Why have the Chinese become more tolerant of "Zao lian"?

A functional analysis

Yilun Dai

Abstract

"Zao Lian" (having a love relationship before college) has been a taboo in China from 1950s to early 2000s, but the Chinese society's attitude towards "Zao Lian" has been changing since late 2000s. This research examines structural factors that might contribute to this change of attitude using computational tools and time series analysis. I used the number of films and movies with "Zao Lian" elements produced in Mainland China each year as a quantitative indicator of the attitude and examined three functional factors that might contribute to the change in this number, China's annual total fertility rate, female's average age when entering the first marriage by year, and the number of Chinese students studying abroad each year. I then built a VAR model between the number of films and TV series and each one of the three exogenous factors and conducted Granger Causality Test to determine the causality relationship between the number of films and TV series and the exogenous factors. A change in the number of films and TV series Granger-causes a change in TFR and a change in female's mean age at first marriage, while a change in the number of students studying abroad Granger-causes a change in the number of films and TV series. Recursive residuals demonstrated the stability of the causality relationships. The attitude towards "Zao Lian" is adjusted by the policy maker to control TFR and the age at first marriage, while diversifying forms of education contributes to the change of this attitude.

1. Introduction

"Zao Lian"(Chinese word: 早恋), Chinse term of puppy love, has been a unique taboo in Socialist China from 1950s to early 2000s. Although often translated as puppy love, the more

direct translation of "Zao Lian" is early(Zao) love(Lian), or premature love. Early love, in China, is most commonly defined as dating or having a love relationship before college. From late 1970s to early 2000s, parents and teachers spared no efforts in preventing and suppressing early love; some schools suspended or even dismissed young lovers; policy makers formulated punitive policies and issued strict censorship on films and TV series containing themes of early love. In recent years (since late 2000s), however, there is an increasing tolerance on puppy love in the Chinese society. Middle school and high school students are holding hands on campus without worrying about being expelled; TV series with young lovers at school are streaming on major TV networks. Past literatures have examined functional reasons for the establishment of this taboo by analyzing documents and by interviews, and there are also several literatures that recorded this recent change of attitude, but few has explored the reason for the increasing tolerance of "Zao lian". This research will use computational tools and statistical analysis to examine several functional factors that might contribute to this change.

2. Literature Review

Past literatures provided historical and structural explanations for the formation and development of this taboo through qualitative analysis and interviews. Yubin Sheng, Scholar at Georgetown University, in *Too young to date! The origins of zaolian (early love) as a social problem in 20th- century China* (2015), examines the factors that lead to this strong opposition from both "Moral and Ideological" and "Structural/Institutional" perspectives. Shen argues that Chinese opposition to early love stems from at least the early 20th century (pp. 90), and that traditional Confucian culture, the Neo-Confucianism in particular, plays an important part in opposition to early love, since the Neo-Confucianism forbids contacts of single men and women and emphasizes marriage through parents' arrangement (pp. 90). However, Neo-Confucianism,

instead of opposing having a love relationship during certain stage of one's life, rejects the idea of romantic relationship completely, and therefore cannot provide a solid explanation for the Chinese negative attitude towards love relationship of teenagers after 1949s, when forced marriage by parents was forbidden by the law. Sheng's structural and institutional explanations, however, provide some insights on the factors that affect Chinese society's attitude early love. These explanations will be discussed later and will be quantified in my research.

In *Rethinking Sexual Repression in Maoist China: Ideology, Structure and the Ownership of the Body*, an article cited by Shen, Everette Yue Hong Zhang (2005), suggested that administrator in the Maoist era executed "no dating rule" partly because of the employment system. From Maoist era to early post-Maoist era, there was no free labor market, and young graduates from school must obey to the assignments. Couples could be assigned to locations far away from each other; although they were required to obey "revolutionary needs", couples who were separated, in reality, would strive to switch work units (the assigned office) (pp. 8). Therefore, Zhang argued that, the purpose of forbidding early love was to reduce the mobility of labor force between work units and to secure the proper functioning of the work unit employment system (pp. 8). While changes in employment system will not be the main focus of my research, my research will incorporate the idea that economic structure or economic situation would be one of the structural explanations, and will examine the difference in attitudes toward early love nowadays based on different economic situation. Specifically, my examination on different economic situation could involve differences in attitudes in urban and rural areas, visualized by head maps.

Shen's article also examines structural reasons for the Socialist China's antipathy towards China, but Shen primarily focuses on three institutional purposes of preventing early love: the prevention of early marriage, the control of fertility rate, and the exam-heavy education system

(pp. 93). Through a comprehensive literature review, Shen demonstrated that from 1950s to early 1980s, the main concern of early love was early marriage, and such concern not only apply to students but to all the young men and women in the country. The *Marriage Law of 1950* raised the legal age of marriage of men and women to 20 and 18 respectively (pp. 95). Shen argued that the two terms "early love" and "early marriage" was combined until the early 1990s, and that the effort to prevent early love was, in fact, the effort to prevent early marriage, since early marriage was seen as not only harmful for health but also harmful for the nation at that time (pp. 95). The Chinese society became more strongly against early love and early marriage after 1978, when One Child Policy was implemented. Shen stated that the repression of early love was indeed a powerful enforcement of policies that eradicate early marriage and control overwhelming childbirth rate, and that as practices of early marriage and uncontrolled birth rate were gradually diminished, early love became the only social problem among the three "early" practices (pp. 97). Like Zhang's study, this article is also a survey and analysis through past articles and documents. In my research, I will quantify the two structural factors, marriage and childbirth, by practicing statistical analysis on marriage rate and total fertility rate data in China since 2000s, and use computational tools to inspect their relationships with the change of attitude towards early love.

Another institutional factor considered by Shen was the education system and policies on education. The compulsory education, introduced in 1986, required that all children and teenagers of school ages attend to school (pp. 97). Therefore, the opposition against early love became a means of preventing dropouts from schools and of increasing education rate in China. Since then, the target of early love switched from uneducated rural youth to teenagers in schools (pp. 97). A newspaper article from *Shanghai Daily* demonstrating the change of attitude towards early love also focuses on the education. According to Jie Wang, author of *Puppy love no longer taboo*

(2013), parents and teachers in middle school and high school start to view early love positively, arguing that early love can be sweet memories and motivation to work hard towards exam, and that adolescence should be more than just exams and piles of homework. The author implied that the cultural reason for this attitude change is that those parents and teachers had bad experience of being forbidden from early love and being controlled and punished by their parents and teachers, and thus more sympathetic towards the new generation of teenagers, but did not provide further structural reasons for this shift in attitude. However, this article implies that the exam score, while still important, gradually ceases to be the one and only core of education in China. Therefore, part of my research will be treating the influence of change in education system in China, and will use the data of Chinese students studying abroad each year, since studying abroad indicates the diversification in forms of education.

Another study that reflects the changing social norms of early love is *Love and Sexuality of Chinese Teenagers: A Random Sampling Based on Census of populations aged 14-17, 2010*, by Suiming Pan and Yingying Huang (2015). Although this study focuses more on sexuality and physical relationships of teenagers, it uses computational tools such as binary logistic regression, to analyze the influence of difference in urban and rural area, the level education and the presence of sexual education at school (pp. 58, Table 6). The data from this study comes from computer-based surveys that were completed by the participants when no one else was present. While those were not online digital surveys, they were computer-administered surveys, adopted a combination of closed and open questions, and used lying-detecting techniques. This article justifies the importance of using computational tools in sociology researches in love relationships of teenagers. Besides the computational tools, this article also demonstrates a more open attitude towards love relationships and sex among urban teenagers with higher level of education and sex educations in

schools (pp. 58), which motivates me to look into urban-rural difference and education systems when I study the structural reasons that bring about changes in attitude.

3. Data and Methods

3.1.Data used

This research uses data from national and international databases as well as digitally-obtained big data. There are four sets of annual data: Chinese TV and film data, China's total fertility rate data, the mean age of first marriage data, and the number of students studying abroad, all of are from 1995 to 2016.

I used Chinese TV and film data, specifically, the number of films and TV series produced in Mainland China with elements of "early love" (referred to as the number of films and TV series below), as an indicator of public tolerance of "early love". I chose film and TV data because China has a strict censorship on contents of film and TV: according to State Administration of Press, Publication, Radio, Film and Television of the People's Republic of China, the works that are in theatres or on TV should have content consistent with mainstream social norms and values, and the Administration will delete any content it considers inappropriate or ban the film or the TV series with such content. Before the late 2000s, Chinese adaptions of Japanese or Korean films and TV series on early love in high school were remade so that the story took place in college (Smith, 2016). Therefore, the number of films and TV series related to "early love" could reflect the society's attitude towards "early love". The data comes from Douban, a Chinese films and TV database that is similar to Rotten Tomatoes. Douban has over 20,000 films and over 15, 000 TV series on record (data from Zhihu), and is constantly updated with latest films, comments and ratings. Therefore, it possesses two important characteristics of Big Data described by Salganik: big enough and "always on" (i.e., is constantly updated).

An alternative to Douban's data is the first love age data from China Marriage and Love Relationship Survey, a nation-wide digital survey conducted by Peking University and baihe.com (a Chinese match-making website) in 2015. According to the survey, the mean age of first love is decreasing by generation. Those who were born in 70s and 80s have a mean age of first love that is greater than 18 years old, while there is a sharp decrease in the number when it comes to those who were born in the 90s (15.18). There is a further decrease in mean age of first love for those who were born during or after 1995 (12.67). While this part of survey data reveals the decreasing trend of first love age, it is generation wise rather than annual data, and therefore not sufficient for the analysis of this survey.

I will use China's annual Total Fertility Rate (TFR) data from the World Bank in this research. China's Total Fertility Rate had been decreasing sharply from the mid-1980s (2.675 in 1986) to late 1990s (1.494 in 1999) due to the One Child Policy. As shown in Figure 1, Total Fertility Rate witnessed a slight increase in the 21st century, and has been around 1.6 since the late 2000s.

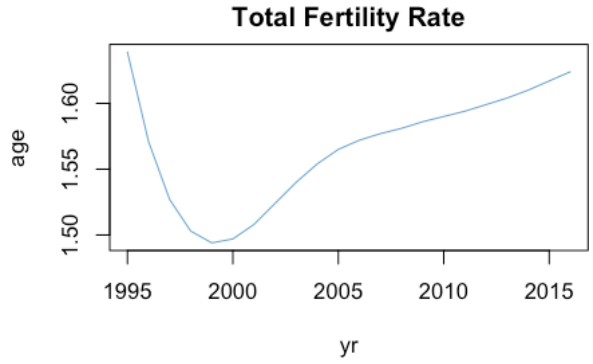

Figure 1: China's TFR from 1995 to 2016

This research uses multiple data sets from the National Bureau of Statistics of the PRC (referred to as the Bureau of Statistics in this paper). The mean age of first marriage from late 1980s to 2010 comes from the sixth China Population Census. While factual marriage (having held wedding and cohabiting without obtaining certificate of marriage) has no legal status since 1994 according to the Marriage Law of China, the sixth Census took factual marriage into account. The legal marriage age in China is 20 for female and 22 for male. However, the census has taken into account of those who get married before 20, since it is likely that at least the majority of this population get married before 1994. The mean age of first marriage data from 2010 to 2016 comes from the annual report of the Bureau of Statistics. Since only female first marriage age is available for this part of data, and male's mean first marriage age and female's mean first marriage age follow the same trend (Trading Economics), this research will use the mean age of first marriage of female. Figure 2 demonstrates that female mean age of first marriage is postponing from 22.85 in 1995 to 26.00 to 2016.

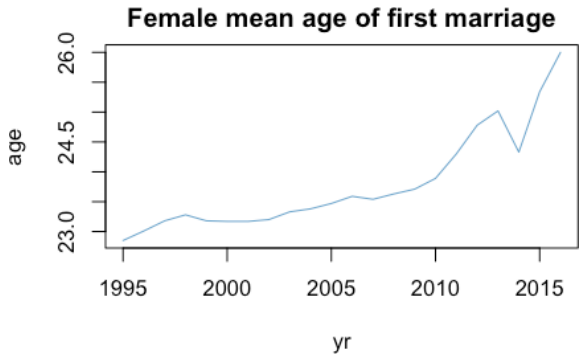

Figure 2: China's female mean age of first marriage from 1995 to 2016

The annual number of students studying abroad also comes from the Bureau of Statistics. Because College Entrance Exam in China is often seen as "sealing the deal" for the life track of teenagers, teenagers are told by their parents and teachers to spare no effort studying and not to get distracted by anything irrelevant to studying for exams. "Early love" is seen as a distraction that will result in poor performance in exams, especially failure in College Entrance Exam. What if College Entrance Exam is no longer the one and only way to high-quality higher education? Will "Early Love" still be regarded as a distraction that impedes teenagers' entrance into undergraduate institutions? Therefore, I will use the annual data of number of students studying abroad. We could see from Figure 3 that the number of students studying abroad fluctuate around 20,000 before 21st century. This number increases to 100,000 at the beginning of 21st century, but it does not begin to increase sharply until 2007.

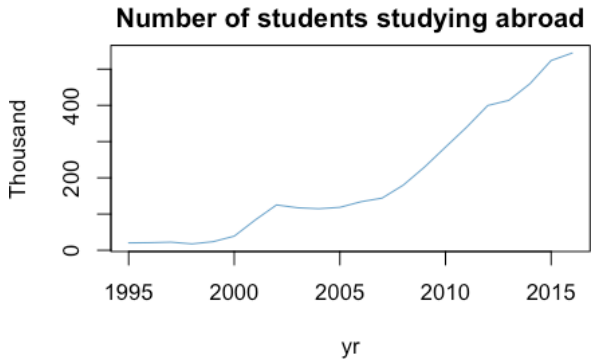

Figure 3: Number of Chinese Students Studying Abroad from 1995 to 2016

3.2 Data Collection

The World Bank and the Bureau of Statistics both provide public access to datasets; I need to conduct web crawling to obtain films and TV data from Douban. Douban has categorized films and TV series by tags, including country of production, genre and theme. Films and TV series with all of the three tags “Mainland China”, “Romance” and “School or Adolescence” are crawled. I used Python’s web crawling tool, Beautiful Soup, to collect each film or TV series’ name and year of production. I then used python to clean the data and obtain the number of films and TV series on teenage romance that are produced in Mainland China each year. Figure 4 is a histogram of number of films and TV series by year.

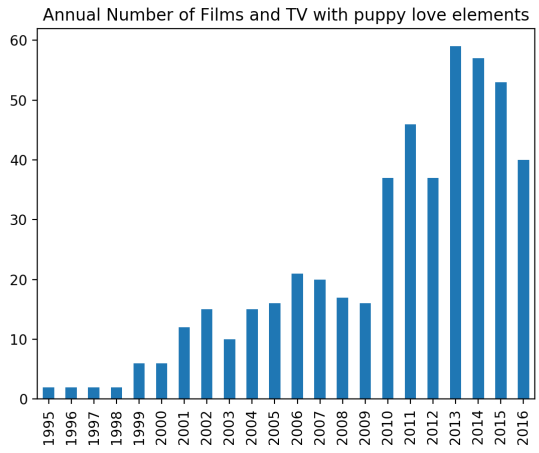

Figure 4: Number of Films and TV series on puppy love produced in China by year

3.3 Methods used and Model Selection

In this research, I will construct time series for the four data sets and analyze possible causal relationships between variables in R Studio. To test the likely relationships between the number of films and TV series and another variable, I will construct a 2-variable vector autoregression model (VAR) between the number of films and TV series and TFR, the number of films and TV series and female's mean age of marriage, and the number of films and TV series and the number of students studying abroad respectively. A two-variable VAR model of order p is a two-variable autoregression in which two equations are estimated; in each equation, the variable on the left-hand side is regressed on p-lags of itself and p-lags of the other variable. For each of the three VAR models I construct, in the first equation, I will regress the other variable on p lags of itself and p lags of the number of films and TV series. In the second equation, I will regress the number of films and TV series on p lags of itself and p lags of the other variable. According to Diebold, one key advantage of a VAR model is that it allows for cross-variable dynamics (pp. 499).

After finishing evaluating VAR models, I will conduct Granger causality test for each of the models. The theory of Granger causality was developed in the 1960s by Clive Granger, and it states that if a variable Y1 "Granger causes" a variable Y2, then the past values of Y1 should improve the prediction of Y2 compared to using the past values of Y2 alone. Granger causality test will reveal the direction of the causal relationship should there exist any.

I will use R's VARselect() function to select the optimal value of p for each model. Taken into account the disadvantage of a more complex model (according to the parsimony principle), I selected the value of lags based on SC. I constructed a VAR (2) model for the number of films and TV series and TFR (equation (1) and (2)), a VAR(3) model for the number of films and TV series

and the female mean first marriage age (equation (3) and (4)), and a VAR(5) model for the number of films and TV series and the number of students studying abroad (equation (5) and (6)).

VAR (2) model:

$$TFR_t = 1.524 * TFR_{t-1} - 0.639 * TFR_{t-2} + 1.052 * 10^{-4} * NumFilmTV_{t-1} + 2.404 * 10^{-4} * NumFilmTV_{t-2} + \varepsilon_t \quad (1)$$

$$NumFilmTV_t = 194.014 * TFR_{t-1} - 97.684 * TFR_{t-2} + 0.606 * NumFilmTV_{t-1} + 0.08967 * NumFilmTV_{t-2} + \varepsilon_t \quad (2)$$

VAR (3) model:

$$age_t = 0.743 * age_{t-1} - 0.675 * age_{t-2} + 0.701 * age_{t-3} + 0.002 * NumFilmTV_{t-1} + 0.034 * NumFilmTV_{t-2} - 0.022 * NumFilmTV_{t-3} + \varepsilon_t \quad (3)$$

$$NumFIlmTV_t = 10.377 * age_{t-1} - 21.713 * age_{t-2} - 18.055 * age_{t-3} + 0.328 * NumFilmTV_{t-1} + 0.129 * NumFilmTV_{t-2} - 0.030 * NumFilmTV_{t-3} + \varepsilon_t \quad (4)$$

VAR (5) model:

$$students_t = 1.584 * students_{t-1} - 0.647 * students_{t-2} - 0.477 * students_{t-3} - 0.437 * students_{t-4} + 0.702 * students_{t-5} - 0.226 * NumFilmTV_{t-1} - 1.487 * NumFilmTV_{t-2} - 1.962 * NumFilmTV_{t-3} - 1.370 * NumFilmTV_{t-4} - 2.037 * NumFilmTV_{t-5} + \varepsilon_t \quad (5)$$

$$NumFilmTV_t = 0.308 * students_{t-1} + 0.009 * students_{t-2} + 0.363 * students_{t-3} - 0.084 * students_{t-4} + 0.375 * students_{t-5} - 1.225 * NumFilmTV_{t-1} - 1.408 * NumFilmTV_{t-2} - 1.696 * NumFilmTV_{t-3} - 0.436 * NumFilmTV_{t-4} - 2.256 * NumFilmTV_{t-5} + \varepsilon_t \quad (6)$$

4. Results

After evaluating the three models, I performed the Granger causality test for each model. Table 1 shows the p-value and the direction of causality of each test. The first row shows the p-value for the first equation in each of the three models (equation (1), (3), and (5)), and the second row shows the second equation in each of the three models (equation (2), (4), and (6)). The smaller p-value of the two indicates the direction of the Granger causality, i.e., past values of the number of films and TV series helps predict the value of TFR at a significant level of less than 0.005; past values of the number of films and TV series helps predict the mean age of first marriage at a significant level of 0.005 to 0.01; past values of the number of students studying abroad helps predict the number of films and TV series relevant to "early love" at a significant level of 0.01 to 0.05.

Equation	*Exogenous Variables*		
	TFR	*Mean Age*	*Number of Students*
Variable ~ lags(NumFilmTV)	$3.623 * 10^{-5}$ ***	0.003387 **	0.4076
NumFilmTV ~ lags(variable)	0.4968	0.1927	0.02798 *

*Table 1: Note: * indicates a causality at a significance level of 0.01 ~ 0.05; ** indicates a causality at a significance level of 0.005 ~ 0.01; *** indicates a causality at a significance level of less than 0.005.*

To visualize the direction of the influence, I plotted the 5-step Impulse-Response Function(IRF) for each of the three models respectively. Figure 5 shows that the number of films and tv series that is 2 lags away has a positive influence on total fertility rate; figure 6 shows that numbers of films and tv series that are 2 lags and 3 lags away have a positive influence on female's average age when entering the first marriage; figure 7 shows that numbers of students that are 3 lags and 5 lags away have a positive influence on the number of films and tv series.

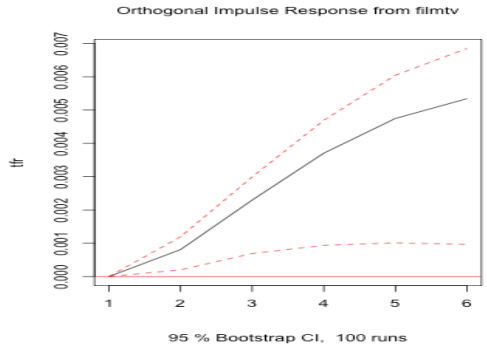

Figure 5: Response of TFR to Impulse from the number of films and TV series

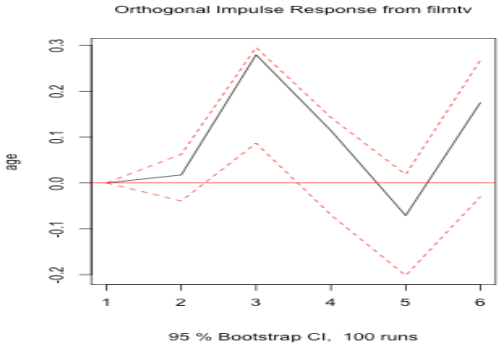

Figure 6: Response of Female's average age at First Marriage to Impulse from the number of films and TV series

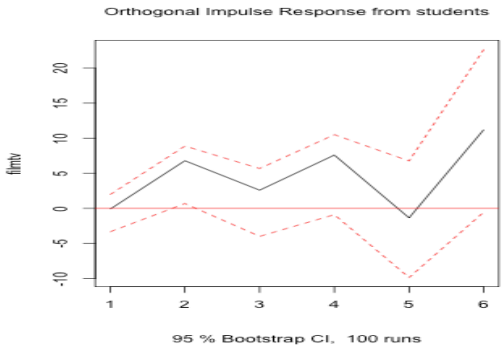

Figure 7: Response of the number of films and tv series to the Impulse from the number of students (in 1,000)

5. Testing

R has a tsCV()function for time series cross validation, but it is not applicable to VAR models with exogenous variables. Therefore, I completed the testing process using R's stability function with the type "Rec CUMSUM". This function is based on recursive residuals and examines if there is any structural change in residuals. Each of the three models has relatively stable and random residuals, which are neither biased upwards or downwards (Figures 5, 6 and 7).

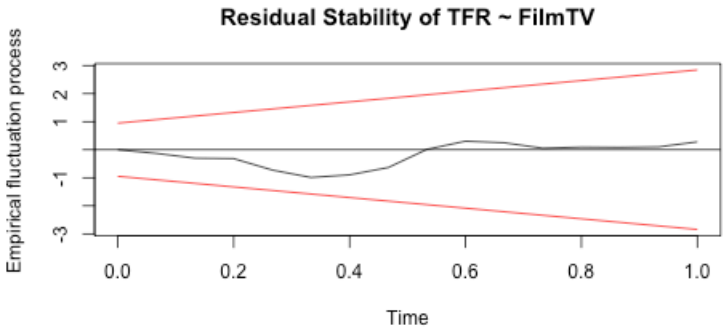

Figure 8: Residual Stability of VAR(2) Eq(1)

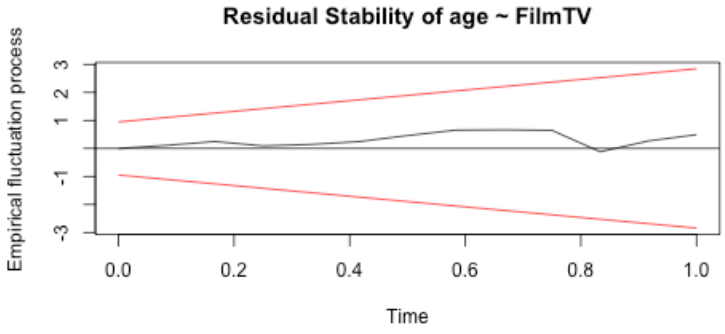

Figure 9: Residual Stability of VAR(3) Eq(3)

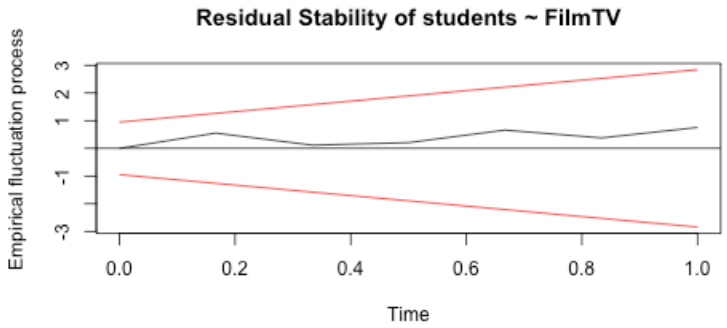

Figure 10: Residual Stability of VAR(5) Eq(6)

6. Conclusion

The Chinese society's attitude towards Zao Lian, specifically, the number of films and TV series with Zao Lian component in previous years, can be used predict China's Fertility Rate and the average female's age of first marriage, contrary to the hypothesis that TFR and the average female's age of first marriage have causal effects on the change in the attitude towards Zao Lian. However, a change in number of students studying abroad has a positive causal effect on the attitude towards Zao Lian. Policy makers are adjusting the society's perception on Zao Lian in order to facilitate new marriage and fertility policies, because average marriage age has been postponing and TFR has been low. On the other hand, an increasing number of Chinese students studying abroad demonstrates diversifying forms of education in China. As ways to obtain high-quality higher education are no longer solely score-based, the purpose to reduce the negative effect of Zao Lian on exam results becomes less significant.

References:

Diebold, F.X. (2015), Forecasting, Department of Economics, University of Pennsylvania, http://www.ssc.upenn.edu/~fdiebold/Textbooks.html

Pan, S., Parish, W., & Huang, Y. (2011). Love and Sex of Chinese Teenagers: A Random Sampling Based on Census of populations aged 14-17. Chinese Youth Studies, 12(7), 55.

Salganik, Matthew J. 2017. Bit by Bit: Social Research in the Digital Age. Princeton, NJ: Princeton University Press. Open review edition.

Shen, Y. (2015). Too young to date! The origins of zaolian (early love) as a social problem in 20th-century China. History of Science, 53(1), 86.

Smith, T. (2016, June 13). A Look At 'Zao Lian' And The Secret Dating Lives Of Teenagers In China [Electronic version]. The Knockturnal.

Wang,J. (2013,August31).Puppy love no longer taboo [Electronic version].Shanghai Daily

Zhang, E. (2005). Rethinking Sexual Repression in Maoist China: Ideology, Structure and the Ownership of the Body. Body and Society, 11(3), 9.