appropriate conditions believe that imitating western freedom is the root of social disorder.

In terms of inconsistent preferences, $\{Q3 = 0, Q5 = 0\}$, *Q6*, *Q17*, *Q48*, and *Q47* are the top features, and they all exert negative influences on the outcomes, as shown in Figure 9. The fresh features include *Q17* (fostering procedural justice), *Q48* (unnecessary to push forward the simplification of Chinese characters), and *Q47* (the superiority of traditional Chinese Medicine). *Q17* leans towards democracy, and the other two lean toward more traditionalist. The possible explanation for that they are the key factors that can readily anchor one's preferences in specific ideologies, like pro-democracy or anti-nationalism. $\{Q3 = 1, Q5 = 1\}$'s top features include *Q10*, *Q27*, *Q1*, and *Q31*. Likewise, all features decrease the likelihood of being inconsistent. What is different is that the key themes here are democracy, market, and socialism. The recurring themes indicate that democracy, market, and nationalism are the keys to understanding today's ideological spectrum.

5. Discussion and Conclusion

After a close look at China's ideological spectrum, government transparency, and freedom, the study is subject to several limitations. The study intends to understand inconsistent preferences, yet it fails to disentangle the issue. We have found that one step further on democracy, market, and nationalism can decrease the likelihood of having inconsistent preferences. These three themes are consistent with other scholars' findings. To some extent, the study confirms ideology being "a system of beliefs" since we do not find any specific question that leads to more inconsistencies.

The other limitation is about the dataset. The majority of respondents are well-educated or high-income,

[1] The percentage of $\{Q3 = 0, Q5 = 0\}$ and $\{Q3 = 1, Q5 = 1\}$ are 16.6% and 25.4% respectively.

and we might fail to understand the underprivileged's ideology– whether they lean toward democracy and freedom or nationalism. The data from 2014 makes us unable to grasp what happened after Xi Jinping assumed office. Lots of scholars believe that China has been experiencing political regressions and rising nationalism. However, our data may never tell that story.

References

Cantoni, D., Chen, Y., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2017). Curriculum and ideology. *Journal of Political Economy*, *125*(2), 338–392.

Chinese political compass. (2018). http://yiqingxu.org/papers/english/2018_Pan_Xu_ideology/replication.zip

Converse, P. (1964). The nature of belief systems in mass publics. In D. Apter (Ed.), *Ideology and discontent* (pp. 206–261). Glencoe, IL: Free Press of Glencoe.

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in china allows government criticism but silences collective expression. *American political science Review*, 326–343.

Leader Maynard, J., & Mildenberger, M. (2018). Convergence and divergence in the study of ideology: A critical review. *British Journal of Political Science*, *48*(2).

National Bureau of Statistics of China. (2015). *Statistical communiqué of the people's republic of china on the 2014 national economic and social development*. Retrieved April 12, 2021, from http://www.stats.gov.cn/english/PressRelease/201502/t20150228_687439.html

Pan, J., & Xu, Y. (2018). China's ideological spectrum. *The Journal of Politics*, *80*(1), 254–273.

Pan, J., & Xu, Y. (2020). Gauging preference stability and ideological constraint under authoritarian rule. *Available at SSRN*. https://ssrn.com/abstract=3679076

Ruger, T. W., Kim, P. T., Martin, A. D., & Quinn, K. M. (2004). The supreme court forecasting project: Legal and political science approaches to predicting supreme court decisionmaking. *Columbia Law Review*, 1150–1210.

Wu, J., & Meng, T. (2017). The nature of ideology in urban china. *21st Century China Center Research Paper*, (2017-08).

Zhao, D. (2009). The mandate of heaven and performance legitimation in historical and contemporary china. *American Behavioral Scientist*, *53*(3), 416–433.

Contents

A Survey Questions

Table 4: Survey Questions Grouped by Themes

Question	Sign	Mean	SD	(1)	(2)	(3)	(4)
Political Institutions (+1: Pro-democracy, -1: Anti-democracy)							
#1 People should not have universal suffrage if they have not been educated about democracy.	+1	2.17	0.86	0.24	0.42	0.28	0.06
#2 Universality of human rights take precedence over sovereignty.	+1	2.66	0.93	0.12	0.30	0.38	0.20
#3 When events that have major repercussions for the safety and security of people occur, the government should freely disseminate information even if information disclosure increases the risks of unrest.	+1	2.62	0.92	0.11	0.36	0.33	0.19
#4 Western multiparty systems are unsuitable for China in its current state.	+1	2.11	0.76	0.20	0.53	0.23	0.04
#5 Indiscriminately imitating (systems of) western-style freedom of speech will lead to social disorder in China.	-1	2.60	0.83	0.09	0.35	0.42	0.13
#10 Even if procedural rules are violated in the process of investigation and evidence gathering, those who have actually committed crimes should be punished.	+1	2.41	0.72	0.11	0.41	0.45	0.03
#12 It is acceptable besmirch the images of national leaders and founding leaders in literary and artistic works.	+1	2.51	0.85	0.13	0.33	0.43	0.11
#13 When laws fail to fully constrain criminal behaviors, people have the right to impose their own punishments for these behaviors.	-1	2.54	0.81	0.10	0.38	0.41	0.11
#14 Media should be allowed to represent the voice of a particular social stratum or interest group.	+1	2.23	0.87	0.24	0.35	0.37	0.05
#17 Lawyers should do their utmost to defend clients even if the client has committed a crime.	+1	2.60	0.96	0.15	0.30	0.35	0.20
Freedom (+1: Pro-freedom, -1: Anti-freedom)							
#6 It is preferable to let universities recruit students by themselves than to have a unified national college entrance examination system.	+1	2.65	0.88	0.09	0.35	0.38	0.18

#7 Religious adherents should be allowed to conduct missionary work in nonreligious spaces.	+1	2.42	0.79	0.13	0.37	0.44	0.06
#8 Primary school, secondary school, and college students should all participate in government organized military training.	+1	2.56	0.87	0.13	0.30	0.44	0.12
#24 Wasting food is an individual freedom.	+1	2.43	0.80	0.11	0.43	0.37	0.08
#44 The fundamental standard to evaluate the value of a work of art is whether it is liked by the masses.	-1	2.58	0.82	0.13	0.25	0.53	0.09
#45 Even with population pressures, the state and the society have no right to interfere in the decision to have a child, or how many children to have.	+1	2.11	0.71	0.18	0.55	0.24	0.02
Market (+1: Pro-market, -1: Pro-state)							
#21 The minimum wage should be set by the state.[2]	-0.5	2.47	0.85	0.11	0.44	0.32	0.13
#25 If the price of pork is too high, the government should intervene.	-1	2.43	0.86	0.16	0.35	0.40	0.09
#27 Education should be public to the greatest extent.	-1	2.40	0.77	0.11	0.46	0.37	0.07
#29 Attempting to control real estate prices will undermine economic development.	+1	2.05	0.71	0.19	0.61	0.17	0.04
#30 The primary means to improve the lives of the low-income people is to give them fiscal subsidies and support.	+1	2.75	0.92	0.08	0.35	0.31	0.26
#37 Individuals should be able to own, buy and sell land.	+1	2.61	0.88	0.09	0.38	0.35	0.18
#40 Natural monopolies that emerge out of market competitions are harmless.	+0.5	2.30	0.76	0.13	0.49	0.33	0.05
Socialism (+1: Pro-socialism, -1: Anti-socialism)							
#22 The fruits of China's economic development since reform and opening up are enjoyed by a small group of people; most people have not received much benefit.	-1	2.58	0.67	0.05	0.37	0.53	0.05
#31 A rich person deserves better medical services.	+1	2.31	0.78	0.12	0.54	0.26	0.08
#32 High income earners should disclose the sources of their income.	-1	2.84	0.91	0.06	0.34	0.32	0.29
#33 People who make money through gains from financial investments contribute less to the society than people make money through labor.	+1	2.71	0.84	0.09	0.26	0.49	0.15
#34 It is better to sell state-owned enterprises to capitalists than to let them go bankrupt.	+1	1.95	0.80	0.31	0.45	0.21	0.03

Globalization (+1: Pro-globalization, -1: Anti-globalization)							
#23 In the decision-making of major (infrastructure) projects, individual interests should give way to social interests.	-1	2.52	0.76	0.11	0.33	0.50	0.06
#26 A high tariff should be imposed on imported goods that are also produced domestically to protect domestic industries.	-1	2.35	0.76	0.14	0.42	0.41	0.04
#28 The interests of state-owned enterprises are part of the national interest.	-1	2.77	0.70	0.03	0.28	0.56	0.12
#35 Sectors related to national security and important to the national economy and people's livelihoods must be controlled by state-owned enterprises.	-1	2.88	0.71	0.04	0.21	0.59	0.16
#39 Foreign capital in China should enjoy the same treatment as national capital.	+1	2.01	0.79	0.28	0.47	0.22	0.03
Traditionalism (+1: Pro-tradition, -1: Anti-tradition)							
#41 Two adults should be free to engage in voluntary sexual behavior regardless of their marital status.	+1	2.20	0.83	0.23	0.38	0.35	0.04
#42 One should not openly comment on the shortcomings of their elders.	+1	2.90	0.84	0.04	0.29	0.40	0.27
#43 The modern Chinese society needs Confucianism.	-1	2.33	0.81	0.12	0.54	0.24	0.10
#46 The Eight Diagrams (Bagua) in The Book of Changes (Zhouyi) can explain many things well.	-1	2.68	0.87	0.09	0.32	0.41	0.18
#47 The perspective of traditional Chinese medicine on human health is superior to that of modern mainstream medical science.	-1	2.75	0.75	0.05	0.29	0.53	0.14
#48 It is unnecessary to push forward the simplification of Chinese characters.	-1	2.21	0.86	0.19	0.48	0.23	0.09
#49 Traditional Chinese classics should be the basic education material for children.	-1	2.81	0.69	0.03	0.27	0.57	0.13
Nationalism (+1: Pro-nationalism, -1: Anti-nationalism)							
#9 National unity and territorial integrity are the highest interest of society.	+1	2.84	0.92	0.08	0.26	0.38	0.27
#15 If it has sufficient state capabilities, China has the right to take any action to defend its national interests.	+1	2.81	0.85	0.06	0.29	0.43	0.22
#16 Force should be used to reunify Taiwan with China if conditions permit.	+1	2.55	0.85	0.09	0.42	0.34	0.15
#18 Chinese citizens should be allowed to hold foreign citizenship.	-1	2.43	0.79	0.08	0.51	0.30	0.11

#19 It is impossible for western countries led by the United States to tolerate the rise of China into a major power.	+1	2.21	0.76	0.13	0.61	0.18	0.08
#20 The state should take measures to train and support athletes so they can win glory for the country in various international competitions.	+1	3.02	0.84	0.06	0.17	0.47	0.31

[2] As most countries have set the minimum wage, disagreeing on this statement gives a smaller weight compared to other statements.

B Parametric Models

B.1 Correlation Matrix

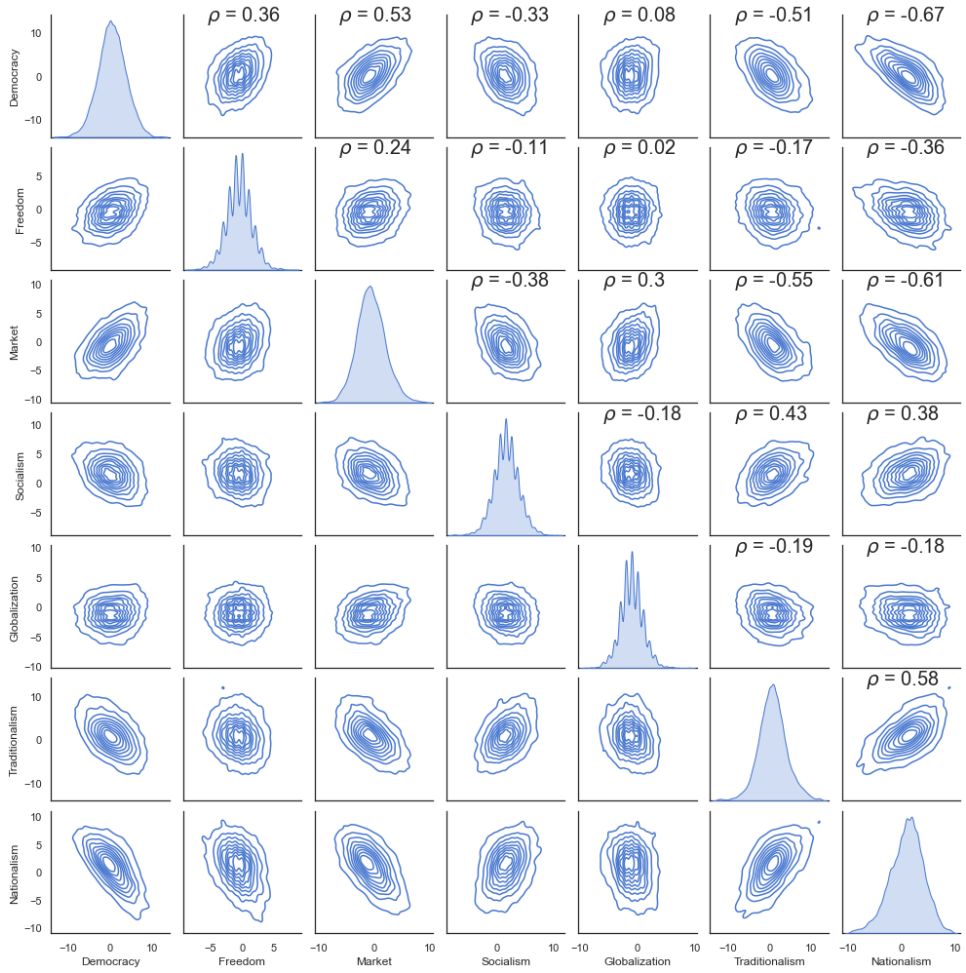


Figure 6: Distributions of Each Index and Their Correlations

B.2 Regression Coefficients

Table 5: Logistic Regression Coefficients (Index)

	Dependent variable:					
	Q3	Q5	Q3=Q5=0	Q3=0, Q5=1	Q3=1, Q5=0	Q3=Q5=1
	(1)	(2)	(3)	(4)	(5)	(6)
Democracy	0.158***	-0.048***	-0.052***	-0.121***	0.113***	0.063***
	(0.011)	(0.009)	(0.012)	(0.011)	(0.011)	(0.011)
Freedom	0.227***	-0.052***	-0.093***	-0.167***	0.142***	0.084***
	(0.014)	(0.012)	(0.016)	(0.014)	(0.014)	(0.013)
Market	0.141***	-0.057***	-0.041***	-0.125***	0.101***	0.032**
	(0.013)	(0.011)	(0.015)	(0.013)	(0.014)	(0.013)
Socialism	0.056***	0.006	-0.048***	-0.024*	0.028**	0.035***
	(0.013)	(0.011)	(0.015)	(0.013)	(0.013)	(0.012)
Globalization	-0.027**	0.049***	-0.033**	0.047***	-0.050***	0.005
	(0.013)	(0.011)	(0.015)	(0.013)	(0.013)	(0.013)
Traditionalism	-0.037***	0.047***	-0.031***	0.056***	-0.044***	0.003
	(0.010)	(0.008)	(0.011)	(0.010)	(0.010)	(0.010)
Nationalism	-0.178***	0.036***	0.099***	0.129***	-0.108***	-0.054***
	(0.012)	(0.010)	(0.014)	(0.013)	(0.012)	(0.012)
Education						
High School	-0.089	0.084	0.164	-0.001	-0.202	0.187
	(0.148)	(0.130)	(0.167)	(0.150)	(0.148)	(0.169)
College	-0.169	0.405***	0.039	0.184	-0.535***	0.460***
	(0.139)	(0.122)	(0.159)	(0.142)	(0.138)	(0.158)
Above college	-0.141	0.708***	-0.195	0.298*	-0.793***	0.708***
	(0.151)	(0.133)	(0.177)	(0.155)	(0.152)	(0.167)
Income (ref. 0-50k)						
50-100k	-0.021	-0.026	0.026	-0.016	0.012	-0.013
	(0.055)	(0.049)	(0.063)	(0.056)	(0.058)	(0.055)
100-150k	-0.157	0.037	0.079	0.081	-0.123	-0.020
	(0.123)	(0.109)	(0.142)	(0.128)	(0.129)	(0.121)
150-300k	0.039	0.274***	-0.400***	0.177*	-0.127	0.161*
	(0.093)	(0.082)	(0.121)	(0.095)	(0.094)	(0.088)
Above 300k	-0.106	-0.015	-0.207	0.175	0.085	-0.226*
	(0.122)	(0.104)	(0.150)	(0.127)	(0.119)	(0.120)
Region						
North	0.091	-0.113	0.064	-0.137	0.121	0.023
	(0.122)	(0.109)	(0.144)	(0.121)	(0.133)	(0.123)
Northeast	0.132	-0.078	0.020	-0.160	0.087	0.077

	(0.126)	(0.113)	(0.148)	(0.125)	(0.138)	(0.127)
East	0.178	-0.157	0.065	-0.223**	0.187	0.061
	(0.113)	(0.101)	(0.133)	(0.112)	(0.123)	(0.114)
Southcentral	0.083	-0.190*	0.094	-0.164	0.189	-0.064
	(0.114)	(0.102)	(0.133)	(0.112)	(0.124)	(0.115)
Southwest	0.316**	-0.443***	0.198	-0.475***	0.460***	-0.100
	(0.131)	(0.117)	(0.151)	(0.133)	(0.139)	(0.133)
Age	0.002	-0.005**	0.002	-0.003	0.005**	-0.004
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Gender	0.159***	-0.219***	-0.032	-0.212***	0.255***	-0.167***
	(0.048)	(0.042)	(0.056)	(0.050)	(0.050)	(0.048)
Constant	0.325	0.231	-1.950***	-1.032***	-0.990***	-1.302***
	(0.200)	(0.177)	(0.233)	(0.204)	(0.207)	(0.214)
Observations	10,000	10,000	10,000	10,000	10,000	10,000
Log Likelihood	-5,394.705	-6,549.558	-4,295.437	-5,176.536	-5,076.838	-5,438.831
AIC	10,833.410	13,143.120	8,634.874	10,397.070	10,197.680	10,921.660

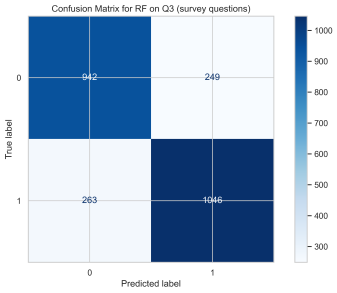
Note: *p<0.1; **p<0.05; ***p<0.01

C Non-Parametric Models

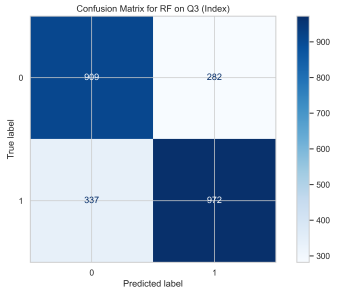
C.1 Confusion Matrix

Figure 7: Confusion Matrix

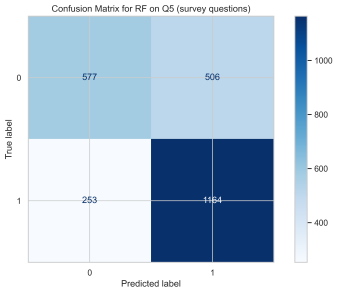
(a) Q3



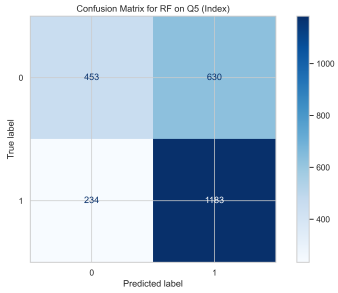
(b) Q3 (Index)



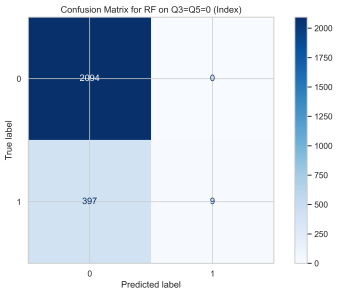
(c) Q5



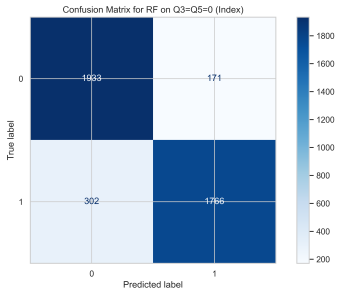
(d) Q5 (Index)



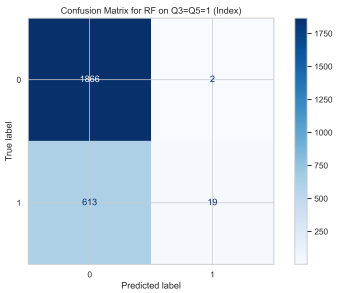
(e) Q3=Q5=0 (Index) without SMOTE



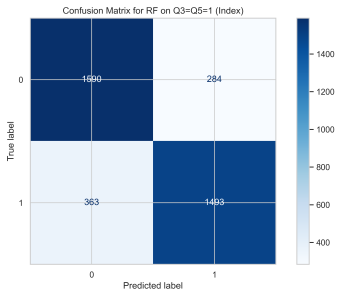
(f) Q3=Q5=0 (Index) with SMOTE



(g) Q3=Q5=1 (Index) without SMOTE



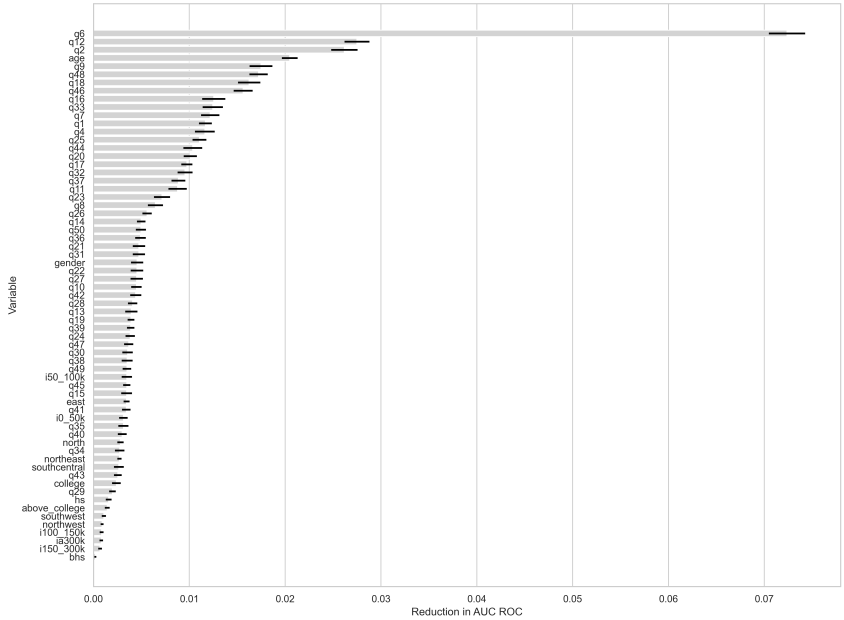
(h) Q3=Q5=2 (Index) with SMOTE



C.2 Permutation Importance

Figure 8: Permutation Importance for Q3 and Q5

(a) Q3



(b) Q5

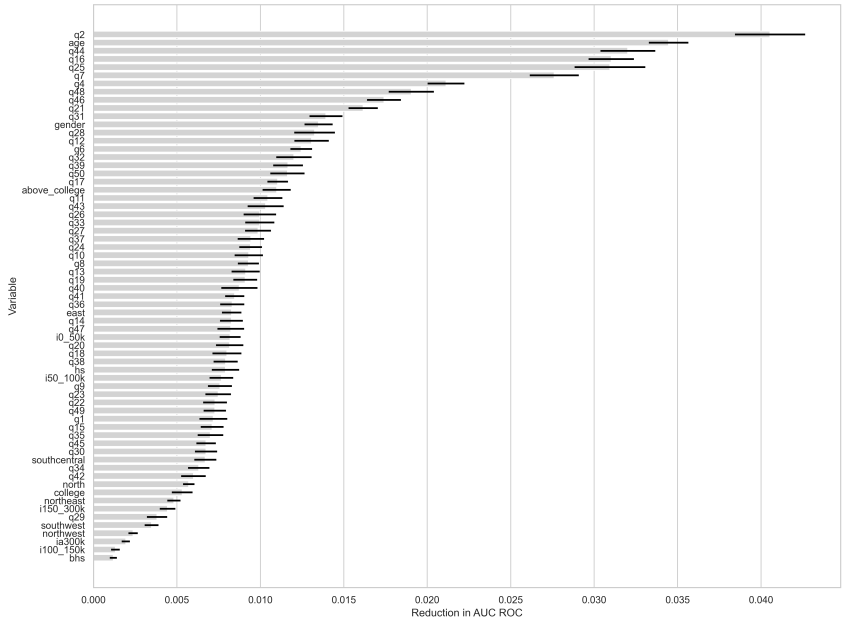
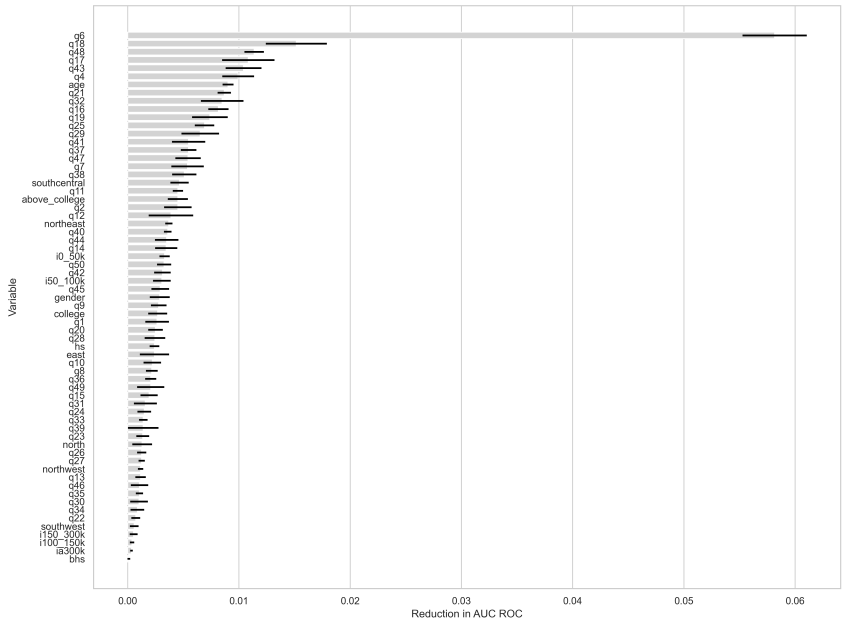
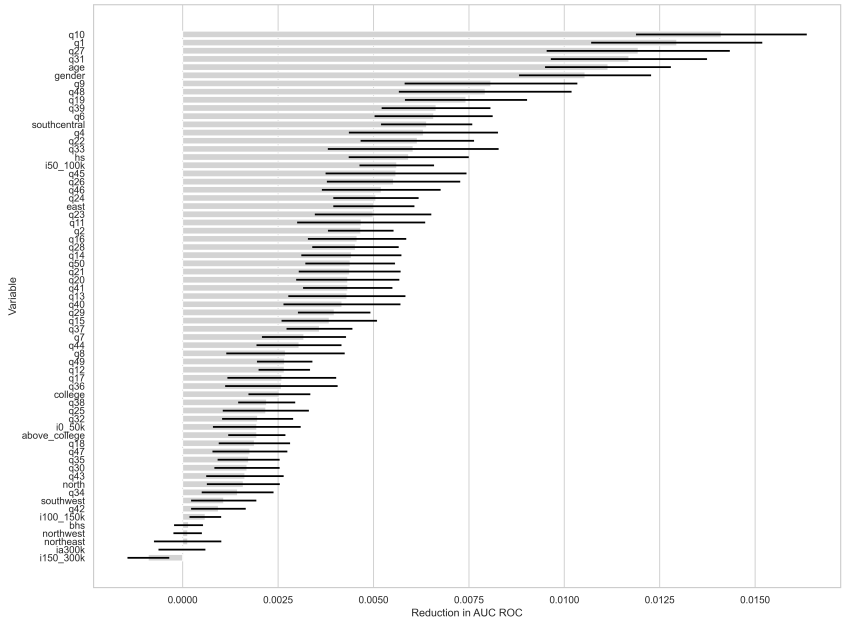


Figure 9: Permutation Importance for Inconsistent Preferences

(a) $\{Q1 = 0, Q2 = 0\}$



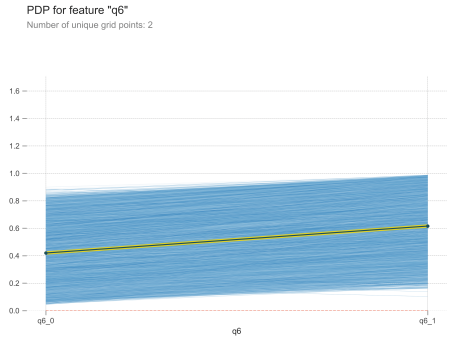
(b) $\{Q1 = 1, Q2 = 1\}$



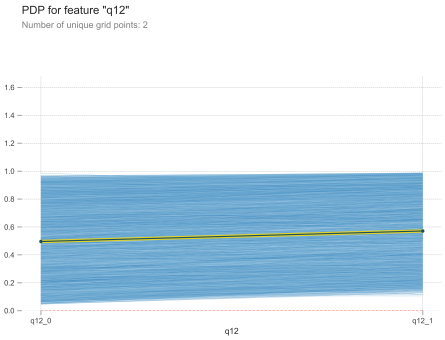
C.3 ICE Plots

Figure 10: ICE Plots for Q3

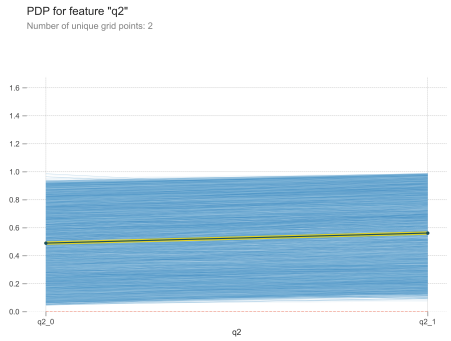
(a) Q6



(b) Q12



(c) Q2



(d) Age

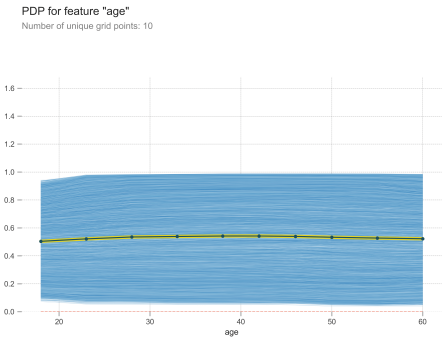
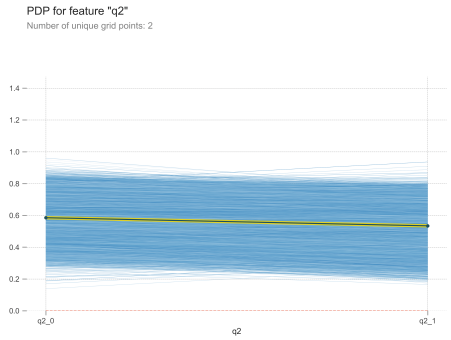
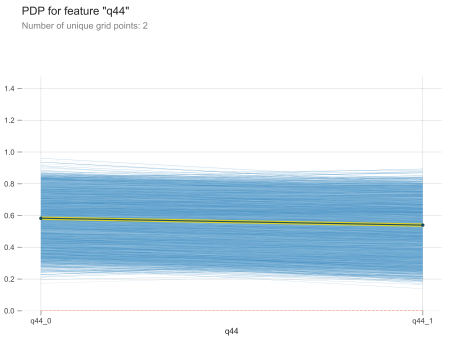


Figure 11: ICE Plots for Q5

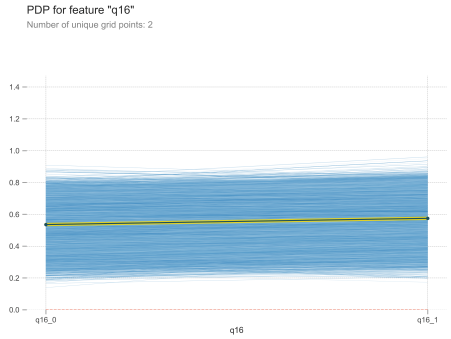
(a) Q2



(b) Q44



(c) Q16



(d) Age

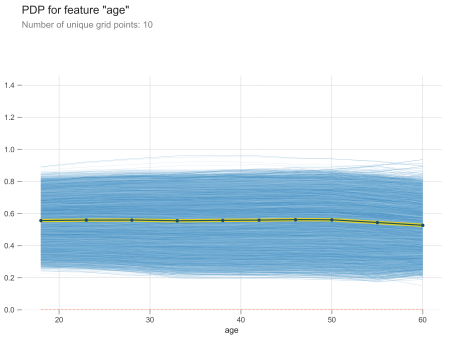
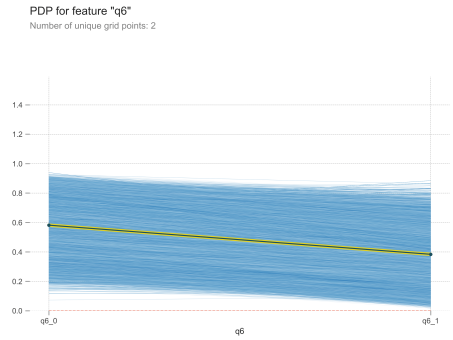
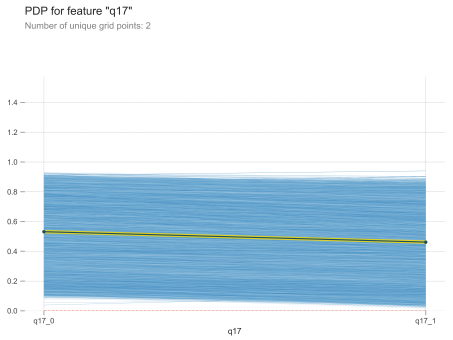


Figure 12: ICE Plots for $\{Q1 = 0, Q2 = 0\}$

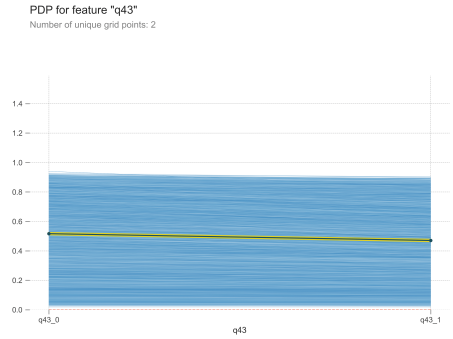
(a) Q6



(b) Q17



(c) Q43



(d) Q48

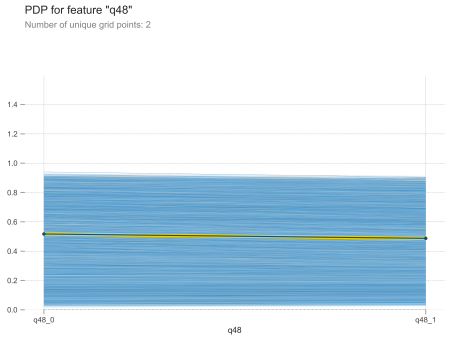
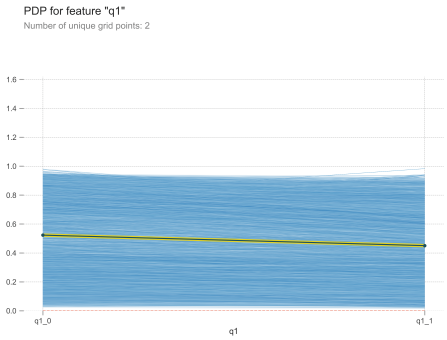
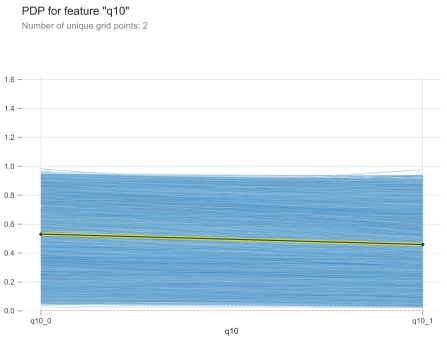


Figure 13: ICE Plots for $\{Q1 = 1, Q2 = 1\}$

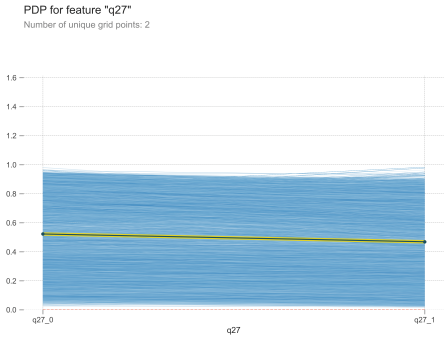
(a) Q1



(b) Q10



(c) Q27



(d) Q31