Analysis of Parenting Style in China: A Logistic Regression Approach

Min Ho Kim [a], David Zhang[b]

[a] Department of Economics, University of Toronto, 1003056266

[b] Department of Economics, University of Toronto, 1003793255

1 Executive Summary

The phrase "tiger mom" was first popularized by Amy Chua, a Yale law professor and a stereotypical "tiger mother" who wrote the *Battle Hymn of the Tiger Mother*, the 2011 international bestseller, that argued for the "cultural traits" as the reason behind the success of certain racial groups. While some cited this book as evidence for adopting a "tiger mom" style of parenting, others treated work like this as an "unscholarly romp" that "relies on anecdote, rather than data." (Lundberg, 2015).

Behind the heated controversy generated by this book is the age-old debate of what works best for raising children, and there is a growing interest among economists to understand this topic better. In their seminal paper published in the reverend *Econometrica*, economists Fabrizio Zilibotti and Matthias Doepke constructed a unifying framework that details the mechanism of how different parental preferences lead to different parental decisions. In their paper, the authors documented the striking trend among developed countries that as income inequality increases, the distribution of parenting styles shifts from a more "permissive" type to more "authoritative" and "authoritarian" parenting. That is, growing inequality incentivizes parents to impose more structure on children, thereby forcing children to sacrifice more for long-term success. Inspired by this, our paper seeks to discover whether there exists a relationship between income inequality and parenting style in a developing economy like China.

We constructed our own measure of "traditional" and "liberal" parents, in a similar spirit to the paper by Doepke and Zilibotti. Our categorization differs from theirs in that we only define two types of parenting style, thus allowing for the construction of a binary response variable, which we then use on a logit regression model to examine what factors increase or decrease

the likelihood of a given parent becoming a "traditional" parent. We also constructed our own measure of income inequality. Due to the lack of accurate and publicly available data on local income inequality in China, we calculated county-level Gini coefficients using our data set.

To our dismay, our Gini coefficient estimates are still wildly inaccurate. Our adopted data set is also limited in both size and scope. As a result, we found no significant relationship between inequality and parenting preferences. On the other hand, we found that parental education level and family social status possess strong and significant predictive power on parenting style. Specifically, as the parent has completed more basic education (which in China includes primary and lower secondary), and as the household rises up the local social status ladder, the parent will be more likely to adopt a "liberal" parenting style.

While we failed to uncover our intended results regarding income inequality, we found some very encouraging signs that certain parental and familial traits may be highly indicative of parenting styles. We conclude our paper by calling for future research, both on the relationship between income inequality and parental preferences within China and on further identifying and validating parental and familial characteristics that are important in shaping parenting styles.

2 Introduction

The importance of parenting has long been an integral part of the study of education. In his review article, Spera (2005) concluded that parental involvement and monitoring predict adolescent achievement. It has also been reported that parental practices and involvement differ significantly by parental socioeconomic status. Hart and Risley (1992) famously documented the "Thirty Million Word Gap" that a child receives between high-SES and low-SES households. Time spent on total child care also differs significantly by parental socioeconomic status (Guryan, Hurst, & Kearney, 2008).

That parental involvement plays an important role in shaping a child's development is well-established and uncontroversial, but what constitutes as good parental involvement? In her *notorious* book titled *Battle Hymn of the Tiger Mother*, Chua (2014) popularized the concept of tiger mothers, a jargon used for parents who practice a "traditional, strict 'Chinese' upbringing". This garnered interest from the public and researchers alike in discovering how different parenting styles may impact education outcomes. Parental types were initially a research focus

of the developmental psychology literature, but recently economists have started to take an interest in this area as well (Doepke, Sorrenti, & Zilibotti, 2019). In their seminal paper, Doepke and Zilibotti (2017) developed a model that captures the decision-making of both parents and children. Parents are categorized into three types - *permissive*, *authoritative*, and *authoritarian* - each implying different preferences on child outcomes and parenting effort to be spent. *Authoritative* parents are pushier than *permissive* parents but also less paternalistic than *authoritarian* parents. The optimal parenting decisions are then endogenously determined in their model.

One thing that motivated Doepke and Zilibotti (2017) is the observed decline in the *authoritarian* style of parenting and the positive correlation between inequality and *authoritarian* parenting. Using data from the WVS survey, Doepke and Zilibotti (2017) showed that as income inequality grows, the return to education becomes higher as well, and this positively correlates with the proportions of *authoritative* and *authoritarian* parents while negatively correlates with that of *permissive* parents.

The trend documented in Doepke and Zilibotti (2017) is at the cross-country level, covering surveys from mostly OECD countries. We are therefore interested in discovering whether similar trends can be found in developing countries with different cultural and economic landscapes. In this paper, we used data from the *2010 China Family Panel Studies* (CFPS) to study factors that could potentially predict the distribution of different parent types. Due to differences in the type of questions asked between the WVS surveys and the CFPS survey, we propose a different measure of parental types. Doepke and Zilibotti (2017) categorized parents based on the qualities that they wish their children to possess. In the CFPS survey, parents were asked whether they agree or disagree about certain statements regarding parenting preferences. The answers usually range from 1 to 5 (Strongly Disagree to Strongly Agree), although there are also binary response questions and more open-ended questions. Given this, we construct a score measure that takes the weighted average of parents' responses.

Because we adopted the score measure for parental types, we realized it would be a lot more appropriate to utilize the score measure as a binary variable. We thus categorized parents into two groups by imposing a cutoff on the score measure. Specifically, we assign parents with scores above the cutoff to the "traditional" group and those with scores below the cutoff to the "liberal" group. It should be noted that in our context, "traditional" and "liberal" are defined along the line of whether the parent's preferences and parenting style are closer to

traditional Chinese values in child-rearing. Parents who exhibited a more paternalistic and utilitarian attitude toward their children's upbringing are considered to be "traditional", while parents who tend to allow for more freedom for their children are considered "liberal". As a raw approximation, our definition of "traditional" parents can be perceived as *authoritarian* in Doepke and Zilibotti (2017), while "liberal" parents resemble *permissive* parents. We do not include *authoritative* in our categorization as that would require another dimension in order to be distinguished from *permissive* and *authoritarian*.

We utilize a logit regression model to examine what factors could predict the likelihood of being a "traditional" parent. In particular, we are chiefly interested in whether income inequality possesses any predictive power on parental types, controlling for parental and familial characteristics. Our results on this front are disappointing, likely due to the poor quality of our Gini coefficient estimates and the limitations of our data set. While we did not uncover any significant correlation between income inequality and parental preferences, we found encouraging results from some parental and familial traits that we included as controls. Noticeably, we concluded that the more basic education (elementary and junior high) that the parent has completed, and the higher the family's social status in the local area, the less likely the parent becomes a "traditional" parent.

2.1 Related Literature

Burton, Phipps, and Curtis (2002) presented a theoretical framework very similar to the one in Doepke and Zilibotti (2017), in which both the child and the parent are utility maximizers with the child having a higher discount rate. They concluded by examining the Canadian data that the relationship between "bad parenting" and poor child outcomes is not unidirectional and that programs such as parenting classes may not be as cost-effective as it appears to be. Agostinelli, Doepke, Sorrenti, and Zilibotti (2020) limited the scope of parental influences through the lens of peer and neighbourhood effects. In particular, parents can put restrictions on children's interaction with peers and move children to another neighbourhood.

Recent literature has also paid more attention to the children's influence on parenting style and other measures of child development. Deng and Tong (2020) discovered that respectful parenting affects children's noncognitive ability. Chung, Xiong, and Zhang (2022) examined whether parenting styles are affected by children's learning delays and developmental deficiencies.

3 Data and Methods

3.1 Data

The data is from the *2010 China Family Panel Studies* conducted by Peking University in Beijing, China. It is a biannual survey of the family members of the sampled households. The survey questionnaire span from adult and family background to child information. Questions include, but are not limited to, objective measures such as education level, income, and the number of children, as well as behavioural questions regarding parents' parenting style. There are three data sets that we draw on from the 2010 CFPS database: *adult*, *child*, and *family economic conditions*, each containing about 33,000, 8,900, and 15,000 observations, respectively.

Table 1:

Statistic	N	Mean	St. Dev.	Min	Max
Net Family Income	1221	40826	52809	20	1145000
# of Child	1221	1.352	0.62	1	6
Distance from Nearest School	1221	15.68	20.86	0.016	150
Highest Education Level of Parent	1221	2.86	NA	1	7
Non-Agricultural Business Owner	1221	0.0884	NA	0	1
Local Gini Coefficient	1221	0.5867	0.0749	0.2988	0.7322

Since we constructed our dependent variable with questions from the *child* dataset, we merged variables from the other two datasets into the *child* dataset based on *family id*. The dataset we obtained after merging them together contains $n = 1,221$ observations at the family level. The reasons for the shrinkage in sample size are twofold. On the one hand, we took out observations from the *child* dataset for children who already had a sibling represented in our final dataset in order to avoid sampling repeated parents. On the other hand, the response rate to most questions that we sampled for the dependent variable is below 30 percent. This rendered most observations from the *child* dataset unusable. As a result of the above, there could be a shift in the distribution of our variables of interest after we merged the datasets. Table 1 shows the summary statistics of some of the important variables for our merged dataset. Figure 1 displays the correlation matrix of the relevant variables, along with the distribution of each variable. The distributions are right-skewed with very thin tails, indicating the lack of variance, and hence the possibility of bias present in the data.

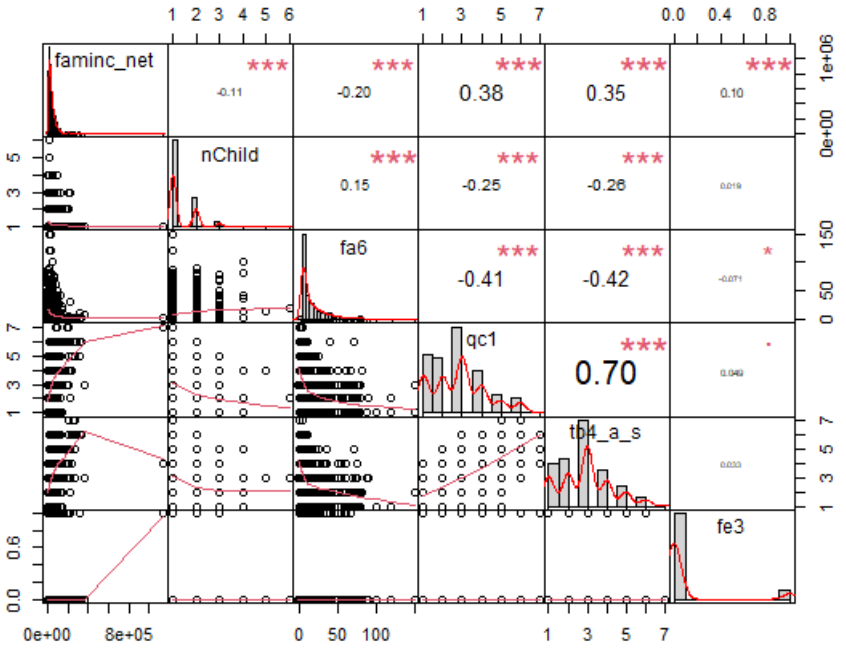

Figure 1: Correlation Matrix

3.2 Empirical Strategy

3.2.1 Measures for Different Parent Types

The main aspect of our paper is to construct a score of how traditional a parent is based on the parents' answers to the behavioural questions regarding their child. For each variable, we decide whether a response falls into the category of "traditional" or "liberal" based on a score measure that we calculated from their answers. We call this measure $score_i \in [0, 1]$. The sum of scores across all variables divided by the number of questions they answer would be their final score. We assign parents whose $score_i$ is larger than or equal to 0.85 to the "traditional" group, and parents with $score_i < 0.85$ to the "liberal" group. We chose the cutoff for $score_i$ to be 0.85 because the distribution of $score_i$ is right-skewed, thereby mandating us to narrow down the scope of our defined "traditional" parents. We then construct a binary response indicator $y_i \in \{0, 1\}$, and then let $y_i = 1$ for "traditional" parents and $y_i = 0$ for "liberal" parents. The construction of such a binary response variable allows us to use the logit model to infer which variables are useful in predicting the log odds pertaining to one of the two parenting types.

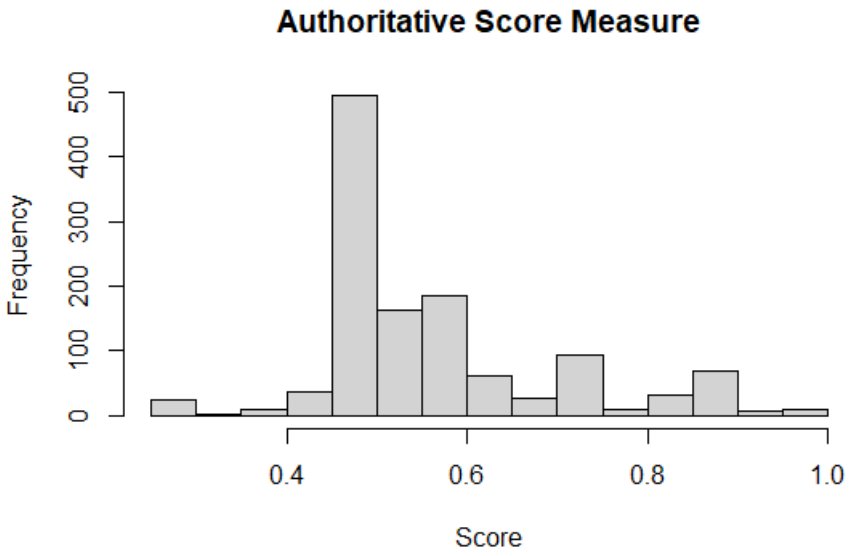

Figure 2: Sample Distribution of Authoritative Score Measure

3.2.2 Measures for Inequality

One of our major obstacles was to obtain a measure of income inequality at the county level. Good macro-level data at a province or county level in China is usually difficult to obtain. While there are publically available local level Gini coefficient estimates such as the ones in Bhattacharya, Palacio-Torralba, and Li (2018), they appear to be unreasonable and inaccurate at first glance with some local jurisdictions estimated at 0.99 while others at 0.06.

We thus turn to our own data and construct our own Gini coefficient estimates. Gini Coefficient is constructed using the personal income of the original data set, grouped by Hukou status, a good indicator of where the respondent legally resides. Figure 3 shows the distribution of the Gini coefficient constructed at the county level. This is still a vast overestimate as the lowest Gini coefficient we obtained is above 0.5. Nonetheless, we hope it will at least capture the rank of inequalities across different counties.

3.2.3 Modelling Framework

Our baseline methodology revolves around fitting a *binary logistic regression* model, commonly referred to as a binary logit model. The logit model is a way of formulating a structural relationship between the log odds – $log(\frac{p_i}{1-p_i})$ where p_i is the probability of some event – and relevant covariates. Such models are also referred to as *limited dependent variable* models,

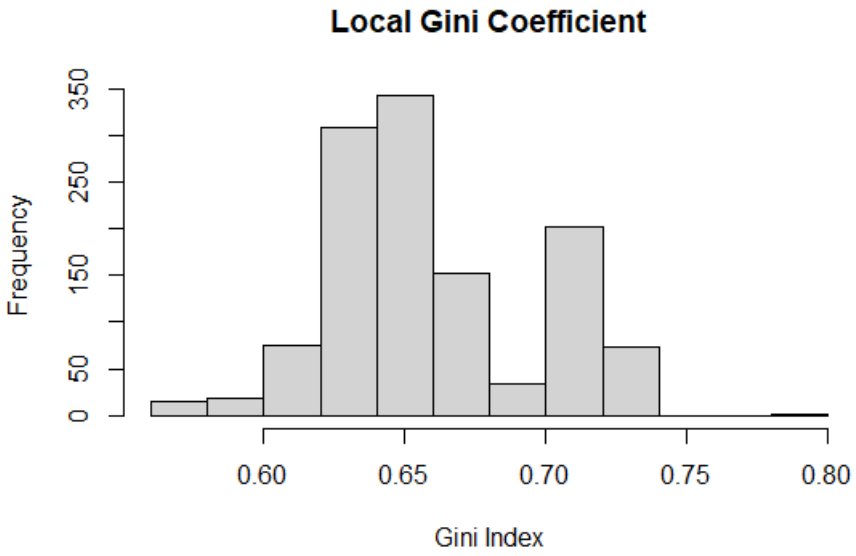

Figure 3: Distribution of Gini Coefficients by Province

where we assume decision-makers face a set of choices, and choose among them to maximize their utility. In our context, families make a binary decision, whether to be a "traditional" or a "liberal" parent. The logit model allows us to predict this binary outcome given the family characteristics. We estimate the model of the following form:

$$log\Big(\frac{p_i}{1-p_i}\Big) = X_i\beta + C_i\Gamma + \epsilon_i \tag{1}$$

X_i are variables of interest and C_i are the control variables. The presence of C_i is to take into account the heterogeneity present in the data. Traditional panel data methods to eliminate heterogeneity, such as fixed and random effects, cannot be applied to (1) due to the non-linear formulation of the logit model. Therefore, when deemed necessary, dummy variables are embedded in C_i to account for the possibility of family and regional heterogeneity.

We assume ϵ_i is i.i.d and follows Type 1 extreme value distribution, also known as the Gumbel distribution. Given our assumptions and $y_i = [0, 1]$ where $y_i = 1$ indicates parent i being a tiger parent, can form the joint density (2), and the coefficients are then estimated using maximum likelihood estimation.

$$f(y|\beta) = \prod_{i=1}^{n} [P(y_i = 1|\beta)]^{y_i} [P(y_i = 0|\beta)]^{1-y_i} \tag{2}$$

As an alternative to measuring the robustness of our results, we run a linear regression model

along with the logit model, where we regress the raw scores instead of the binary measure of parenting style.

$$score_i = X_i\beta + C_i\Gamma + u_i \tag{3}$$

However, we expect estimates from the regression to be small and not statistically significant. This is for several reasons. First, the distribution of $score_i$ is not well-centred. Moreover, our remaining sample after the merging process is quite small considering that some of the variables we adopted are county-level, and very likely are not a good representation of the local population distribution.

4 Results

Given all the problems aforementioned of our merged data set and Gini coefficients, we do not expect to obtain meaningful results regarding income inequality and parenting type. However, if our model is correctly specified, there is still hope to observe potential correlations between parenting types and parental/familial traits that we included as controls. We report our results in Table 2 below.

The first column of Table 2 presents estimates from the linear regression (3), where $score_i$ is regressed on the controls. The only statistically significant result occurs for Junior HS, an indicator for whether the parent has completed junior high, but the magnitude of the estimate is rather small. This echoes our prediction in the previous section that a linear regression with a continuous dependent variable may not be the appropriate choice of model.

We now turn to the logit model. Column 2 presents results without the *Gini coefficient* vector as the regressor, and column 3 shows results with the Gini coefficients. Not surprisingly, we obtained no significant relationship between the Gini coefficient and parenting type, but we were able to uncover some parental and familial characteristics that predict parenting type with statistical significance. These characteristics are listed in Table 2.

For all regressors presented in Table 2 excluding the Gini coefficient and the constant term, we found their estimates to be practically the same with or without the Gini coefficient added to the model, in terms of both point estimates and level of significance. We found a positive, but very small, relationship between family net income and the log probability of being a "traditional"

parent. Such findings are interesting, considering that one would have thought parents to be more "liberal" as their income goes up. However, given that the estimates are so small, it is not clear whether we can gauge any meaningful implications from this positive, albeit statistically significant, estimate.

What is perhaps the most important finding of this paper is that parental level of education and family social status in the local area possess strong predictive power on the likelihood of a parent being a "traditional". *Primary School* and *Junior High* are indicators for whether the parent has completed primary school and junior high school education, respectively. We found that parents who have completed each level of education are much less likely to become "traditional". Completing a primary school education drops the log-likelihood of becoming a "traditional" parent by 0.68, while the completion of junior high education decreases that log-likelihood by 0.63. Similarly, we also found that in families that are of a higher social status in the local area, parents are less likely to be "traditional". Since *social status* is a discrete variable with responses ranging from 1 to 5 (low to high), the results can be interpreted as a one-level increase in family local social status decreasing the log-likelihood of being a "traditional" parent by 0.26.

5 Concluding Remarks

One issue that we encountered is that most questions that we used to construct the parental type measure have a rather low response rate, and as a result, our final sample may be too small in size and not representative enough of the true population distribution. This, coupled with the low-quality estimate of local income inequality, may very well be the reason for our disappointing results regarding income inequality and parental preferences. In spite of this, our strong and significant estimates of the predictive power of education level and family social status on parental types can serve as a great first step in discovering the forces at play in shaping parental preferences.

Our paper could be extended by further work in two dimensions. On the one hand, we believe that there could still be a strong relationship between income inequality and parenting style. We call for further research with higher-quality estimates of income inequality and more refined field surveys. On the other hand, our paper could also lead to further validation of how much certain parental and familial traits could predict parenting preferences, and what could be the mechanism at play here.

Table 2:

	Dependent variable:		
	Score	Authoritative	
	OLS	*Logistic*	
	(1)	(2)	(3)
Gini Coefficient			−1.436
			(3.344)
# Children	0.007	−0.160	−0.156
	(0.007)	(0.202)	(0.202)
Family Net Income	−0.000	0.00001*	0.00001*
	(0.00000)	(0.00000)	(0.00000)
Primary School	−0.013	−0.666*	−0.681*
	(0.013)	(0.371)	(0.373)
Junior HS	−0.029**	−0.625*	−0.634*
	(0.013)	(0.359)	(0.359)
Social Status	−0.006	−0.264*	−0.263*
	(0.005)	(0.145)	(0.145)
School Distance	0.0003	0.010*	0.011*
	(0.0002)	(0.006)	(0.006)
Constant	0.571***	−1.764	−0.773
	(0.069)	(1.573)	(2.792)
Observations	1,221	1,221	1,221
R^2	0.028		
Adjusted R^2	0.006		
Log Likelihood		−293.549	−293.456
Akaike Inf. Crit.		643.097	644.912
Residual Std. Error	0.133 (df = 1193)		
F Statistic	1.264 (df = 27; 1193)		
Note:			*p<0.1; **p<0.05; ***p<0.01

References

Agostinelli, F., Doepke, M., Sorrenti, G., & Zilibotti, F. (2020). It takes a village: The economics of parenting with neighborhood and peer effects.
DOI: 10.3386/w27050

Bhattacharya, P., Palacio-Torralba, J., & Li, X. (2018). On income inequality within china's provinces. *Chinese Studies*, *07*(02), 174–182. DOI: 10.4236/chnstd.2018.72015

Burton, P., Phipps, S., & Curtis, L. (2002). All in the family: A simultaneous model of parenting style and child conduct. *American Economic Review*, *92*(2), 368–372. DOI: 10.1257/000282802320191633

Chua, A. (2014). *Battle hymn of the tiger mother*. Bloomsbury Publishing.

Chung, A., Xiong, Y., & Zhang, J. (2022). An economic analysis of tiger parenting: Evidence from child developmental delay or learning disability. *Labour Economics*, *78*, 102226. DOI: 10.1016/j.labeco.2022.102226

Deng, L., & Tong, T. (2020). Parenting style and the development of noncognitive ability in children. *China Economic Review*, *62*, 101477. DOI: 10.1016/j.chieco.2020.101477

Doepke, M., Sorrenti, G., & Zilibotti, F. (2019). The economics of parenting. *Annual Review of Economics*, *11*(1), 55–84. DOI: 10.1146/annurev-economics-080218-030156

Doepke, M., & Zilibotti, F. (2017). Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, *85*(5), 1331–1371. DOI: 10.3982/ecta14634

Guryan, J., Hurst, E., & Kearney, M. S. (2008). Parental education and parental time with children.
DOI: 10.3386/w13993

Hart, B., & Risley, T. R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology*, *28*(6), 1096–1105. DOI: 10.1037/0012-1649.28.6.1096

Lundberg, S. (2015). Tiger parenting and american inequality: An essay on chua and rubenfeld's the triple package: How three unlikely traits explain the rise and fall of cultural groups in america. *Journal of Economic Literature*, *53*(4), 945–960. DOI: 10.1257/jel.53.4.945

Spera, C. (2005). A review of the relationship among parenting practices, parenting styles, and adolescent school achievement. *Educational Psychology Review*, *17*(2), 125–146. DOI: 10.1007/s10648-005-3950-1