Ironies of Modernization:

China's Ideological Spectrum Revisited

Replication of Pan and Xu 2018

Jingtian Chen Nora Chen

May 2, 2018

Abstract

The study of ideological cleavages in authoritarian settings contributes to our understanding of the sources of regime legitimacy. In this paper, we replicate Pan and Xu (2018)'s paper on China's ideological spectrum. By making small changes in survey item deletion and sample selection using confirmatory factor analysis, we find an alternative three-dimensional configuration of preferences as China's ideological spectrum, which indicates that nationalism is a core element of regime support in China. Using a hierarchical graded response model, we suggest that modernization has differential effects on political, economic, and social preferences in China, with stronger effects on economic and social preferences than political preferences.

Introduction

In their recent paper published in the *Journal of Politics*, Pan and Xu (2018) use data from a large-scale opt-in online survey [1] (hereafter the *zuobiao* survey) between 2012 to 2014 to study how public preferences are configured and constrained in China. Using principal component analysis (PCA) and exploratory factor analysis (EFA), Pan and Xu (2018) find that public preferences in China are bounded together by some constraint, but the constraint is weaker than that in competitive democracies. Second, Pan and Xu (2018) group the 50 questions in the *zuobiao* survey into seven categories corresponding to the main debates in political, economic, and social domains in China. They test all the possible combinations of the seven categories using confirmatory factor analysis (CFA) and find that the best fit model for ideological configuration in China is a three dimensional one, where the three latent traits are highly correlated: (i) the first dimension reflects views on political institutions and individual freedom; (ii) the second dimension on market economy, capital and labor/social justice, economic sovereignty and traditionalism; and (iii) the third dimension on nationalism. They refer to this configuration of preferences along *political*, *economic/social*, and *nationalism* dimensions as "China's ideological spectrum". Third, they estimate the relationships between the three latent traits and individual, regional-level covariates, suggesting that preferences for more liberal, nontraditional, and non-nationalist are associated with higher education, higher income, and higher levels of regional economic development.

Our replication paper centers on the configuration of public preferences in China. [2] We summarize our methodological changes to Pan and Xu (2018) and key findings as follows. First, by making small changes in survey item deletion and sample selection using CFA, we find an alternative three-dimensional model that fits the data well, indicating that the

[1] zuobiao.me.

[2] For clarification, configuration refers to "the minimum number of coordinates of a space (dimensions) that can best capture divisions and clustering in beliefs and attitudes" (Pan and Xu 2018).

configuration of preferences in China fall along these three dimensions: (i) the first dimension reflects views on political institutions, individual freedom, and nationalism; (ii) the second dimension on market economy, capital and labor/social justice, and economic sovereignty; and (iii) the third dimension on traditional social values. We also stratify respondents based on their education and age group, and find that older and better-educated individuals are more likely to constrain their preferences to less dimensions. Next, to improve the precision in estimating latent traits, we turn to item response theory (IRT) and use a hierarchical graded response model. Rather than an almost identical relationship between three latent traits and individual level covariates in Pan and Xu (2018), we show that individual covariates on education, income and age have a stronger association with pro-market and non-traditional values than with political liberalism.

We are able to replicate all the results in Pan and Xu (2018) and do not take issue with the presentation of results in their paper, and therefore focus to improve the results in Pan and Xu (2018) using controlled methodological experiments and IRT. The rest of the paper is organized as follows. Section 2 introduces the original data, and discuss its limitations as well as the justifications for our methodological changes. Section 3 presents results from exploring the ideological dimensionality and configuration in China. Section 4 demonstrates how latent traits of each dimension are associated with individual-level covariates. The last section concludes.

Data and Methods

The *zuobiao* survey is an opt-in online survey that was initially circulated around online bulletins and most popular among Chinese university students, and attracted 460,532 respondents between 2012 and 2014. It asks 50 questions using a 4-point scale (from strongly disagree, disagree, to agree, strongly agree), and does not have missing data. To address the

problem of poor representativeness associated with an opt-in sample, Pan and Xu (2018) construct a 10,000 observation sample, weighted by province (according to the IP address of each respondent), age cohorts, gender, and their interactions based on the urban population of China. [3] All analyses in Pan and Xu (2018) and our paper are conducted on this 10,000 observation sample. [4]

Our methodological changes are motivated by concerns over the use of CFA. Both CFA and IRT are initially used in educational and psychological testing to estimates latent traits - certain characteristics or abilities that are unobservable - and more recently, used in political science (for example, Imai, Lo and Olmsted 2016). Pan and Xu (2018) use CFA to test all possible structures of latent traits (a total of 877 models) and select the model that has the best fitness statistics. There are two problems with this approach.

First, because CFA requires a pre-specified model, Pan and Xu (2018) group survey items into seven substantive categories to reduce the total number of candidate models, which are (1) political institutions, (2) individual freedom, (3) market economy, (4) capital and labor, (5) economic sovereignty and globalization, (6) nationalism, and (7) traditionalism. Their CFA analysis requires two assumptions: "(1) each question is driven by only a single factor, and (2) questions within the same category (as described above) are driven by the same factor". These two assumptions however, are likely to be violated if certain questions in the survey used cannot easily fit into the substantive categories that researchers introduce. We wish to see if the best model produced is susceptible to item-wise changes. In addition, we want to see if levels of ideological constraint are associated with individual covariates.

Second, CFA typically treats ordinal response categories of each item (from 1 to 4 for example) as interval data, which assumes that the different response categories are evenly

[3] Using the 2005 One-Percent Intercensal Population Survey. Pan and Xu (2018) also exclude three provinces that have fewer than 1,000 respondents in the raw data.

[4] The developers of *zuobiao* do not wish to release their raw data to the public now. As a result, we are unable to obtain the raw data of Pan and Xu (2018).

spaced. This makes item response models are a more preferable approach to scaling categorical data (Zhou 2017). Pan and Xu (2018) use the diagonally weighted least squares (DWLS) estimator to estimate CFA models to address this problem (Li 2016). Alternatively, we use IRT to estimate the extent to which each item loads onto a given latent trait and individuals' ideal points on each latent trait. Specifically, we use a hierarchical graded response model, which jointly estimates item parameters and the effects of observed individual covariates (Zhou 2017). This approach is suited for our ordinal data, and can increase statistical efficiency and offer better asymptotic inference compared with the two-step IRT approach which estimates ideal points and conducts subsequent analyses separately. We use a R package, `hIRT` developed by Zhou (2017) in this paper.

Ideological Configuration and Constraint

Configuration: Nationalism and legitimation

As stated above, CFA requires that each item is driven by only one factor. However, some questions are likely to violate this assumption. We particularly disagree with the categorization of question 6 in the *zuobiao* survey, which asks "It is preferable to let universities recruit students by themselves than to have a unified national college entrance examination system". This question is placed under the "individual freedom" category by Pan and Xu (2018), and ultimately belongs to the political dimension in their best fit model (model A in the paper). However, the debate on college admission in China mostly centers on two issues over social equality: (1) the regional disparity in the quota of elite universities admissions in different provinces, which systematically favors more developed regions such as Beijing and Shanghai. And (2) a higher level of socioeconomic inequality associated with the "independent recruitment" of universities, compared with the national college entrance exam, which systematically favors more privileged students (there is a large body of sociological

literature on this issue, for example, see Wu and Li 2017). Either way, we think this is not about individual choices and freedom.

We drop Question 6 from the dataset and run the same CFA analysis. The best-fit result is a three-dimensional model different from Pan and Xu (2018)'s best-fit model. Instead, our three-dimensional model suggests that public preferences in China fall along these three dimensions: (i) political institutions, individual freedom, and nationalism (ii) market economy, economic sovereignty and globalization, capital and labor/social justice (iii) traditionalism. We refer to Pan and Xu (2018)'s three-dimensional model as "the original model", and our three-dimensional model as "our model".

Recognizing the ambiguities in some of the other questions in the *zuobiao* survey, and in order to avoid further assumptions on categorization of survey items, we drop one item at a time to construct 50 datasets and repeat the CFA analysis (using the same DWLS estimator) for each of the dataset. For each 49-item dataset, we conduct a complete search of all 877 models and select the best model in terms of two goodness of fit measures, chi-squared χ^2 and the root mean square error of approximation (RMSEA), as used by Pan and Xu (2018). We find that in 20% of the 49-item datasets (10 of 50) yield a different best-fit model from what Pan and Xu (2018) find. Table 1 summarizes the results from this exercise and each of the questions that lead to such different results when it is dropped. 60% of the changes produce two different four-dimensional models and 40% of the changes result in the three-dimensional model ("our model") we obtain by dropping Question 6. In addition, the model with the best fit yielded by data in year 2014 alone is our model. Like any internet product (most notably, Facebook), the *zuobiao* survey would be more diffused in later years, when respondents entered the survey are less likely to come from privileged backgrounds .

The three latent traits in the original model are highly correlated with one another (Pan and Xu 2018), meaning that those who prefer authoritarianism are more likely to be more nationalistic and prefer state intervention in market. We distinguish our finding from

No.	Question	Result when a question is dropped
Q3	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.	Dimension 1: Political institutions, individual freedom, nationalism; Dimension 2: Market economy, capital & labor/social justice, economic sovereignty; Dimension 3: Traditionalism.
Q6	It is preferable to let universities recruit students by themselves than to have a unified national college entrance examination system.	
Q9	National unity and territorial integrity are the highest interest of society.	
Q15	If it has sufficient state capabilities, China has the right to take any action to defend its national interests.	
Q16	Force should be used to reunify Taiwan with China if conditions permit.	Dimension 1: Political institutions, individual freedom; Dimension 2: Market economy, capital & labor/social justice, economic sovereignty; Dimension 3: Traditionalism; Dimension 4: Nationalism.
Q17	Lawyers should do their utmost to defend clients even if the client has committed a crime.	
Q18	Chinese citizens should be allowed to hold foreign citizenship.	
Q48	It is unnecessary to push forward the simplification of Chinese characters.	
Q35	Sectors related to national security and important to the national economy and people's livelihoods must be controlled by state-owned enterprises.	Dimension 1: Political institutions, individual freedom; Dimension 2: Market economy, economic sovereignty; Dimension 3: Traditionalism, capital & labor/social justice; Dimension 4: Nationalism.
Q39	Foreign capital in China should enjoy the same treatment as national capital.	

Table 1: CFA results from datasets with item-wise deletion

theirs by showing that the support for authoritarian rule and nationalism are driven by the same latent trait. Existing scholarly work on authoritarian durability typically associates legitimacy with performance or recent institutional development (Nathan 2003; Zhao 2009; Gandhi 2008; Svolik 2012; Truex 2016; Manion 2016). Our finding suggests that nationalism is not only a derivative of performance legitimacy, but also core to the legitimation of the CCP (the Chinese Communist Party) regime in China: those who support CCP rule do so not only because of China's economic development over the *recent four decades*, but also because of the CCP's track record for the *past seven decades* in promoting the project of "national rejuvenation" and the regime narrative that the rejuvenation of the Chinese nation

requires a strong centralized regime. The finding that the support for authoritarian rule and nationalism are driven by the same latent trait reflects the success of the regime strategy to link regime legitimacy to "national rejuvenation".

Dimensionality

We also split the 10,000 observation sample into subsamples based on the self-reported education and age, and find that the ideological configuration of the youngest age group (from 18 or older to 24) is best characterized by a four-dimensional model, and the ideological configuration of the best-educated individuals (graduate school or above) is best characterized by a two-dimensional model. The rest of the subsamples are best characterized by three-dimensional models. In the U.S., the more politically sophisticated individuals are more likely to possess a unidimensional ideological spectrum (Lupton, Myers and Thornton 2015). We uncover that in the absence of electoral competition and open political cues from elites, those who are better-educated and older also tend to organize their ideological preferences along less dimensions, namely, more constrained in their thinking about politics. This finding challenges that of Wu and Meng (2017)'s study on China.

Individual-level Variation

In this section, We explore the relationships between ideology and individual covariates using latent traits estimated by IRT, and compare that with the results using CFA. We estimate both the original three-dimensional model using 50 survey items and our three-dimensional model using 49 survey items (dropping Question six). Estimated coefficients for each question using IRT on our model with Question 6 dropped are shown in the appendix.

With regards to individual-level variation, the results from our three sets of analyses show some differences from the result in Pan and Xu (2018). As in Pan and Xu (2018), we see

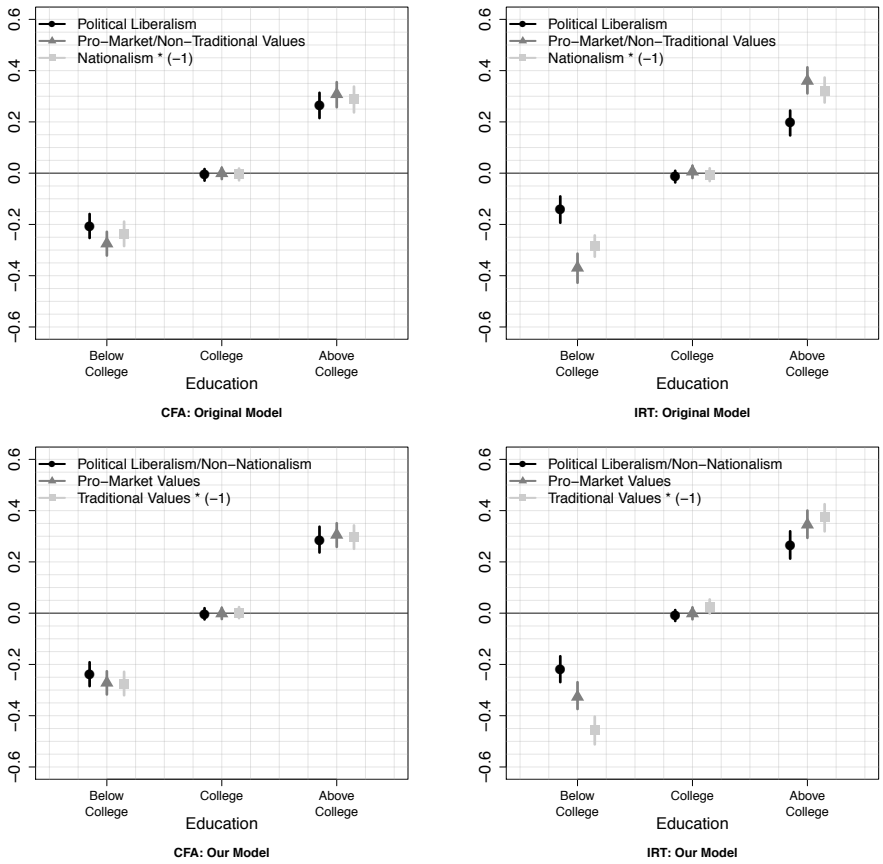

Figure 1: Ideology and education

that both higher education and higher levels of income are associated with politically liberal, pro-market, non-traditional and non-nationalist attitudes. However, when we directly account for the potentially unevenly spaced ordinal response categories and the heterogeneous discriminatory power of different items through IRT, there emerges an interesting set of differences with Pan and Xu (2018)'s result. Under both sets of IRT analyses (with the original model and our model), as shown in Figure 1 and Figure 2, demographic groups divided along social and economic lines exhibit much more homogenous preferences on political issues than non-political issues. Higher education and higher income are much more strongly associated with pro-market economic values and non-traditional social values. Their relationship with politically liberal values, however, is much more muted. This result suggests that rising lev-

els of education and income as a result of the pro-market reforms for the past four decades might have brought about significant changes in social and economic values, but not so much in political values.

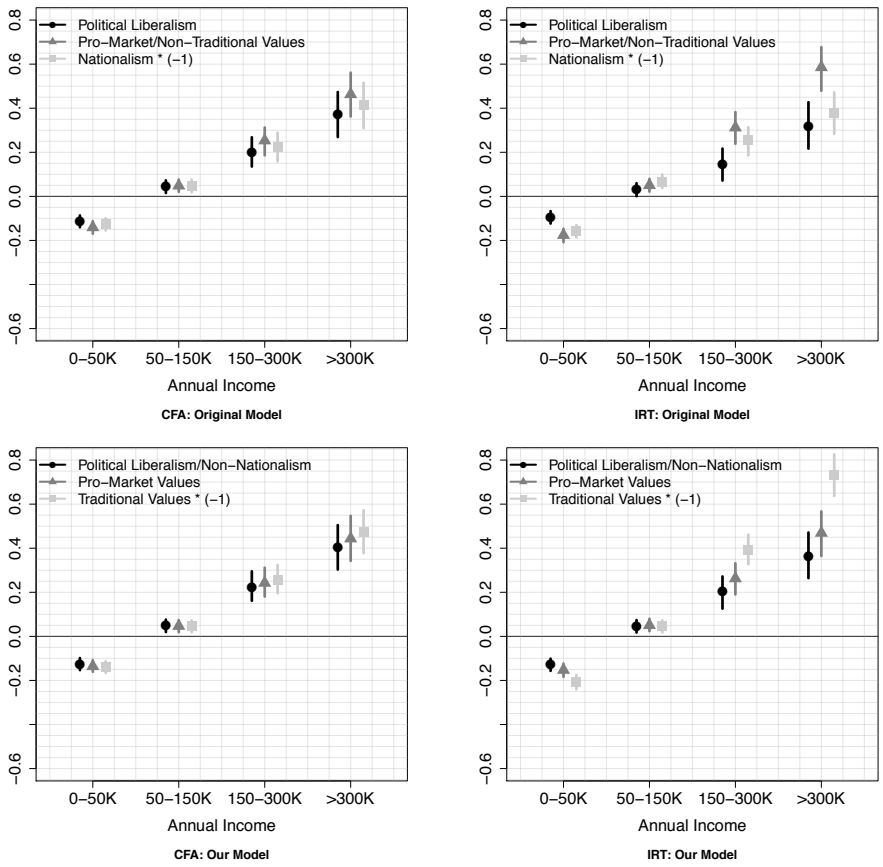

Figure 2: Ideology and income

As for age, the IRT analyses show a clear linear decline in non-traditional values for those with age of 27 or older, a linear decline in pro-market values for those with age of 35 or over, as presented in Figure 3. However, there is no clear linear relationship between age and either non-nationalist views or political liberalism. While it seems that those with age over 48 and below 24 tend to exhibit more nationalist and politically less liberal views, political liberalism and nationalism fluctuates with age for the those that fall between the ages of 24 and 48. As is the case with both education and income, political liberalism

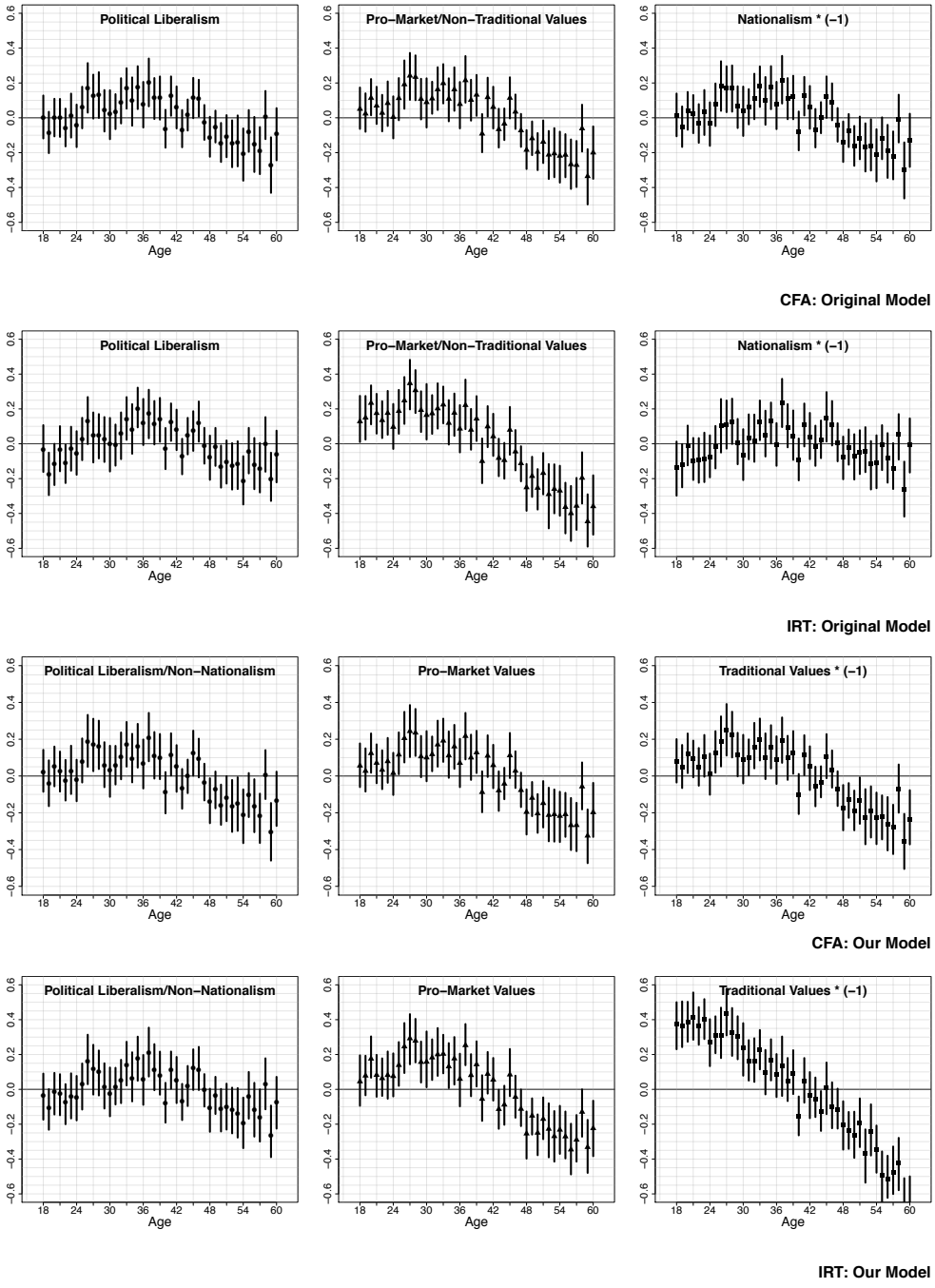

Figure 3: Ideology and age

has a much weaker relationship with age compared to values on other dimensions. While non-nationalism is more strongly correlated with higher education and higher income, its

relationship with age is more ambiguous. In order to understand the explanation for this phenomenon, additional research based on data spanning a longer period of time is needed to disaggregate age effect from cohort effect.

Conclusion

While we are able to replicate Pan and Xu (2018)'s result, we find that the CFA result on the composition of ideological dimensions in China are susceptible to one-item deletion in their dataset. This presents a problem as some survey items are ambiguous and potentially load onto more than one latent traits. We derive an alternative three dimensional model with preferences configured along *political liberalism/non-nationalism*, *economic*, and *social traditionalism* dimensions. Using IRT as an alternative measurement approach, we find that individual covariates associated with modernization theory have a stronger relationship with pro-market values, traditional values, but much weaker relationship with political liberalism.

References

[1] Gandhi, J. 2008. *Political Institutions Under Dictatorship*. Cambridge University Press.

[2] Imai, K., lo, J., and Olmsted, J. 2016."Fast Estimation of Ideal Points with Massive Data." *American Political Science Review* 110(4): 631-656.

[3] Lupton, R. N., Myers, W. M., and Thornton, J. R. 2015. "Political Sophistication and the Dimensionality of Elite and Mass Attitudes, 1980? 2004." *Journal of Politics* 77(2): 368-380.

[4] Manion, M. 2016. *Information for Autocrats: Representation in Chinese Local Congresses*. Cambridge University Press.

[5] Nathan, A. J. 2003. "Authoritarian Resilience." *Journal of Democracy*14(1): 6-17.

[6] Pan, J., and Xu, Y. 2018. "China?s Ideological Spectrum." *Journal of Politics*80(1): 254-273.

[7] Svolik, M. W. 2012. *The Politics of Authoritarian Rule*. Cambridge University Press.

[8] Truex, R. 2016. *Making Autocracy Work: Representation and Responsiveness in Modern China*. Cambridge University Press.

[9] Wu, J. Y., and Meng, T. 2017. "The Nature of Ideology in Urban China." Working Paper.

[10] Wu, X., and Li, Z. 2017. "Independent Freshman Admission Program (IFAP) and Talent Selection in Higher Education in China: Evidence from National Elite Universities." *Chinese Journal of Sociology* 37(5): 139-164.

[11] Zhao, D. 2009. "The Mandate of Heaven and Performance Legitimation in Historical and Contemporary China." *American Behavioral Scientist* 53(3): 416-433.

[12] Zhou, X. "Hierarchical Item Response Models for Analyzing Public Opinion." Working Paper.

Appendix

A.1 Original Categorization and Summary Statistics Please see Pan and Xu's (2018) online appendix at: `http://jenpan.com/jen_pan/ideology_appendix.pdf`.

A.2 Estimated coefficients from IRT on our model Using IRT on our model with Question 6 dropped, the estimated coefficients for each question is as follows:

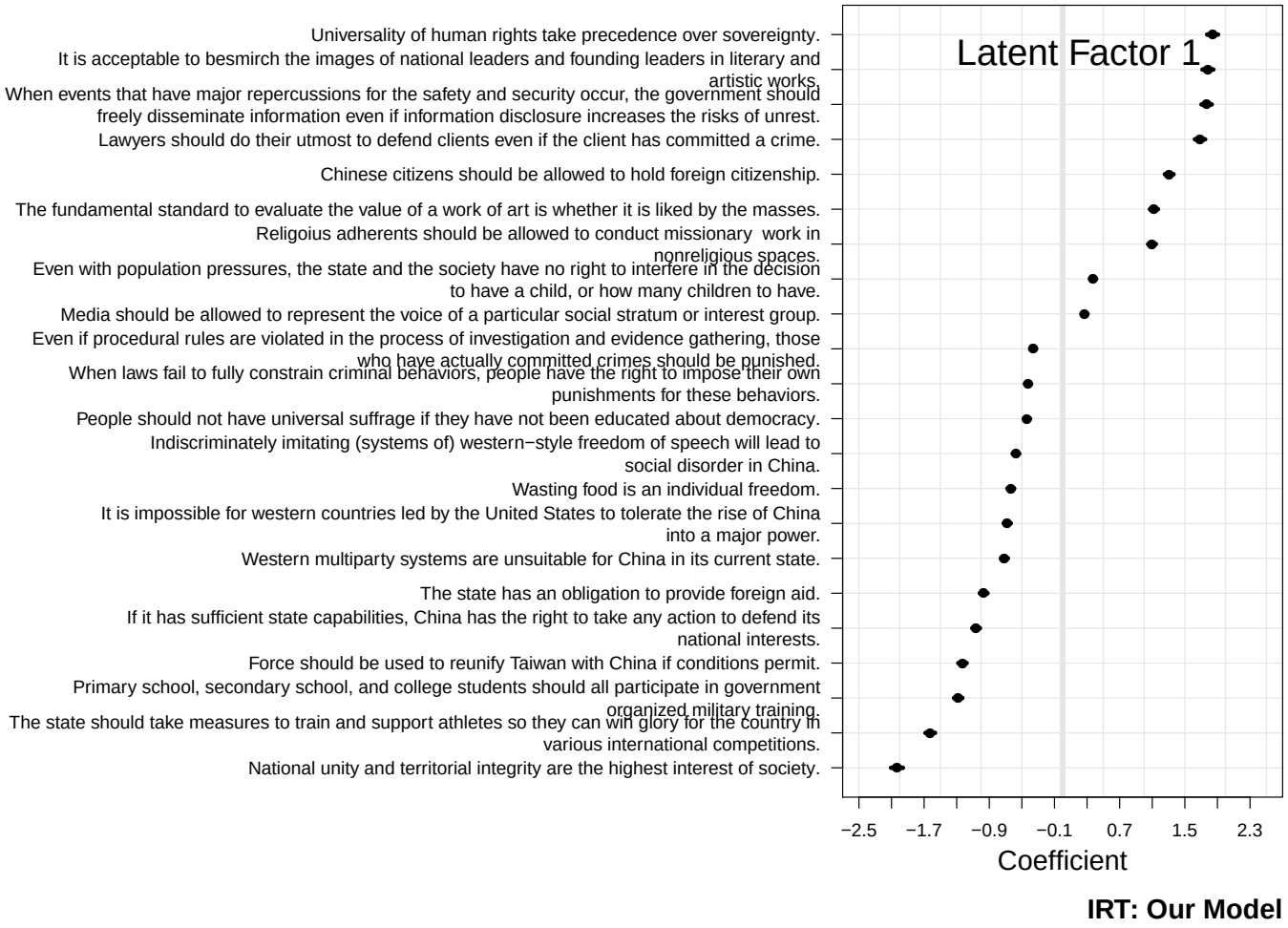

Figure 4: Estimated coefficients: first latent factor

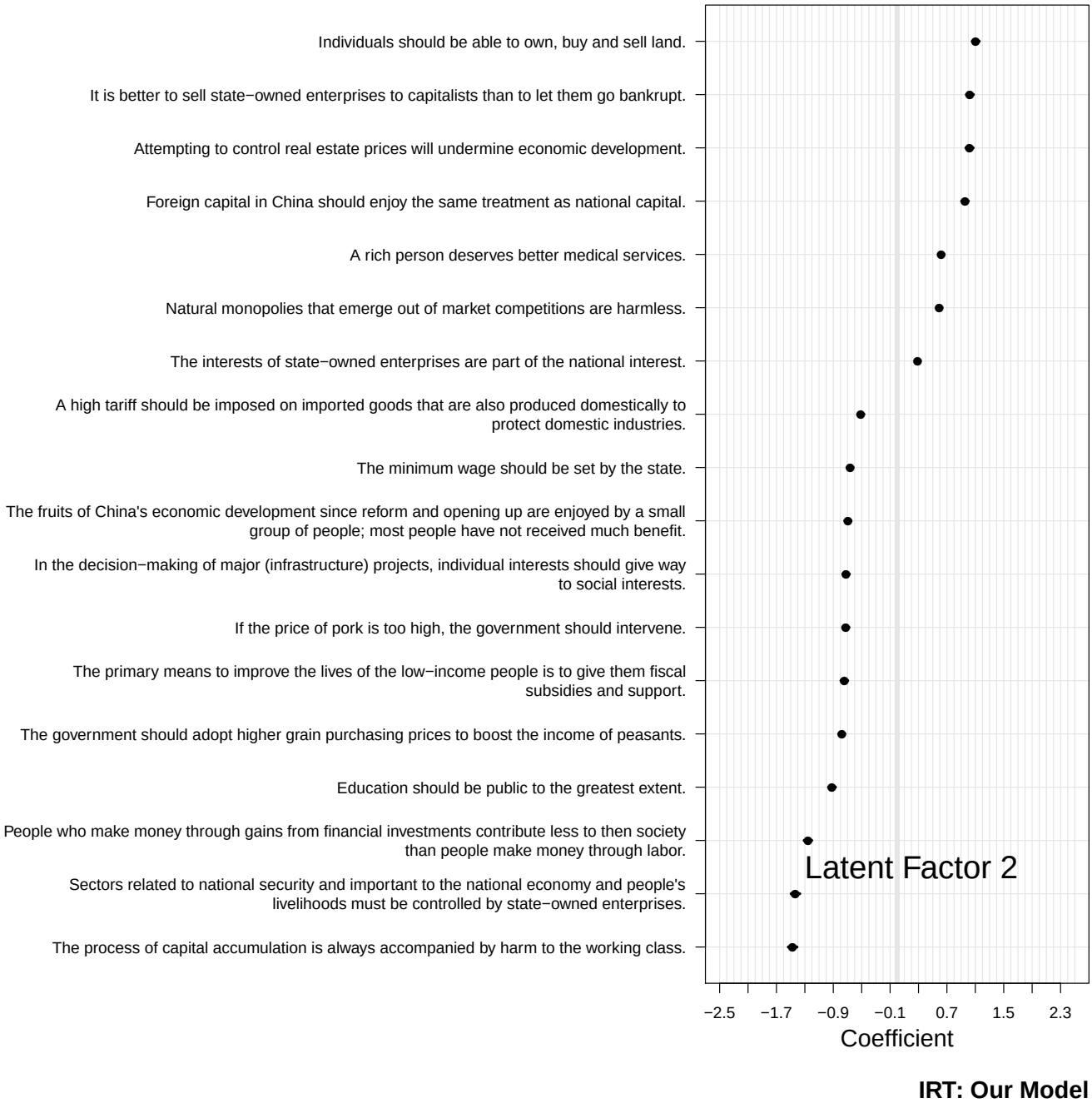

Figure 5: Estimated coefficients: second latent factor

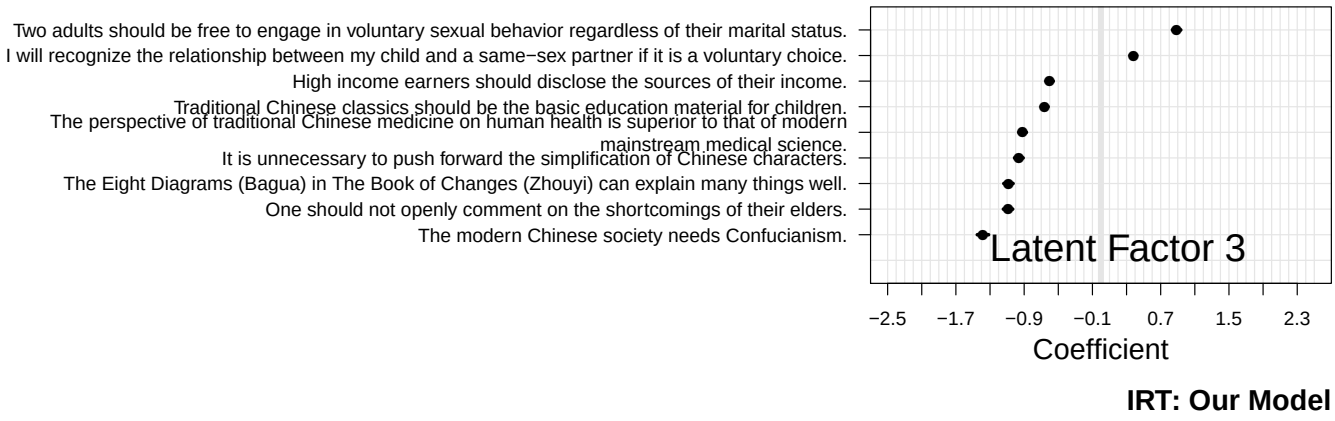

Figure 6: Estimated coefficients: third latent factor